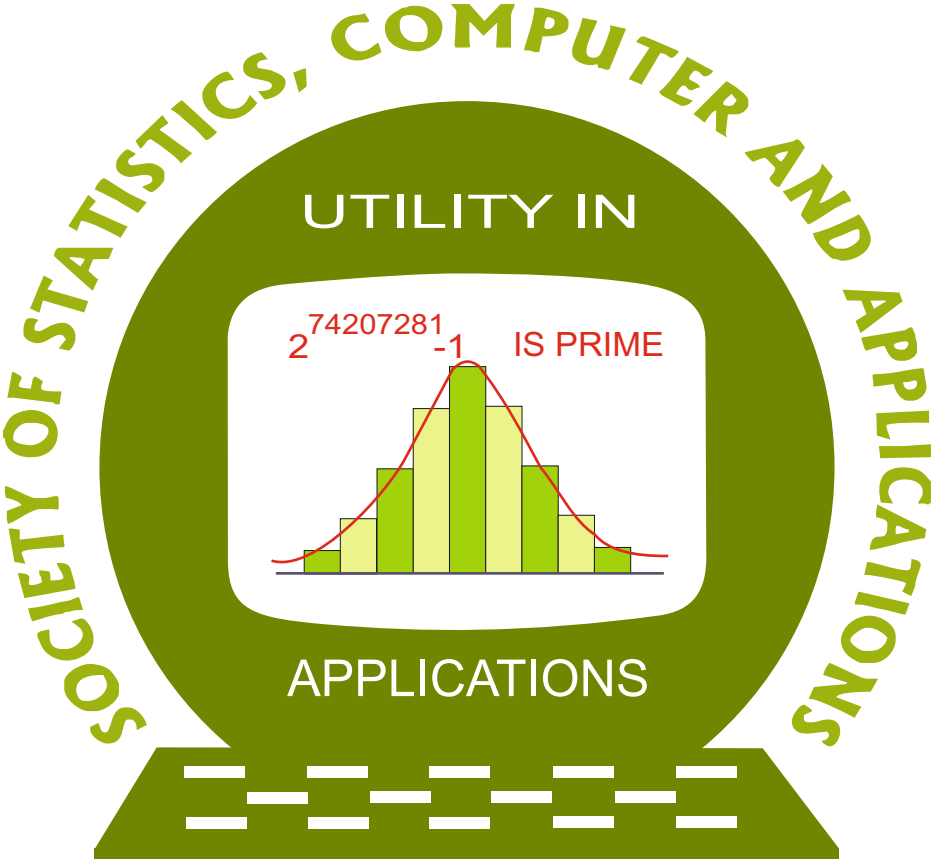
ISSN 2454-7395(online)

STATISTICS AND APPLICATIONS



FOUNDED 1998

Journal of the Society of Statistics, Computer and Applications
https://ssca.org.in/journal.html
Vol. 19, No. 1 (Special Issue), 2021 (New Series)

Society of Statistics, Computer and Applications

Council and Office Bearers

Founder President
Late M.N. Das

President
V.K. Gupta

Executive President
Rajender Parsad

Patrons

A.C. Kulshreshtha
A.K. Nigam
Bikas Kumar Sinha
D.K. Ghosh
K.J.S. Satyasai
P.P. Yadav
Pankaj Mittal
R.B. Barman
R.C. Agrawal
Rahul Mukerjee
Rajpal Singh

Vice Presidents

A. Dhandapani
Lal Mohan Bhar
P. Venkatesan
Ramana V. Davuluri
S.D. Sharma
V.K. Bhatia

Secretary
D. Roy Choudhury

Foreign Secretary
Sudhir Gupta

Treasurer
Ashish Das

Joint Secretaries

Aloke Lahiri
Shibani Roy Choudhury

Council Members

Alka Arora
Anil Kumar Yadav
Anshu Bhardwaj
B. Re. Victor Babu
Manish Sharma
Manisha Pal
Piyush Kant Rai
Praggya Das
Ranjit Kumar Paul
Raosaheb V. Latpate
S.A. Mir
Sanjeev Panwar
V. Srinivasa Rao
V.M. Chacko
Vishal Deo

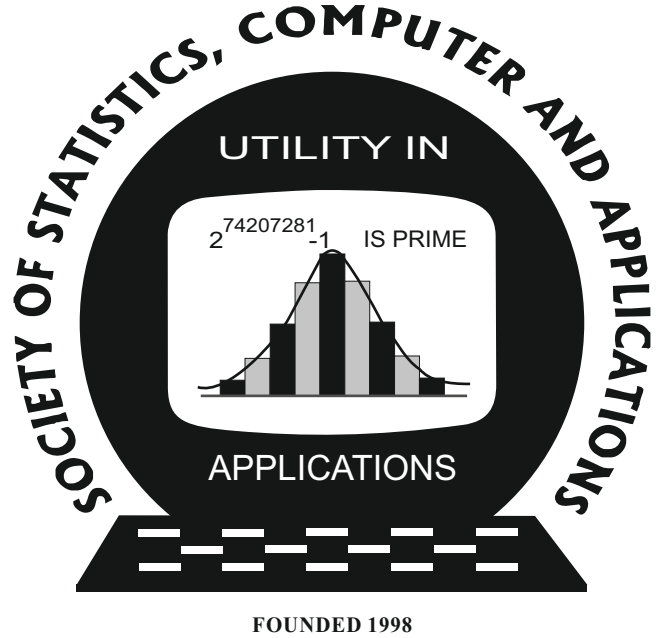
Ex-Officio Members (By Designation)

Director General, Central Statistics Office, Government of India, New Delhi
Director, ICAR-Indian Agricultural Statistics Research Institute, New Delhi
Chair Editor, Statistics and Applications
Executive Editor, Statistics and Applications

Society of Statistics, Computer and Applications, B - 133, Ground Floor, C.R. Park, New Delhi - 110019

Statistics and Applications

ISSN 2454-7395(online)



Journal of the Society of Statistics, Computer and Applications

https://ssca.org.in/journal.html

Volume 19, No. 1 (Special Issue), 2021 (New Series)

Special Issue on "Contributions to Combinatorics, Design of Experiments, Linear Algebra and Related Areas" in the memory of Late Professor Aloke Dey who left for his heavenly abode on 10 February 2020

Statistics and Applications

Volume 19, 1 (Special Issue), 2021 (New Series)

Editorial Panel

CONTENTS

Statistics and Applications ISSN 2454-7395 (online)
Volume 19, No. 1 (New Series) 2021

Special Volume in Memory of Professor Aloke Dey

PREFACE

The Society of Statistics, Computer and Applications (SSCA) was founded in 1998 to honour great legendary Professor M.N. Das. Since then, the SSCA has been organizing National / International Conferences every year along the length and breadth of the country. It has organized thus far 23 conferences. The SSCA, among other scientific activities, also brings out this journal called *Statistics and Applications*. This is an open access journal and readers can view full paper online, download it, save it and print it without any cost. The journal publishes research papers after a stringent review process. The journal is available at https://ssca.org.in/journal.html.

This Issue (No. 1) of the Volume 19 of *Statistics and Applications* has been brought out as a tribute to Professor Aloke Dey, a world leader in the area of designs of experiments and combinatorics, who at the age of 74 years and 7 months, left for his heavenly abode after a brief illness, on 10 February 2020. Incidentally, Late Professor Dey happened to be a Ph.D. student of Late Professor Das.

The Editorial Board of the journal decided to have this special issue in memory of Late Professor Aloke Dey and titled it as **Contributions to Combinatorics, Design of Experiments, Linear Algebra and Related Areas**. The Editorial Board also invited Vinod Kumar Gupta, Sudhir Gupta, Rajender Parsad and Ashish Das to act as the Guest Editors of this special issue. Vinod Gupta shared the key responsibility of coordinating with the authors, the reviewers and the other guest editors. The authors invited to contribute to this issue were essentially Professor Dey's collaborators, his students, friends and colleagues around the globe. It gives us unbounded pleasure to mention here that the response obtained was overwhelming.

Professor Dey was a legendary statistician and his contributions to the advancement of statistical sciences have been colossal. Although he was world renowned for his contributions in the area of design of experiments, Professor Dey had also researched in other diverse areas like survey sampling, combinatorial theory, linear algebra and cryptology. Spread over half a century, his fundamental research, reflecting an amazing versatility, depth and originality immensely enriched the theory and applications of statistics. He also authored 5 books on block designs, fractional factorial designs and cross over designs. Professor Dey not only had been a great researcher, he had also been a brilliant teacher and mentor for many young upcoming researchers. With his strong desire to conduct quality basic research, he influenced the thoughts of many young researchers and inspired them along the journey of their research. With his immense love, affection and willingness to always help, he made an unforgettable mark on the lives of many of his colleagues, students and young researchers. It is a matter of great pride for all the guest editors of this issue to add here that they were also mentored by Late Professor Dey. The research journey of all the four guest editors was also influenced by the guidance of Late Professor Dey. In his demise the statisticians' fraternity has lost a brilliant, celebrated and globally recognized statistician. Professor Dey was a pious soul and a true humane - a gem.

Professor Aloke Dey was fellow of the Indian National Science Academy (INSA), the National Academy of Sciences, India (NASI), and an Elected Member of the International Statistical Institute. He served as an editor of Sankhya, the Indian Journal of Statistics, during 2002–2005, and was also the chair editor of Statistics and Applications, during 2009–2020.

Every year, on June 29th, India celebrates the birth anniversary of Late Professor Prasanta Chandra Mahalanobis, a recognized global leader in the discipline of statistics, as National Statistics Day. On the National Statistics Day in 2010, for his lifetime contributions in the field of Statistics, Late Professor Dey was conferred the Government of India's prestigious National Award in Statistics that had been instituted in honour of Professor P.V. Sukhatme.

This memorial issue has the tributes to this stalwart Professor Dey in the form of 31 research papers from authors across the globe. These papers cover a broad spectrum of the research interests of Professor Dey. Additionally, there is an article on Professor Aloke Dey - A tribute by Vinod Kumar Gupta, Rahul Mukerjee, Mausumi Bose and Arun Nigam. The guest editors feel honoured to have been able to produce an issue that is a befitting homage to their GURU, a tall statistician with immense contributions.

We would like to sincerely express our gratefulness to all the authors for their contribution to this memorial issue. The reviewers, an unobservable layer without whom the process of journal publication cannot function at all, have also been prompt and thorough. Their suggestions helped in improving the quality and presentation of the contents. We are indebted to all the reviewers and thank them sincerely for their support. We would like to place on record our highest admiration for the Executive Council of SSCA and the Editorial Board of Statistics and Applications for their support and for entrusting their faith on the Guest Editors for bringing out this special issue as a tribute to Professor Dey. The Guest Editors, in turn, are greatly honoured by this responsibility. The help received from Dr. B.N. Mandal, Managing Editor for bringing the papers in the format of the journal is highly appreciated.

This issue contains papers of high academic standards covering a wide spectrum of statistical research. We are confident that the readers would find these papers enjoyable and a resource for generating newer ideas for advancing research in statistical sciences. This issue is our endeavour to pay rich tributes to this giant of statistical sciences with a towering stature, filled with traits of humanity like gentleness, kindness, humbleness and gratitude.

Vinod Kumar Gupta
Sudhir Gupta
Rajender Parsad
Ashish Das

May 2021

Statistics and Applications {ISSN 2454-7395(online)}
Volume 19, No. 1, 2021 (New Series), pp 1-10

Tribute to Professor Aloke Dey – A Noble and Pious Soul

(12 September 1945 - 10 February 2020)

Vinod Kumar Gupta[1], Rahul Mukerjee[2], Mausumi Bose[3] and Arun K. Nigam[4]
[1]*Formerly National Professor at IASRI, New Dehi*
[2]*Professor, Indian Institute of Management, Kolkata*
[3]*Professor, Indian Statistical Institute, Kolkata*
[4]*Formerly Professor at IASRI, New Delhi*

Received: 02 December 2020; Revised: 20 January 2021; Accepted: 04 February 2021

A Hebrew Proverb (Translation by Rabindranath Tagore)

A death is not the extinguishing of a light,
but the putting out of the lamp
because the dawn has come.

Life is given to us,
we earn it by giving it.
Let the dead have the immortality of fame,
but the living the immortality of love.

Life's errors cry for the merciful beauty
that can modulate their isolation into a
harmony with the whole.

Life, like a child, laughs,
shaking its rattle of death as it runs.

1. An Eulogy

We are all united not only in our desire to pay our homage to Aloke Dey but rather in our need to do so. For such was his extraordinary appeal in the community of statisticians across the globe that all of us feel that we have lost a soul mate of ours.

Aloke Dey was the very essence of wisdom, of dedication, of duty, of style, of sincerity, of academic honesty, of humbleness, of compassion, of friendship, of care. His associates had always been charmed by not only his knowledge but also his rectitude and integrity. And talking of his dedication, even with problems in his eyes, he used to spend hours and hours of his time working on computers and writing manuscripts.

Aloke Dey has left for his heavenly abode, and in the entire academic world, there is a feeling of having been left desolate and forlorn. All of us sense that feeling, and we do not know when we shall be able to overcome it. And, at the same time, there is a proud thankfulness to God for allowing

Corresponding Author: V.K. Gupta
Email: vkgupta_1751@yahoo.co.in

us of this generation to be associated with this towering personality who not only made monumental contributions towards the advancement of statistical sciences, but also remained, at the same time, so down to earth and so compassionate. While his jovial nature endeared him to everyone, he carried an aura reflecting the deepest respect that he earned from people around him for his unbounded talent, his perfect intuition and his sharp memory. Indeed, even an apparently simple problem posed by him often had deep underpinnings, leading to the excitement of discovery. He was truly a friend, philosopher and guide whose very presence filled the minds of his associates with joy of learning, confidence and happiness.

Today is our chance to say thank you, Aloke Dey, for the way you brightened science and our lives, even though God granted you but a short life and took you away so early. Only now that you are gone do we truly appreciate what we are now without and we want you to know that life without you is very, very difficult. It is only through the strength of the message that you gave us over the years that we are slowly gaining the strength to move forward. The days that we spent together and what we had learned from you will remain in our hearts as our most cherished treasure.

2. A Birds Eye View of the Accolades

Professor Dey's attainments are so towering and encompass such a multitude of directions that we find it really hard to present them here in adequate detail. Yet, we make a valiant effort to highlight just a few of these.

Professor Aloke Dey…

[A brilliant scholar]

- Master's degree (1964) in mathematics, then master's degree (1966) from the Indian Agricultural Statistics Research Institute (IASRI) with a first rank, followed by a Ph.D. degree (1969), from IASRI.

[Official positions]

- Joined IASRI in 1970 as a faculty; became a senior professor in 1977 at the early age of 32 years.
- Joined the Indian Statistical Institute (ISI) in 1989 as a professor and continued there till his formal retirement in 2007.
- Senior Scientist of the Indian National Science Academy (2007-12) and the National Academy of Sciences, India (2012-17), both hosted by ISI.
- Held senior academic positions at many institutions abroad, including those in USA, Canada and Taiwan.

[A researcher par excellence]

- A world leader in statistics for fundamental research in diverse areas, for excellent dissemination of ideas through elegantly written books, and for influential editorial work.
- Research, spread over five decades, encompassing multiple areas pertaining to not only statistics but also mathematics, and reflecting an amazing depth and versatility.
- Extensive publications, with numerous citations, in the very best journals – no wonder though, given the profound depth of his findings.
- Areas particularly enriched through his work include design of experiments, survey sampling, combinatorial theory, linear algebra and cryptology.

- [*Research in design of experiments*]
  Here alone, path-breaking contributions to such diverse areas as factorial designs, varietal block and row-column designs, weighing designs, response surface designs, crossover designs, designs for biological assays and diallel crosses, and so on.
  Specifically:
  (A) Research on orthogonal fractional factorial plans and related orthogonal arrays, with emphasis on the practically important but mathematically difficult asymmetric case, blending theoretical elegance with immediate applicability, notably in industrial experimentation and quality control. This includes, in particular:
    (i) Work on the hard problem of obtaining fractional factorial plans when certain interactions are important, coming up with an ingenious solution *via* the use of tools from finite projective geometry.
    (ii) Deep results on optimal main effect plans under nonorthogonal blocking, opening up a whole new area.
  (B) High impact results, in both statistics and combinatorics, on other topics of experimental design, such as
    (i) a new class of incomplete block designs with nested structure,
    (ii) universal optimality and nonoptimality of certain row-column designs (well-known for counterintuitive findings),
    (iii) optimal designs for biological assays and diallel crosses, as well as optimal weighing designs (now classics in the respective fields),
    (iv) crossover designs (including a recent authoritative review).
- [*Research in other areas*]
  Very remarkable contributions to many other areas such as
    (i) unequal probability sampling plans,
    (ii) characterization problems *via* conditional expectations,
    (iii) tactical configurations,
    (iv) diagonally range dominant matrices,
    (v) efficient key pre-distribution schemes for distributed sensor networks,
    and so on.
  These include elegant statistical proofs of several results in matrix algebra.
- [*Books*]
  All real gems that received many accolades from the statistical community; all from major international publishers.

[Honours and awards]

- Fellow of the Indian National Science Academy (INSA) and the National Academy of Sciences, India (NASI).
- Honoured with the prestigious *Professor P.V. Sukhatme National Award in Statistics* (2010), by the Ministry of Statistics and Programme Implementation, Government of India, for lifetime contributions to the field of statistics.
- Elected Member of the International Statistical Institute.

[An editor of eminence]

- Editor, Sankhya, the Indian Journal of Statistics (2002-05); under his eminent leadership and through his painstaking efforts, the journal attained new heights.
- Chair Editor, Statistics and Applications (2009-2020); under his research administrative capability, the journal witnessed a boost in its stature and started becoming visible globally.

[A great teacher]

- While being a researcher par excellence, always mindful of his responsibilities as a teacher; successive generations of students benefited themselves under the tutelage of the great teacher in him.
- Co-author of an INSA sponsored book *Understanding Mathematics* that aimed at the promotion of mathematics among senior school students and first year college students.

[A great mentor]

- Supervised more than 15 Ph.D. students and inspired them to reach their full potential.
- Over the years, also acted generously as a mentor to many other statisticians apart from his direct PhD students; they all benefited academically from his counsel at various stages of their careers.

[A great friend]

- While being a celebrated teacher and an inspiring mentor, was also an extremely caring human being and a wonderful friend.
- A scintillating conversationalist who took an active interest in many areas beyond academics; many statisticians can vouch for the fact that, besides being interested in their academic affairs, he was also concerned about their overall well-being.
- Often, the professional association developed into a much closer bond where he became like a family member whom it was a joy to spend time with.

3. Research Publications

The arrangement is chronological so as to reflect the research interests of Professor Dey over the years. Within each year, the arrangement is alphabetical according to the authors' surnames.

3.1. Books Published

1. A. Dey (1985). *Orthogonal Fractional Factorial Designs*. John Wiley.
2. A. Dey (1986). *Theory of Block Designs*. John Wiley/ Halsted Press.
3. A. Dey and R. Mukerjee (1999). *Fractional Factorial Plans*. John Wiley.
4. M. Bose and A. Dey (2009). *Optimal Crossover Designs*. World Scientific.
5. A. Dey (2010). *Incomplete Block Designs*. Hindustan Book Agency/ World Scientific.

3.2. Research Papers Published

1. M. N. Das and A. Dey (1967). Group divisible rotatable designs. *Annals of the Institute of Statistical Mathematics* 1**9**, 337–347; corrections *ibid* (1968), **20**, 337.
2. A. Dey (1968). On response surface designs with equispaced doses. *Calcutta Statistical Association Bulletin*, **19**, 135–143.
3. A. Dey and A. K. Nigam (1968). Group divisible rotatable designs. Some further considerations. *Annals of the Institute of Statistical Mathematics*, **20**, 477–481.
4. A. Dey (1970). On construction of balanced *n*-ary block designs. *Annals of the Institute of Statistical Mathematics*, **22**, 389–393.
5. A. Dey and M. N. Das (1970). On blocking second order rotatable designs. *Calcutta Statistical Association Bulletin*, **17**, 75–85.

6. A. Dey and G. M. Saha (1970). Main effect plans for n^k factorials with blocks. *Annals of the Institute of Statistical Mathematics*, **22**, 381–388.
7. A. C. Kulshreshtha and A. Dey (1970). A new weighing design. *Australian Journal of Statistics*, **12**, 166–168.
8. A. K. Nigam and A. Dey (1970). Four and six level second order rotatable designs. *Calcutta Statistical Association Bulletin*, **19**,155–167.
9. A. C. Kulshreshtha, G. M. Saha and A. Dey (1971). On circular designs. *Annals of the Institute of Statistical Mathematics*, **23**, 491–497.
10. A. Dey, A. C. Kulshreshtha and G. M. Saha (1972). Three symbol partially balanced arrays. *Annals of the Institute of Statistical Mathematics*, **24**, 525–528.
11. A. C. Kulshreshtha, A. Dey and G.M. Saha (1972) Balanced designs with unequal replications and unequal block sizes. *Annals of Mathematical Statistics*, **43**, 1342–1345.
12. G. M. Saha and A. Dey (1973a). On construction and uses of balanced *n*-ary designs. *Annals of the Institute of Statistical Mathematics*, **25**, 439–445.
13. G. M. Saha, A. C. Kulshreshtha and A. Dey (1973b). On a new type of *m*-class cyclic association scheme and designs based on the scheme. *Annals of Statistics*, **1**, 985–990.
14. A. Dey and C. K. Midha (1974). On a class of PBIB designs. *Sankhyā*, **B36**, 320–322.
15. A. Dey and G. M. Saha (1974). An inequality for tactical configurations. *Annals of the Institute of Statistical Mathematics*, **26**, 171–173
16. A. K. Banerjee, A. Dey and G. M. Saha (1975). Some main effect plans for 3^n factorials. *Annals of the Institute of Statistical Mathematics*, **27**, 159–165.
17. A. Dey (1975). A note on balanced designs. *Sankhyā*, **B37**, 461–462.
18. T. K. Gupta and A. Dey (1975). On some new second order rotatable designs. *Annals of the Institute of Statistical Mathematics*, **27**, 167–175.
19. R. Chakravarty and A. Dey (1976). On the construction of balanced and orthogonal arrays. *Canadian Journal of Statistics*, **4**, 109–117.
20. A. Dey and C. K. Midha (1976). Generalised balanced matrices and their applications. *Utilitas Mathematica*, **10**, 139–149.
21. R. Gopalan and A. Dey (1976). On robust experimental designs. *Sankhya*, **B38**, 297–299.
22. A. Dey (1977). Construction of regular group divisible designs. *Biometrika*, **64**, 647–649.
23. A. Dey and S. C. Gupta (1977). Singular weighing designs and estimation of total weight. *Communications in Statistics: Theory and Methods*, **A6**, 289–295.
24. A. Dey and G. V. S. Ramakrishna (1977). A note on orthogonal main-effect plans. *Technometrics*, **19**, 511–512.
25. A. Chacko and A. Dey (1978). On the estimation of total weight in chemical balance weighing designs. *Australian Journal of Statistics*, **20**, 83–86.
26. M. Singh and A. Dey (1978). Two-way elimination of heterogeneity. *Journal of the Royal Statistical Society*, **B40**, 58–63.
27. A. Chacko and A. Dey (1979). Weighing designs optimum for the estimation of total weight. *Sankhyā*, **B41**, 270–276.
28. M. Singh and A. Dey (1979a). On analysis of some augmented block designs. *Biometrical Journal*, **21**, 87–92.
29. M. Singh and A. Dey (1979b). Block designs with nested rows and columns. *Biometrika*, **66**, 321–326.
30. M. Singh, A. Dey and A. K. Nigam (1978). Two-way elimination of heterogeneity. II. *Sankhyā*, **B40**, 227–235.
31. A. Dey and M. Singh (1980). Some series of efficiency balanced designs. *Australian Journal of Statistics*, **22**, 364–367.

32. K. Win and A. Dey (1980). Incomplete block designs for parallel-line assays. *Biometrics*, **36**, 487–492.
33. A. Dey, M. Singh and G. M. Saha (1981). Efficiency balanced block designs. *Communications in Statistics: Theory and Methods*, **A10**, 237–247.
34. S. K. Agarwal, P. Kumar and A. Dey (1982). On unequal probability sampling of two units without replacement. *Journal of the Royal Statistical Society*, **B44**, 43–46.
35. V. Agrawal and A. Dey (1982). A note on orthogonal main effect plans for asymmetrical factorials. *Sankhyā*, **B44**, 278–282.
36. V. K. Gupta, A. K. Nigam and A. Dey (1982). Orthogonal main effect plans for asymmetrical factorials. *Technometrics*, **24**, 135–137.
37. K. Sinha and A. Dey (1982). On resolvable PBIB designs. *Journal of the Statistical Planning and Inference*, **6**, 179–181.
38. A. Dey, V. K. Gupta and M. Singh (1983). Optimal change over designs. *Sankhya*, **B45**, 223–239.
39. K. Sinha and A. Dey (1983). A series of truly self-dual PBIB designs. *Mathematische Operationsforschung und Statistik Series Statistics*, **14**, 53–54.
40. A. Dey and V. Agrawal (1985). Orthogonal fractional plans for asymmetrical factorials derivable from orthogonal arrays. *Sankhyā*, **B47**, 56–66.
41. A. Dey and A.K. Nigam (1985). Construction of group divisible designs. *Journal of the Indian Society of Agricultural Statistics*, **37**, 163–166.
42. A. Dey, U. S. Das and A. K. Banerjee (1986). Construction of nested balanced incomplete block designs. *Calcutta Statistical Association Bulletin*, **35**, 161–167.
43. A. Dey and V. K. Gupta (1986). Another look at the efficiency and partially efficiency balanced designs. *Sankhya*, **B48**, 437–438.
44. M. Singh and A. Dey (1987). A note on incomplete block designs for symmetrical parallel line assays. *Communications in Statistics: Theory and Methods*, **16**, 3555–3564.
45. A. Dey (1988). Some new partially balanced designs with two associate classes. *Sankhyā*, **B50**, 90–94.
46. A. Dey and S. P. Dhall (1988). Robustness of augmented BIB designs. *Sankhya,* **B50**, 376–381.
47. A. Das and A. Dey (1989a). A generalisation of distinct representatives and its applications. *Calcutta Statistical Association Bulletin*, **38**, 57–63.
48. A. Das and A. Dey (1989b). A note on balanced block designs. *Journal of Statistical Planning and Inference*, **22**, 265–268.
49. A. Dey and A. Das (1989). On some E-optimal block designs. *Metrika*, **36**, 269–278.
50. N. K. Nguyen and A. Dey (1989). Computer aided-construction of D-optimal 2^m fractional factorial designs of resolution V. *Australian Journal of Statistics*, **31**, 111–117.
51. A. Das and A. Dey (1990a). Optimality of row-column designs. *Calcutta Statistical Association Bulletin*, **39**, 63–72.
52. A. Das and A. Dey (1990b). A note on construction of Graeco-Latin square of order $2n+1$. *Journal of the Indian Society of Agricultural Statistics*, **42**, 247–249.
53. N. K. Nguyen and A. Dey (1990). Computer-aided construction of small (M,S)-optimal incomplete block designs. *Australian Journal of Statistics*, **32**, 399–410.
54. R. Srivastava, V. K. Gupta and A. Dey (1990). Robustness of some designs against missing observations. *Communication in Statistics: Theory and Methods*, **19**, 121–126.
55. R. B. Bapat and A. Dey (1991). A-optimal block designs with minimal number of observations. *Statistics and Probability Letters*, **11**, 399–402.

56. A. Das and A. Dey (1991). Optimal variance- and efficiency-balanced designs for one- and two-way elimination of heterogeneity. *Metrika*, **38**, 227–238.
57. A. Dey, and K. Balasubramanian (1991). Construction of some families of group divisible designs. *Utilitas Mathematica*, **40**, 283–290.
58. A. Dey, R. Srivastava and V. K. Gupta (1991). Robust designs - A review and bibliography**.** *Cahiers du* CERO, **33**, 51**–**62.
59. V. K. Gupta, A. Das and A. Dey (1991). Universal optimality of block designs with unequal block sizes. *Statistics and Probability* Letters, **11***,* 177**–**180.
60. R. Srivastava, V. K. Gupta and A. Dey (1991). Robustness of some designs against missing data. *Journal of Applied Statistics*, **18**, 303**–**308.
61. K. Balasubramanian, A. Dey and P. Bhimasankaram (1992). Diagonally range dominant matrices. *Linear Algebra and Its Applications*, **176**, 45–60.
62. A. Das and A. Dey (1992). Universal optimality and nonoptimality of some row-column designs. *Journal of Statistical Planning and Inference*, **31**, 263–271.
63. A. Dey (1993a) Some orthogonal arrays with variable symbols. *Journal of Combinatorics and Information and System Sciences*, **18**, 209–215.
64. A. Dey (1993b). Robustness of block designs against missing data. *Statistica Sinica*, **3**, 219–231.
65. A. Dey, S. Hande and M. L. Tiku (1994). Statistical proofs of some matrix results. *Linear and Multilinear Algebra*, **38**, 109–116.
66. A. Dey, G. C. Chawla, and G. Balachandran (1995). Cyclic change-over designs. *Journal of the Indian Statistical Association*, **33**, 71–76.
67. A. Dey, K. R. Shah and A. Das (1995). Optimal block designs with minimal and nearly minimal number of units. *Statistica Sinica*, **5**, 547**–**558.
68. C. K. Midha and A. Dey (1995). Cyclic group divisible designs. *Calcutta Statistical Association Bulletin*, **45**, 253–257.
69. K. Balasubramanian and A. Dey (1996). D-optimal designs with minimal and nearly minimal number of units. *Journal of the Statistical Planning and Inference*, **52**, 255–262.
70. A. Dey and C. K. Midha (1996a). Construction of some asymmetrical orthogonal arrays. *Statistics and Probability Letters*, **28**, 211–217.
71. A. Dey and C. K. Midha (1996b). Optimal block designs for diallel crosses. *Biometrika*, **83**, 484–489.
72. A. Dey, C. K. Midha, and D. C. Buchthal (1996). Efficiency of the residual design under the loss of observations in a block. *Journal of the Indian Society of Agricultural Statistics*, **49**, 237–248.
73. C. K. Midha and A. Dey (1996). A note on resolvable incomplete block designs. *Journal of the Indian Statistical Association*, **34**, 81–84.
74. K. Balasubramanian and A. Dey (1997). Distributions characterized through conditional expectations. *Metrika*, **45**, 189–196.
75. G. C. Chawla and A. Dey (1998). A series of balanced ternary change over designs. *Journal of the Indian Society of Agricultural Statistics*, **51**, 42–50.
76. A. Das, A. Dey and A. M. Dean (1998). Optimal designs for diallel cross experiments. *Statistics and Probability Letters*, **36**, 427–436.
77. A. Dey and C. K. Midha (1998). Addition or deletion? *Statistics and Probability Letters*, **37**, 409–414.
78. A. Dey and R. Mukerjee (1998). Techniques for constructing asymmetric orthogonal arrays. *Journal of Combinatorics, Information and System Sciences* (Professor J.N. Srivastava Dedication Volume), **23**, 351–366.

79. A. Dey and K. R. Shah (1998). Optimal asymmetric fractional factorial plans for estimating main effects and selected two-factor interactions. *Statistical Theory and Applications. International Journal of Mathematical and Statistical Sciences*, **7**, 27–38.
80. G. M. Saha, A. Dey and C. K. Midha (1998). Construction of nested incomplete block designs. *Calcutta Statistical Association Bulletin*, **48**, 195–205.
81. A. Dey, K. Balasubramanian and S. Gupta (1999). Incomplete block designs for slope ratio assays. First NIU Statistical Sciences Symposium (De Kalb, IL, 1996). *Journal of Statistical Planning and Inference*, **78**, 369–383.
82. A. Dey and R. Mukerjee (1999). Inter-effect orthogonality and optimality in hierarchical models. *Sankhya*, **B61**, 460-468.
83. A. Das, A. Dey and S. Gupta (2000). A-efficient block designs for slope ratio assays. *Calcutta Statistical Association Bulletin*, **50**, 255–263.
84. F. S. Chai, A. Das and A. Dey (2001). A-optimal block designs for parallel line assays. *Journal of Statistical Planning and Inference*, **96**, 403–414.
85. A. Dey and C. Y. Suen (2001). Further asymmetric orthogonal arrays. *Statistics and Applications*, **3**, 61–64.
86. C. Y. Suen, A. Das and A. Dey (2001). On the construction of asymmetrical orthogonal arrays. *Statistica Sinica*, **11**, 241–260.
87. K. Chatterjee, A. Das and A. Dey (2002). Quasi-orthogonal arrays and optimal fractional factorial plans. *Statistica Sinica*, **12**, 905–916.
88. A. Dey (2002). Optimal designs for diallel crosses. *Journal of the Indian Society of Agricultural Statistics*, **55**, 1–16.
89. A. Dey and C. Y. Suen (2002). Optimal fractional factorial plans for main effects and specified two-factor interactions: a projective geometric approach. *Annals of Statistics*, **30**, 1512–1523.
90. R. Mukerjee, A. Dey and K. Chatterjee (2002). Optimal main effect plans with nonorthogonal blocking. *Biometrika*, **89**, 225–229.
91. M. Bose and A. Dey (2003). Some small and efficient cross-over designs under a nonadditive model. *Utilitas Mathematica*, **63**, 173–182.
92. F. S. Chai, A. Das and A. Dey (2003). Block designs for symmetric parallel line assays with block size odd. *Sankhyā*, **65**, 689–703.
93. A. Das, A. Dey and C. K. Midha (2003). On a property of orthogonal arrays and optimal blocking of fractional factorial plans. *Metrika*, **57**, 127–135.
94. A. Dey and R. Mukerjee (2003). *Symmetrical Factorial Experiments: A Mathematical Theory* (a tribute to R.C. Bose). In R. Bhatia Ed. Connected at Infinity, Hindustan Book Agency, New Delhi, 3–17.
95. C. Y. Suen and A. Dey (2003). Construction of asymmetric orthogonal arrays through finite geometries. *Journal of Statistical Planning and Inference*, **115**, 623–635.
96. A. Das and A. Dey (2004a). Optimal main effect plans with non-orthogonal blocks. *Sankhyā*, **66**, 378–384.
97. A. Das and A. Dey (2004b) Designs for diallel cross experiments with specific combining abilities. *Journal of the Indian Society of Agricultural Statistics*, Special Volume **57**, 247–256.
98. A. Dey (2004). Obituary: Sujit Kumar Mitra. *Sankhyā*, **66**, 211–212.
99. A. Das, A. Dey, S. Kageyama and K. Sinha (2005). A-efficient balanced treatment incomplete block designs. *Australasian Journal of Combinatorics*, **32**, 243–252.
100. A. Das, A. Dey and P. Saha (2005). Small asymmetric fractional factorial plans for main effects and specified two-factor interactions. *Metrika*, **62**, 33–52.
101. A. Dey (2005a). Orthogonal arrays. *Mathematics Student*, **74**, 145–152.

102. A. Dey (2005b). Projection properties of some orthogonal arrays. *Statistics and Probability Letters*, **75**, 298–306.
103. A. Dey and C. Y. Suen and A. Das (2005). Asymmetric fractional factorial plans optimal for main effects and specified two-factor interactions. *Statistica Sinica*, **15**, 751–765.
104. A. Das, A. Dey and C. K. Midha (2006). Allocating factors to the columns of an orthogonal array when certain interactions are important. *Statistics and Probability Letters*, **76**, 1570–1577.
105. M. Bose and A. Dey (2006). Combined intra-inter unit analysis of crossover designs and related optimality results. *Journal of the Indian Society of Agricultural Statistics*, **60**, 144–150.
106. A. Dey and R. Mukerjee (2006). D-optimal designs for covariate models. *Statistics*, **40**, 297–305.
107. A. Das, A. Dey, L. Y. Chan and K. Chatterjee. (2008). On $E(s^2)$-optimal supersaturated designs. *Journal of Statistical Planning and Inference*, **138**, 3749–3757.
108. A. Dey (2008). Canonical efficiency factors and related issues revisited. *Journal of the Indian Society of Agricultural Statistics*, **62**, 169–173.
109. A. Dey (2009). Orthogonally blocked three-level second order designs. *Journal of Statistical Planning and Inference*, **139**, 3698–3705.
110. K. Chatterjee and A. Dey (2010). A class of saturated row-column designs. *Indian Journal of Pure and Applied Mathematics*, **41**, 293–302.
111. A. Dey (2010). Construction of nested orthogonal arrays. *Discrete Mathematics*, **310**, 2831–2834.
112. A. Dey (2011). On the construction of two-level fractional factorial designs when some combinations are debarred. *Journal of the Indian Society of Agricultural Statistics*, **65**, 221–224.
113. A. Dey (2012a). On the construction of nested orthogonal arrays. *Australasian Journal of Combinatorics*, **54**, 37–48.
114. A. Dey (2012b). Recent developments in fractional factorial designs. *Journal of the Indian Society of Agricultural Statistics*, **66**, 251–258.
115. A. Dey and R. Mukerjee (2012a). Development of research in experimental design in India. *International Statistical Review*, **80**, 231–252.
116. A. Dey and R. Mukerjee (2012b). Efficiency factors for natural contrasts in partially confounded factorial designs. *Statistics and Probability Letters*, **82,** 2180–2188.
117. M. Bose, A. Dey and R. Mukerjee (2013). Key predistribution schemes for distributed sensor networks via block designs. *Designs, Codes and Cryptography*, **67**, 111–136.
118. A. Dey (2013). *Orthogonal Latin squares and the Falsity of Euler's Conjecture*. In R. Bhatia *et al*. Eds. Connected at Infinity II, Texts and Readings in Mathematics, 67, Hindustan Book Agency, New Delhi, 1–17.
119. A. Dey and B. Kole (2013). Small three-level second-order designs with orthogonal blocks. *Journal of Statistical Theory and Practice*, **7**, 745–752.
120. M. Bose and A. Dey (2015). *Crossover Designs*. In Angela M. Dean, Max Morris, John F. Stufken and Derek Bingham Eds. Handbook of Design and Analysis of Experiments. Boca Raton: Chapman & Hall/CRC Handbooks of Modern Statistical Methods Series, 159–195.
121. N. K. Nguyen and A. Dey (2015). A catalog of orthogonally blocked three-level second-order designs with run sizes ≤ 100. *Journal of Statistical Theory and Practice*, **9**, 537–543.
122. A. Dey (2016). Efficiency factors for natural contrasts. *Statistics and Applications* (New Series), **14**, 1–8.

123. A. Dey and D. Sarkar (2016). A note on the construction of orthogonal Latin hypercube designs. *Journal of Combinatorial Designs*, **24**, 105–111.
124. T. F. Zhang, Y. Y. Zong and A. Dey (2016). On the construction of asymmetric orthogonal arrays. *Journal of Statistical Planning and Inference*, **170**, 77–82.
125. A. Dey and D. Sarkar (2017). A new family of orthogonal Latin hypercube designs. *Australasian Journal of Combinatorics*, **69**, 58–62.
126. A. Dey, R. Singh and A. Das (2017). Efficient paired choice designs with fewer choice pairs. *Metrika*, **80**, 309–317.
127. T. F. Zhang, Q. Deng and A. Dey (2017). Construction of asymmetric orthogonal arrays of strength three via a replacement method. *Journal of Combinatorial Designs*, **25,** 339–348.
128. A. Dey and D. Sarkar (2019). D-efficient composite-type second order designs via computer search. *Statistics and Applications* **17** (New Series) "Special Volume to Felicitate Professor Arun Kumar Nigam on his 75th Birthday", 33–40.
129. T. F. Zhang, G. Wu and A. Dey (2019). Construction of some new families of nested orthogonal arrays. *Communications in Statistics: Theory and Methods*, **48**, 774–779.

Statistics and Applications {ISSN 2454-7395(online)}
Volume 19, No. 1, 2021 (New Series), pp 11-28

Unbalanced Two-symbol $E(s^2)$-Optimal Designs

Dursun A. Bulutoglu[1], Kashinath Chatterjee[2], Stelios D. Georgiou[3], Christos Koukouvinos[4], Kenneth J. Ryan[5] and Stella Stylianou[3]

[1]*Department of Mathematics and Statistics, Air Force Institute of Technology, WPAFB, OH 45433*

[2]*Department of Statistics, Visva Bharati University, Santiniketan, West Bengal, India*

[3]*Department of Mathematics, School of Science, RMIT University, Melbourne, VIC 3001, Australia*

[4]*Department of Mathematics, National Technical University of Athens, Zografou 15773, Athens, Greece*

[5]*Department of Statistics, West Virginia University, Morgantown, West Virginia 26506, USA*

Received: 27 April 2020; Revised: 15 June 2020; Accepted: 20 June 2020

Abstract

Two-symbol supersaturated designs (SSDs) are two-symbol arrays in which the number of rows is no larger than the number of columns. In this paper, a lower bound for the $\mathrm{E}(s^2)$ value of SSDs that are not necessarily balanced is derived. The sharpness of the newly derived lower bound is analyzed theoretically by using constructions of $\mathrm{E}(s^2)$-optimal SSDs and computationally by using the NOA_4 algorithm in Ryan and Bulutoglu (2007). Applications of the newly derived $\mathrm{E}(s^2)$ lower bound to searching for D-optimal designs and equiangular lines are discussed.

Key words: Two-symbol unbalanced SSD; $\mathrm{E}(s^2)$-optimal designs; D-optimal designs; Lower bound.

AMS Subject Classifications: 62M20, 62F15

1. Introduction

Two-symbol supersaturated designs (SSDs) are two symbol arrays with the following properties. The number of rows N does not exceed the number of columns m and no pair of columns are fully aliased, *i.e.*, there exists no pair of identical columns up to permuting the symbols within a column. A two-symbol array is *balanced* if each of the two symbols in a column appears the same number of times when the number of rows N is even or the absolute difference between the frequencies of the occurrences of the two symbols in each column is 1 when N is odd. A two symbol array that is not balanced is called *unbalanced*. Two-symbol SSDs are commonly coded with symbol set $\{0, 1\}$ or $\{\pm 1\}$ and are particularly useful in screening experimentation due to their row-size economy (Georgiou, 2014). It has long been assumed in the literature that SSDs should be balanced. However, unbalanced SSDs are of interest to practitioners who are willing to compromise on the balance property due to high costs. In particular, unbalanced SSDs are useful when restrictions embedded in the problem at hand makes it infeasible to use a balanced SSD. Such SSDs are also

Correponding Author: Dursun Bulutoglu
Email: dursun.bulutoglu@gmail.com

preferable in cases where certain symbols of some columns need to be examined but are expensive to set.

WLOG, assume that each column of the N-row array (SSD) with m columns each with symbols from $\{0, 1\}$ has 0 at most $\lfloor N/2 \rfloor$ times, where $\lfloor . \rfloor$ is the floor function. Define k_l to be the number of columns in which 0 appears l times for $l = 0, \ldots, \lfloor N/2 \rfloor$. Clearly, $\sum_l k_l = m$. Denote this class of arrays by $\mathcal{D}(N, 2^m, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$. Let $\mathcal{D}^{\pm}(N, 2^m, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$ denote the corresponding class of arrays when -1 is used instead of 0. For even N, if we choose $k_l = 0$ for $0 \leq l < \lfloor N/2 \rfloor$ and $k_{\lfloor N/2 \rfloor} = m$, we get a balanced array (SSD), and we get an unbalanced array if $k_l \neq 0$ for at least one l with $l < \lfloor N/2 \rfloor$. We call the vector $(k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor})$ the *balancedness structure* of each $\mathbf{D} \in \mathcal{D}(N, 2^m, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$ or each $\mathbf{D} \in \mathcal{D}^{\pm}(N, 2^m, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$.

Example 1: Consider the 4-row SSD

$$\mathbf{D} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \in \mathcal{D}(4, 2^4, (0, 1, 3)).$$

This SSD is unbalanced with $k_1 = 1$, $k_2 = 3$ and $\sum_l k_l = m = 4$.

For $m \geq N$, an N row, m column, two-symbol array $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m]$ with entries from $\{\pm 1\}$ is a supersaturated design if it has no two columns $\mathbf{x}_i$ and $\mathbf{x}_j$ such that $i \neq j$ and $\mathbf{x}_i^\top \mathbf{x}_j \in \{-N, N\}$. The $\mathrm{E}(s^2)$ value of $\mathbf{X}$ is defined as

$$\mathrm{E}(s^2) = \frac{\sum_{i \neq j} s_{ij}^2}{m(m-1)},$$

where $s_{ij} = \mathbf{x}_i^\top \mathbf{x}_j$ for $1 \leq i \neq j \leq m$. A two-symbol SSD is mapped to a $\{\pm 1\}$ SSD by assigning $+1$ to one symbol and -1 to the other symbol in each column. We call a resulting $\{\pm 1\}$ SSD a corresponding $\{\pm 1\}$ SSD. Then each of the concepts defined for a $\{\pm 1\}$ SSD is defined for a two-symbol SSD via one of its corresponding $\{\pm 1\}$ SSDs. The $\mathrm{E}(s^2)$ value is used to compare two-symbol SSDs with the same number of rows and columns (Georgiou, 2014). An SSD with a smaller $\mathrm{E}(s^2)$ value is more desirable (Georgiou, 2014), and an SSD with the smallest possible $\mathrm{E}(s^2)$ value is called $\mathrm{E}(s^2)$-optimal. For a detailed review of the $\mathrm{E}(s^2)$ optimality criteria for two-symbol SSDs, the reader is referred to (Georgiou, 2014).

Ryan and Bulutoglu (2007) and Das *et. al.* (2008) gave the sharpest known lower bound for balanced SSDs with even N. Bulutoglu and Ryan (2008) and Suen and Das (2010) derived an improved $\mathrm{E}(s^2)$ lower bound for two-symbol SSDs with odd N. For unbalanced SSDs, the best known $\mathrm{E}(s^2)$ lower bounds are not applicable. We generalize the results in Bulutoglu and Ryan (2008) and Suen and Das (2010) to unbalanced SSDs. We derive a lower bound for the $\mathrm{E}(s^2)$ value of unbalanced two-symbol SSDs and present some families of $\mathrm{E}(s^2)$-optimal unbalanced SSDs. Part of our derivation is based on an adaptation of the derivation in Bulutoglu and Ryan (2008).

For an SSD $\mathbf{X} \in \mathcal{D}^{\pm}(N, 2^m, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$ let $s_{\max} = \max_{i<j} |s_{ij}|$ and $f_{s_{\max}}$ be the frequency of $s_{\max}$ in $\{s_{ij}\}_{i<j}$. Then $\mathbf{X}$ is called *minimax-optimal* if no other SSD in $\mathcal{D}^{\pm}(N, 2^m, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$ has a smaller $s_{\max}$ or smaller $f_{s_{\max}}$ at the smallest possible $s_{\max}$. For balanced SSDs, Ryan and Bulutoglu (2007) and Bulutoglu and Ryan (2008) used minimax optimality as a secondary

criterion for picking an SSD among $\mathrm{E}(s^2)$-optimal SSDs. Finding a minimax-optimal and balanced SSD among $\mathrm{E}(s^2)$-optimal and balanced SSDs is a very difficult problem as mentioned in Morales and Bulutoglu (2018). Some of the unbalanced infinite families of $\mathrm{E}(s^2)$-optimal SSDs in Section 3 are also minimax-optimal.

There are no known theories or construction methods of SSDs for any choice of $k_l, 0 \leq l \leq \lfloor N/2 \rfloor$, because an achievable lower bound for $\mathrm{E}(s^2)$ is not known for the general case and it is not possible to prove $\mathrm{E}(s^2)$-optimality without resorting to full enumeration. Therefore, there is a need to develop a general sharp lower bound that also covers unbalanced arrays. In this paper, we generalize the best known $\mathrm{E}(s^2)$ lower bound for balanced two-symbol SSDs to all two-symbol SSDs with a given balancedness structure. Additionally, we describe how our newly derived $\mathrm{E}(s^2)$ lower bound can be used to speed up search algorithms for finding two-symbol D-optimal designs in general.

An SSD can be thought as a frame, *i.e.*, a spanning set for its column space. Moreover, certain $\mathrm{E}(s^2)$-optimal SSDs are tight frames (Morales and Bulutoglu, 2018). Another motivation for generalizing the best known $\mathrm{E}(s^2)$ lower bound for balanced two-symbol SSDs to all two-symbol SSDs with a given balancedness structure is that there is no balancedness requirement for frames. Furthermore, certain $\mathrm{E}(s^2)$-optimal and minimax-optimal SSDs are equiangular tight frames and imply the existence of certain strongly regular graphs (Morales and Bulutoglu, 2018; Waldron, 2009).

This paper is organized as follows. In Section 2, we derive a previously unknown lower bound for the $\mathrm{E}(s^2)$ value of an unbalanced two-symbol SSD with symbols from $\{\pm 1\}$ given its column sums. After providing a naive $\mathrm{E}(s^2)$ lower bound, Section 3 theoretically analyzes the Section 2 bound in terms of its achievability and provides families of unbalanced $\mathrm{E}(s^2)$-optimal SSDs achieving the Section 2 bound. Some of these SSDs are optimal with respect to the minimax criterion as well. Section 4 provides computational test results obtained by using the NOA_4 algorithm in Ryan and Bulutoglu, (2007) for the achievability of the Section 2 bound. Finally, in Section 5, we discuss two possible applications of our newly derived $\mathrm{E}(s^2)$ lower bound. In particular, in Section 5.1, we provide an application to searching for D-optimal designs. Moreover, in Section 5.2, for a given t such that $0 < t$, we discuss an application to finding upper bounds on the maximum number of columns for a two-symbol $\{\pm 1\}$ SSD with N rows whose each pairwise column angle is in $[\arccos(t/N), \arccos(-t/N)]$.

2. A General Lower Bound

In this section, we derive a previously unknown lower bound for the $\mathrm{E}(s^2)$ value of unbalanced SSDs. We first provide some definitions and lemmas that will be useful in proving the desired lower bound.

For a 2-symbol array let $c_{i_1 i_2}$ be the number of coincidences in the i_1'th and i_2'th rows for $1 \leq i_1 \neq i_2 \leq N$. The following lemma provides the number of coincidences in all the different pairs of rows in a given SSD.

Lemma 1: For $\mathbf{D} \in \mathcal{D}(N, 2^m, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$ or $\mathbf{D} \in \mathcal{D}^{\pm}(N, 2^m, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$,

$$2 \sum_{i_1 \neq i_2} c_{i_1 i_2} = \sum_{l} (N - 2l)^2 k_l + mN(N-2).$$

Proof: For given $k_l, 0 \leq l \leq \lfloor N/2 \rfloor$ and $\mathbf{D} \in \mathcal{D}(N, 2^m, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$, we have

$$\mathbf{X}_0\mathbf{X}_0^\top = \begin{pmatrix} m & 2c_{12}-m & \cdots & 2c_{1N}-m \\ 2c_{12}-m & m & \cdots & 2c_{2N}-m \\ \vdots & \vdots & \vdots & \vdots \\ 2c_{1N}-m & 2c_{2N}-m & \cdots & m \end{pmatrix}.$$

Therefore,

$$\begin{aligned} \mathbf{1}_N^\top \mathbf{X}_0\mathbf{X}_0^\top \mathbf{1}_N &= 2\sum_{i_1 \neq i_2} c_{i_1 i_2} - mN(N-2) \\ &= \sum_l (N-2l)^2 k_l. \end{aligned}$$

Let $\mathbf{D} \in \mathcal{D}(N, 2^m, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$. For $1 \leq j \leq m$, let n_α^j be the number of times $\alpha \in \{0, 1\}$ appears in the j'th column of $\mathbf{D}$. Also, for $1 \leq j_1 \neq j_2 \leq m$, $\alpha, \beta \in \{0, 1\}$, let $n_{\alpha\beta}^{j_1 j_2}$ be the number of times the symbol combination (α, β) appears as rows of the $N \times 2$ array obtained by concatenating j_1'th and j_2'th columns of $\mathbf{D}$. Then

$$s_{j_1 j_2}^2 = 4\sum_{\alpha,\beta} (n_{\alpha\beta}^{j_1 j_2})^2 - 2\sum_\alpha (n_\alpha^{j_1})^2 - 2\sum_\beta (n_\beta^{j_2})^2 + N^2. \tag{1}$$

Based on (1), we can express the $\mathrm{E}(s^2)$ value in a convenient form so as to obtain another lower bound.

Lemma 2: For $\mathbf{D} \in \mathcal{D}(N, 2^m, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$,

$$\mathrm{E}(s^2) = \frac{4\sum_{j_1 \neq j_2, \alpha, \beta} (n_{\alpha\beta}^{j_1 j_2})^2 - 4(m-1)\sum_l [l^2 + (N-l)^2] k_l + m(m-1)N^2}{m(m-1)}.$$

Proof: For $\mathbf{D} \in \mathcal{D}(N, 2^m, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$, we have

$$\begin{aligned} m(m-1)\mathrm{E}(s^2) &= \sum_{j_1 \neq j_2} s_{j_1 j_2}^2 \\ &= \sum_{j_1 \neq j_2} \left[4\sum_{\alpha,\beta} (n_{\alpha\beta}^{j_1 j_2})^2 - 2\sum_\alpha (n_\alpha^{j_1})^2 - 2\sum_\beta (n_\beta^{j_2})^2 + N^2 \right] \\ &= 4\sum_{j_1 \neq j_2, \alpha, \beta} (n_{\alpha,\beta}^{j_1 j_2})^2 - 2(m-1)\left[\sum_{j_1, \alpha} (n_\alpha^{j_1})^2 + \sum_{j_2, \beta} (n_\beta^{j_2})^2 \right] + m(m-1)N^2 \\ &= 4\sum_{j_1 \neq j_2, \alpha, \beta} (n_{\alpha\beta}^{j_1 j_2})^2 - 4(m-1)\sum_l [l^2 + (N-l)^2] k_l + m(m-1)N^2. \end{aligned}$$

For a given $\mathbf{D} \in \mathcal{D}(N, 2^m, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$, let l_{j_i} be the number of zeros in the j_i'th

column of **D**. Then we have

$$\begin{aligned} n_{00}^{j_1 j_2} &= l_{j_2} - n_{10}^{j_1 j_2} \\ n_{01}^{j_1 j_2} &= l_{j_1} - l_{j_2} + n_{10}^{j_1 j_2} \\ n_{11}^{j_1 j_2} &= N - l_{j_1} - n_{10}^{j_1 j_2}. \end{aligned} \tag{2}$$

Let $\sum_{\alpha,\beta}(n_{\alpha\beta}^{j_1 j_2})^2 = f(l_{j_1}, l_{j_2}, n_{10}^{j_1 j_2})$. Then by equations (2),

$$f(l_{j_1}, l_{j_2}, n_{10}^{j_1 j_2}) = 4(n_{10}^{j_1 j_2})^2 + (4l_{j_1} - 4l_{j_2} - 2N)n_{10}^{j_1 j_2} + l_{j_2}^2 + (l_{j_1} - l_{j_2})^2 + (N - l_{j_1})^2.$$

For fixed l_{j_1}, l_{j_2}, $f(l_{j_1}, l_{j_2}, n_{10}^{j_1 j_2})$ is a convex function of $(n_{10}^{j_1 j_2})$. By differentiating $f(l_{j_1}, l_{j_2}, n_{10}^{j_1 j_2})$ with respect to $n_{10}^{j_1 j_2}$ we see that $f(l_{j_1}, l_{j_2}, n_{10}^{j_1 j_2})$ is minimized at

$$\widehat{n}_{10}^{j_1 j_2}(l_{j_1}, l_{j_2}) = \left[\frac{N + 2l_{j_2} - 2l_{j_1}}{4}\right],$$

where $[x]$ is the integer closest to x. Define

$$\widehat{n}_{10}(i, j) = \left[\frac{N + 2j - 2i}{4}\right].$$

Then we have

$$\begin{aligned} \sum_{j_1 \neq j_2, \alpha, \beta} (n_{\alpha,\beta}^{j_1 j_2})^2 &\geq \sum_i k_i(k_i - 1) f(i, i, \widehat{n}_{10}(i, i)) + \sum_{i \neq j} k_i k_j f\,(i, j, \widehat{n}_{10}(i, j)) \\ &= \theta_1^* = \sum_i k_i(k_i - 1) f\left(i, i, \left[\frac{N}{4}\right]\right) + \sum_{i \neq j} k_i k_j f\left(i, j, \left[\frac{N + 2j - 2i}{4}\right]\right). \end{aligned} \tag{3}$$

The following lemma, whose proof follows from Lemma 2 and inequality (3), provides a lower bound for the $\mathrm{E}(s^2)$ value of $\mathbf{D} \in \mathcal{D}(N, 2^m, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$.

Lemma 3: For $\mathbf{D} \in \mathcal{D}(N, 2^m, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$,

$$\mathrm{E}(s^2) \geq \mathrm{LB}_1,$$

where

$$\mathrm{LB}_1 = \frac{4\theta_1^* - 4(m-1)\sum_l [l^2 + (N-l)^2] k_l + m(m-1)N^2}{m(m-1)}.$$

By routine algebra we get the following result.

Lemma 4: For balanced and two-symbol arrays (SSDs)

$$\mathrm{LB}_1 = \mathrm{LB}_1(N, (0, \ldots, 0, m)) = \begin{cases} 1, & \text{if } N \text{ is odd}, \\ 0, & \text{if } N = 0 \pmod 4, \\ 4, & \text{if } N = 2 \pmod 4. \end{cases}$$

When $N = 2 \pmod 4$, the bound in Lemma 4 for $m = N+1$ is achievable if a skew-symmetric Hadamard matrix of order $N+2$ exists (Morales *et al.* 2019). Such a Hadamard matrix is conjectured to exist for each N divisible by 4 (Koukouvinos and Stylianou, 2008). It is plain to verify the following remark.

Remark 1: For balanced SSDs, $m \geq N+3$ or $N = 0 \pmod 4$, $\mathrm{LB}_1(N, (0, \ldots, 0, m))$ is strictly smaller than the $\mathrm{E}(s^2)$ lower bound in Das *et al.* (2008) and Ryan and Bulutoglu (2007). For balanced SSDs, $N = 2 \pmod 4$ and $m \leq N+2$, $\mathrm{LB}_1(N, (0, \ldots, 0, m))$ is the same as the $\mathrm{E}(s^2)$ lower bound in Das *et al.* (2008) and Ryan and Bulutoglu (2007). For balanced SSDs and odd N, $\mathrm{LB}_1(N, (0, \ldots, 0, m))$ cannot be sharper than the $\mathrm{E}(s^2)$ lower bound in Bulutoglu and Ryan (2008). In particular, when $N = 3 \pmod 4$ and $m = N$, $\mathrm{LB}_1(N, (0, \ldots, 0, m))$ equals to the $\mathrm{E}(s^2)$ lower bound in Bulutoglu and Ryan (2008). A numerical check suggests that these are the only odd N cases for which equality is satisfied.

For $\mathbf{D} \in \mathcal{D}^{\pm}(N, 2^m, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$, let $\Delta = \sum_l (N-2l)^2 k_l$ and

$$F(p) := 8p^2 + 4N^2 - 8Np - 4N + 4\max\{|-mN + \Delta + qN(N-1)| - 4p^2 - 2N^2 + 4Np + 2N, 0\}.$$

Then

$$F(p) = \begin{cases} F_1(p) & \text{if } |-mN + \Delta + qN(N-1)| - 4p^2 - 2N^2 + 4Np + 2N \leq 0, \\ F_2(p) & \text{otherwise,} \end{cases}$$

where

$$\begin{aligned} F_1(p) &= 8p^2 + 4N^2 - 8Np - 4N \\ F_2(p) &= -8p^2 - 4N^2 + 8Np + 4N + 4|-mN + \Delta + qN(N-1)|. \end{aligned}$$

The following theorem provides another $\mathrm{E}(s^2)$ lower bound for $\mathbf{D}$.

Theorem 1: There is a unique q such that $-2N \leq (Nm-\Delta)/(N-1) - qN < 2N$ and $m+q \equiv 2 \pmod 4$. Let

$$p_-^* = \frac{N - \sqrt{2N - N^2 + |-mN + \Delta + qN(N-1)|}}{2},$$

$$p_+^* = \frac{N + \sqrt{2N - N^2 + |-mN + \Delta + qN(N-1)|}}{2},$$

$g(q) := (m+q)^2 N - q^2 N^2 - mN^2 - 2q\Delta$, and $K = \sum_i k_{2i+1}$. For odd N, let

$$\mathrm{LB}_2 = \begin{cases} \dfrac{16\left\lceil \frac{g(q) + F_1\left(\left\lfloor \frac{N}{2} \right\rfloor\right) - m(m-1)}{16} \right\rceil^+ + m(m-1)}{m(m-1)}, & \text{if } \left|\frac{Nm-\Delta}{N-1} - qN\right| < N, \\ \dfrac{16\left\lceil \frac{g(q) + \min\{F_1(\lfloor p_-^* \rfloor), F_2(\lceil p_-^* \rceil)\} - m(m-1)}{16} \right\rceil^+ + m(m-1)}{m(m-1)}, & \text{otherwise,} \end{cases}$$

where $\lceil . \rceil$ is the ceiling function. For $N \equiv 0 \pmod 4$, let

$$\mathrm{LB}_2 = \begin{cases} \dfrac{32\left\lceil \frac{g(q)+F_1\left(\left\lfloor \frac{N}{2} \right\rfloor\right)-8K(m-K)}{32} \right\rceil^+ +8K(m-K)}{m(m-1)}, & \text{if } \left|\frac{Nm-\Delta}{N-1} - qN\right| < N, \\ \dfrac{32\left\lceil \frac{g(q)+\min\{F_1(\lfloor p_-^* \rfloor), F_2(\lceil p_-^* \rceil)\}-8K(m-K)}{32} \right\rceil^+ +8K(m-K)}{m(m-1)}, & \text{otherwise.} \end{cases}$$

For $N \equiv 2 \pmod 4$, let

$$\mathrm{LB}_2 = \begin{cases} \dfrac{\theta\left\lceil \frac{g(q)+F_1\left(\left\lfloor \frac{N}{2} \right\rfloor\right)-4K(K-1)-4(m-K)(m-K-1)}{\theta} \right\rceil^+ +4K(K-1)+4(m-K)(m-K-1)}{m(m-1)}, & \text{if } \left|\frac{Nm-\Delta}{N-1} - qN\right| < N, \\ \dfrac{\theta\left\lceil \frac{g(q)+\min\{F_1(\lfloor p_-^* \rfloor), F_2(\lceil p_-^* \rceil)\}-4K(K-1)-4(m-K)(m-K-1)}{\theta} \right\rceil^+ +4K(K-1)+4(m-K)(m-K-1)}{m(m-1)}, & \text{otherwise,} \end{cases}$$

where

$$\theta = \begin{cases} 64 & \text{if } K = 0, \\ 32 & \text{otherwise,} \end{cases}$$

and $\lceil x \rceil^+ = \max\{0, \lceil x \rceil\}$. Then

$$\mathrm{E}(s^2) \geq \mathrm{LB}_2 \,.$$

Proof: The proof is an adaptation of the proof of Theorem 1 in Bulutoglu and Ryan (2008). For general $\mathbf{D} \in \mathcal{D}^{\pm}(N, 2^m, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$, the $F(p)$ in Bulutoglu and Ryan (2008) becomes

$$F(p) := 8p^2 + 4N^2 - 8Np - 4N + 4\max\{|-mN + \Delta + qN(N-1)| - 4p^2 - 2N^2 + 4Np + 2N, 0\},$$

so that $F(p)$ is continuous. Moreover,

$$F'(p) = \begin{cases} 16p - 8N, & \text{if } |-mN + \Delta + qN(N-1)| - 4p^2 - 2N^2 + 4Np + 2N < 0, \\ -16p + 8N, & \text{if } |-mN + \Delta + qN(N-1)| - 4p^2 - 2N^2 + 4Np + 2N > 0. \end{cases}$$

Hence, $F(p)$ has all of its local minima at p such that

$$-4p^2 - 2N^2 + 4Np + 2N + |-mN + \Delta + qN(N-1)| = 0,$$

and these p's are p_-^* and p_+^*. By the reflection symmetry of $F_1(p)$ and $F_2(p)$ along the axis $y = F_1(p_-^*)$, both of these local minima are in fact global minima, and $F(p)$ for $p \in \{0, 1, \ldots, (N+1)/2\}$ is minimized at $p = \lfloor p_-^* \rfloor$ or $p = \lceil p_-^* \rceil$. The result now follows from

$$s_{ij} \equiv \begin{cases} N \pmod 4, & \text{if the } i\text{'th and } j\text{'th columns of } \mathbf{D} \text{ have both even} \\ & \text{or odd number of } -1\text{'s,} \\ (N+2) \pmod 4, & \text{otherwise.} \end{cases}$$

Remark 2: LB_2 is cheaper to compute compared to the corresponding lower bound in Bulutoglu and Ryan (2008). This is because computing LB_2 requires computing only $F_1(\lfloor p^*_- \rfloor)$ and $F_2(\lceil p^*_- \rceil)$ instead of computing $F(p)$ for $p = 1, 2, \ldots, (N+1)/2$.

By Lemma 3, we obtain the following theorem.

Theorem 2: For $\mathbf{D} \in \mathcal{D}(N, 2^m, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$,

$$\mathrm{E}(s^2) \geq \mathrm{LB} = \mathrm{LB}(N, (k_0, k_1, \ldots, k_{\lfloor \frac{N}{2} \rfloor})) = \max\{\mathrm{LB}_1, \mathrm{LB}_2\}. \tag{4}$$

For balanced SSDs with an odd number of rows, it is easy to show that LB_2 reduces to the $\mathrm{E}(s^2)$ lower bound provided in Bulutoglu and Ryan (2008). (We numerically verified this for $7 \leq N \leq 41$ and $N \leq m \leq 4N$.) For balanced SSDs with an even number of rows, LB is still a valid lower bound; however, it cannot be sharper than the $\mathrm{E}(s^2)$ lower bound of Ryan and Bulutoglu (2007) or of Das *et al.* (2008). This is because the assumed set of constraints for the hypothetical SSD in the derivation of LB_2 are satisfied by the hypothetical SSD in the derivation of the lower bound Ryan and Bulutoglu (2007) and that in Das *et al.* (2008). In fact, we observed in 700 of the 1,314 even N cases with $8 \leq N \leq 40$ and $N \leq m \leq 4N$ that the $\mathrm{E}(s^2)$ lower bound of Ryan and Bulutoglu (2007) or of Das *et al.* (2008) is sharper than the lower bound in Bulutoglu and Ryan (2008). Hence, for balanced and even N SSDs, LB should not be used. Numerically comparing LB_2 to the bound in Ryan and Bulutoglu (2007) and that in Das *et al.* (2008) provided us with a check for the correctness of LB_2. We implemented a similar check for the correctness of LB_1. We observed that LB_1 is most useful when m is close to N and is dominated by LB_2 as m increases.

3. Theoretical Achievability

In this section, we first derive a naive $\mathrm{E}(s^2)$ lower bound for an N row, m column, two-symbol SSD. This bound does not depend on the column sums of the SSD. Then we show that if this naive $\mathrm{E}(s^2)$ lower bound is achievable then LB_2 is also achievable.

Let $\mathbf{X}$ be an N row, m column SSD whose column symbols are from $\{\pm 1\}$. Then the off-diagonal entries of $\mathbf{X}\mathbf{X}^\top$ are odd if and only if m is odd. When m is divisible by 4, it is possible for $\mathbf{X}$ to have mutually orthogonal rows. When $m = 2 \pmod 4$, $\mathbf{X}$ can have at most $\lfloor N/2 \rfloor \lceil N/2 \rceil$ pairs of orthogonal rows. From these facts, we immediately get the following naive $\mathrm{E}(s^2)$ lower bound

$$\mathrm{E}(s^2) \geq \begin{cases} \frac{N(N-1)-mN^2+Nm^2}{m(m-1)}, & \text{if } m \text{ is odd,} \\ \frac{-mN^2+Nm^2}{m(m-1)}, & \text{if } m = 0 \pmod 4, \\ \frac{4\left(\lfloor \frac{N}{2} \rfloor \left(\lfloor \frac{N}{2} \rfloor - 1\right) + \lceil \frac{N}{2} \rceil \left(\lceil \frac{N}{2} \rceil - 1\right)\right) - mN^2 + Nm^2}{m(m-1)}, & \text{if } m = 2 \pmod 4. \end{cases} \tag{5}$$

The lower bound $\mathrm{LB}_2(N, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$ is based on a derivation where $\mathbf{X}\mathbf{X}^\top$ has off-diagonals from the set $\{-4, -2, 0, 2\}$ or the set $\{-2, 0, 2, 4\}$ for even $\sum_i k_i$ and from the set $\{-3, -1, 1, 3\}$ for odd $\sum_i k_i$, where the off-diagonal entries sum to $\sum_{l=0}^{n} (N - 2l)^2 k_l - Nm$. For an SSD $\mathbf{X}$ achieving the naive $\mathrm{E}(s^2)$ lower bound (5), $\mathbf{X}\mathbf{X}^\top$ has off-diagonals from the set $\{-2, 0, 2\}$ for even $\sum_i k_i$ and from the set $\{\pm 1\}$ for odd $\sum_i k_i$. Moreover, the entries of such an $\mathbf{X}\mathbf{X}^\top$ sum to $\sum_{l=0}^{n} (N - 2l)^2 k_l$. Since $\{\pm 1\} \subseteq \{-3, -1, 1, 3\}$ and $\{-2, 0, 2\} \subseteq \{-2, 0, 2, 4\} \cap \{-4, -2, 0, 2\}$, we conclude that $\mathrm{LB}_2(N, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$ is always at least as sharp as the naive $\mathrm{E}(s^2)$ lower bound (5). Hence,

an SSD $\mathbf{X}$ achieving the naive $\mathrm{E}(s^2)$ lower bound (5) also achieves $\mathrm{LB}_2(N, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$, i.e., if the naive $\mathrm{E}(s^2)$ lower bound (5) is achievable, then $\mathrm{LB}_2(N, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$ is equal to the naive $\mathrm{E}(s^2)$ lower bound (5).

There are cases in which LB_2 is strictly larger than the naive $\mathrm{E}(s^2)$ lower bound (5). In particular, for balanced SSDs when N is odd LB_2 reduces to the $\mathrm{E}(s^2)$ lower bound derived in Bulutoglu and Ryan (2008), and there are SSDs ($\mathbf{X}$'s) achieving this lower bound such that the off-diagonal entries of $\mathbf{X}\mathbf{X}^\top$ are not all from the set $\{\pm 1\}$.

For each N, m combination, an $\mathrm{E}(s^2)$-optimal SSD achieving the naive bound (5) can be constructed by using Hadamard matrices. A $t \times t$ matrix $\mathbf{H}_t$ of ± 1's is called a Hadamard matrix if $\mathbf{H}_t^\top \mathbf{H}_t = t\mathbf{I}_t$, where $\mathbf{I}_t$ is the $t \times t$ identity matrix. It is well-known that t must be divisible by 4 for a $t \times t$ Hadamard matrix to exist. It is conjectured that $t \times t$ Hadamard matrix exists whenever t is divisible by 4. Let $\mathbf{1}_r$ be the column of all $1s$ of length r. It is easy to show that any Hadamard matrix can be put in the form

$$\mathbf{H}_t = \left(\begin{array}{cc|cc} 1 & 1 & \mathbf{1}_{\frac{t}{2}-1}^\top & \mathbf{1}_{\frac{t}{2}-1}^\top \\ 1 & -1 & \mathbf{1}_{\frac{t}{2}-1}^\top & -\mathbf{1}_{\frac{t}{2}-1}^\top \\ \hline \mathbf{1}_{\frac{t}{2}-1} & \mathbf{1}_{\frac{t}{2}-1} & \multicolumn{2}{c}{\multirow{2}{*}{$\mathbf{A}$}} \\ \mathbf{1}_{\frac{t}{2}-1} & -\mathbf{1}_{\frac{t}{2}-1} & & \end{array}\right) \tag{6}$$

by applying signed column and/or row permutations (by right and/or left multiplying with permutation matrices that are right or left multiplied by ± 1 diagonal matrices).

Let $\mathbf{X}$ be an N row, m column $\mathrm{E}(s^2)$-optimal SSD achieving the naive $\mathrm{E}(s^2)$ lower bound (5). In what follows, we describe how $\mathbf{X}$ can be constructed provided that the Hadamard conjecture is true. If m is divisible by 4, then $\mathbf{X}$ can be taken to be any N rows of a Hadamard matrix $\mathbf{H}_m$. If $m = 1 \pmod 4$ ($m = 3 \pmod 4$), then $\mathbf{X}$ can be constructed by first adding (deleting) any column with entries in $\{-1, 1\}$ to (from) a Hadamard matrix $\mathbf{H}_{m-1}$ ($\mathbf{H}_{m+1}$) followed by picking any N rows from the resulting matrix. If $m = 2 \pmod 4$, let $\mathbf{H}_{m+2}$ be a Hadamard matrix. Let $\mathbf{A}$ be obtained from $\mathbf{H}_{m+2}$ after putting $\mathbf{H}_{m+2}$ in form (6). Then any N rows of $\mathbf{A}$ can be taken to be $\mathbf{X}$.

Two arrays $\mathbf{D}_1$ and $\mathbf{D}_2$ are *equivalent* if $\Pi_1 \mathbf{D}_1 \Pi_2 = \mathbf{D}_2$ for some signed permutation matrices Π_1 and Π_2 (i.e., permutation matrices that are right or left multiplied by ± 1 diagonal matrices). If $\mathbf{X}_1$ and $\mathbf{X}_2$ are equivalent SSDs, then $\mathrm{SS}(\mathbf{X}_1 \mathbf{X}_1^\top) = \mathrm{SS}(\mathbf{X}_2 \mathbf{X}_2^\top)$, where $\mathrm{SS}(\mathbf{M})$ is the sum squares of the entries of a matrix $\mathbf{M}$. Hence, if $\mathbf{X}_1$ is an $\mathrm{E}(s^2)$-optimal SSD achieving the naive $\mathrm{E}(s^2)$ lower bound (5), then any other SSD equivalent to $\mathbf{X}_1$ is also $\mathrm{E}(s^2)$-optimal and achieves the naive $\mathrm{E}(s^2)$ lower bound (5).

Not every $N, m, k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}$ combination allows a $\mathbf{D} \in \mathcal{D}^{\pm}(N, 2^m, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$ that achieves the naive $\mathrm{E}(s^2)$ lower bound (5). If $\mathbf{D} \in \mathcal{D}^{\pm}(N, 2^m, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$ achieves the naive $\mathrm{E}(s^2)$ lower bound (5), then the fact that $\sum_{l=0}^{n} (N - 2l)^2 k_l$ must be equal to the sum of the entries of $\mathbf{D}\mathbf{D}^\top$ implies the constraint

$$\sum_{l=0}^{n}(N-2l)^2k_l = \begin{cases} Nm+2\alpha_1-2\beta_1, & \text{if } m \text{ is odd,} \\ Nm, & \text{if } m=0 \pmod 4, \\ Nm+4\alpha_2-4\beta_2, & \text{if } m=2 \pmod 4 \end{cases} \tag{7}$$

on $N, m, k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}$ for some non-negative integers $\alpha_1, \beta_1, \alpha_2, \beta_2$. The next theorem follows from the derivation of $\mathrm{LB}_2(N, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$.

Theorem 3: The naive $\mathrm{E}(s^2)$ lower bound (5) is equal to $\mathrm{LB}_2(N, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$ if and only if the constraint in equation (7) is satisfied.

Circulant matrices can be used to construct unbalanced, $\mathrm{E}(s^2)$-optimal, minimax-optimal SSDs and D-optimal designs. A matrix is called *circulant* if each row vector is shifted one element to the right relative to the preceding row vector. A circulant matrix $\mathbf{A} = circ(\mathbf{a})$ is determined by its first row $\mathbf{a}$. Each row of $\mathbf{A}$ is a cyclic shift of the vector $\mathbf{a}$ to the right.

The dual Gram matrix and the Gram matrix of a matrix $\mathbf{A}$ are defined to be $\mathbf{A}\mathbf{A}^\top$ and $\mathbf{A}^\top\mathbf{A}$. If $\mathbf{x} = (x_0, x_1, \ldots, x_{t-1})$ is a vector of length t, the *periodic autocorrelation function* $P_{\mathbf{x}}(s)$ (abbreviated as PAF) is defined, reducing $i + s$ modulo t, as

$$P_{\mathbf{x}}(s) = \sum_i x_i x_{i+s} \quad \text{for } s = 0, 1, \ldots, t-1.$$

The (dual) Gram matrix of a circulant matrix is also circulant and can be calculated by using the periodic autocorrelation function of its first row.

Let t be odd and $\mathbf{A}, \mathbf{B}$ be $t \times t$ circulant matrices with entries in $\{\pm 1\}$. Let $\mathbf{a} = (a_0, a_1, \ldots, a_{t-1})$ and $\mathbf{b} = (b_0, b_1, \ldots, b_{t-1})$ be the first rows of $\mathbf{A}$ and $\mathbf{B}$. Also, let

$$P_{\mathbf{a}}(s) + P_{\mathbf{b}}(s) = \gamma_s \quad \text{for } s = 1, 2, \ldots, t-1, \tag{8}$$

where $|\gamma_s| = \gamma$ is a constant positive real number. Then

$$\mathbf{A}\mathbf{A}^\top + \mathbf{B}\mathbf{B}^\top = (2t-\gamma)\mathbf{I}_t + \gamma\mathbf{J}_t,$$

where $\mathbf{I}_t$ is the $t \times t$ identity matrix and $\mathbf{J}_t$ is a $t \times t$ matrix of ± 1's whose diagonal entries are all 1's. If $\gamma = 2$ and $\mathbf{J}_t$ is the $t \times t$ matrix of 1's, then the $2t \times 2t$ matrix

$$\mathbf{C}_2 = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{B}^\top & \mathbf{A}^\top \end{pmatrix} \tag{9}$$

has the maximum determinant (see, Ehlich 1964) among all the $2t \times 2t$ $\{\pm 1\}$-matrices, i.e., D-optimal.

Theorem 4: Let $N = 2t$, t odd and $\mathbf{a}, \mathbf{b}$ two vectors of length t with entries from $\{\pm 1\}$ satisfying equation (8) for $\gamma_s \in \{-2, 2\}$. Let $a = \sum_i a_i$ and $b = \sum_i b_i$. Then an unbalanced, two-symbol, $\mathrm{E}(s^2)$-optimal and minimax-optimal SSD with N rows and $m = N$ columns achieving the lower bound $\mathrm{LB}(N, (0, \ldots, 0, k_{t-(a+b)/2} = t, 0, \ldots, 0, k_{t-(a-b)/2} = t, 0, \ldots, 0)$ can be constructed. If $\gamma_s = 2$ for $s = 1, 2, \ldots, t-1$, then the constructed design is also D-optimal.

Proof: Let $\mathbf{J}_t$ be a $t \times t$ matrix of ± 1's. Use circulant matrices $\mathbf{A}$ and $\mathbf{B}$ given in (9). Then $\mathbf{C}_2$ in (9) satisfies

$$\mathbf{C}_2\mathbf{C}_2^\top = \mathbf{C}_2^\top\mathbf{C}_2 = \begin{pmatrix} (2t-2)\mathbf{I}_t + 2\mathbf{J}_t & \mathbf{0}_{t\times t} \\ \mathbf{0}_{t\times t} & (2t-2)\mathbf{I}_t + 2\mathbf{J}_t \end{pmatrix}.$$

Thus, $\max_{i<j}|s_{ij}| = 2$, and $\mathbf{C}_2$ achieves the naive $\mathrm{E}(s^2)$ lower bound (5). Moreover, if $\mathbf{J}_t$ is the $t \times t$ matrix of 1's, then $\mathbf{C}_2$ is D-optimal.

Since $\mathbf{a} = (a_0, a_1, \ldots, a_{t-1})$ and $\mathbf{b} = (b_0, b_1, \ldots, b_{t-1})$ satisfy equation (8), t should be odd with $2(2t-1) = a^2 + b^2$. WLOG, we may assume that $a \geq b > 0$. Also, 1 and -1 appear $(t+a)/2$ and $(t+b)/2$ times in $\mathbf{a}$ and $(t-a)/2$ and $(t-b)/2$ times in $\mathbf{b}$. Each of the first t columns of the generated SSD by construction (9) has a column sum of $(a-b)$. So, the number of -1's in each of the first t columns is $l_2 = t - (a-b)/2$. Similarly, each of the last t columns has a column sum of $(a+b)$. So, the number of -1's in each of the last t columns is $l_1 = t - (a+b)/2$. Thus, $k_{l_2} = k_{t-(a-b)/2} = t$, $k_{l_1} = k_{t-(a+b)/2} = t$.

In the examples below, we denote 1 by + and -1 by $-$.

Example 2: Let

$$\begin{aligned} \mathbf{a} &= (+,+,+,+,+,+,+,+,-,+,+,-,-) \\ \mathbf{b} &= (+,+,+,-,-,-,+,-,+,+,-,+,-). \end{aligned}$$

Vectors $\mathbf{a}$ and $\mathbf{b}$ satisfy $P_{\mathbf{a}}(s) + P_{\mathbf{b}}(s) = 2$, $s = 1, 2, \ldots, t-1$. Moreover, $a = \sum_i a_i = 7$ and $b = \sum_i b_i = 1$. Let $\mathbf{A} = circ(\mathbf{a})$ and $\mathbf{B} = circ(\mathbf{b})$. Using construction (9) we obtain $\mathbf{D}$ such that $\mathbf{D} \in \mathcal{D}^{\pm}(26, 2^{26}, (k_9 = 13, k_{10} = 13))$ and $\mathrm{E}(s^2) = \mathrm{LB} = \mathrm{LB}_2 = 1.92$. Hence, $\mathbf{D}$ is an unbalanced, $\mathrm{E}(s^2)$-optimal, minimax-optimal SSD and a D-optimal design.

Example 3. Let

$$\begin{aligned} \mathbf{a} &= (+,+,+,+,+,-,+,-,-,+,+,+,-) \\ \mathbf{b} &= (+,+,+,+,+,-,+,-,-,+,+,+,-). \end{aligned}$$

Vectors $\mathbf{a}$ and $\mathbf{b}$ satisfy $P_{\mathbf{a}}(s)+P_{\mathbf{b}}(s) = 2$ for $s = 1, 2, \ldots, t-1$. Moreover, $a = \sum_i a_i = \sum_i b_i = 5$. Let $\mathbf{A} = circ(\mathbf{a})$ and $\mathbf{B} = circ(\mathbf{b})$. Using construction (9), we obtain $\mathbf{D}$ such that $\mathbf{D} \in \mathcal{D}^{\pm}(26, 2^{26}, (k_8 = 13, k_{13} = 13))$ and $\mathrm{E}(s^2) = \mathrm{LB} = \mathrm{LB}_2 = 1.92$. Hence, $\mathbf{D}$ is an unbalanced, $\mathrm{E}(s^2)$-optimal, minimax-optimal SSD and a D-optimal design.

4. Testing Achievability Computationally

We implemented a computational study to test the achievability of $\mathrm{LB}(N, (k_0, \ldots, k_{\lfloor N/2 \rfloor}))$. In searching for an SSD $\mathbf{D} \in \mathcal{D}^{\pm}(N, 2^m, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$ achieving $\mathrm{LB}(N, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$, we assumed that $k_0 = 1$. First we show that this assumption can be made without losing generality. To do this, we need the following two lemmas.

Lemma 5: Let $\mathbf{D} \in \mathcal{D}^{\pm}(N, 2^m, (0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$. Then $[\mathbf{1}_N\, \mathbf{D}] \in \mathcal{D}^{\pm}(N, 2^{(m+1)}, (1, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$, and $\mathbf{D}$ is $\mathrm{E}(s^2)$-optimal if and only if $[\mathbf{1}_N\, \mathbf{D}]$ is $\mathrm{E}(s^2)$-optimal.

Proof: Observe that

$$N^2 + \mathrm{SS}(\mathbf{D}^\top\mathbf{D}) + 2\sum_l k_l(N-2l)^2 = \mathrm{SS}([\mathbf{1}_N\, \mathbf{D}]^\top[\mathbf{1}_N\, \mathbf{D}]).$$

Hence, $\mathrm{SS}(\mathbf{D}^\top\mathbf{D})$ is minimized if and only if $\mathrm{SS}([\mathbf{1}_N\,\mathbf{D}]^\top[\mathbf{1}_N\,\mathbf{D}])$ is minimized. The result follows because for any SSD $\mathbf{X} \in \mathcal{D}^{\pm}(N', 2^{m'}, (k'_0, \ldots, k'_{\lfloor N'/2 \rfloor}))$ we have

$$\mathrm{E}(s^2) = \frac{(\mathbf{X}^T\mathbf{X}) - N'^2m'}{m'(m'-1)}.$$

Hence, $\mathbf{X}$ is $\mathrm{E}(s^2)$-optimal if and only if $\mathrm{SS}(\mathbf{X}^\top\mathbf{X})$ is minimized.

Lemma 6: Let $k_0 = 0$ and $\sum_l k_l = m$. Then

$$(m+1)m\,\mathrm{LB}(N, (1, k_1, \ldots, k_{\lfloor \frac{N}{2} \rfloor})) = m(m-1)\,\mathrm{LB}(N, (0, k_1, \ldots, k_{\lfloor \frac{N}{2} \rfloor})) + 2\sum_l k_l(N-2l)^2.$$

Proof: Each $\mathrm{LB}_i(N, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$ for $i = 1, 2$ is derived based on a hypothetical $\mathbf{D}_i^* \in \mathcal{D}^{\pm}(N, 2^m, (0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$, where

$$\mathrm{LB}_i(N, (k_0, k_1, \ldots, k_{\lfloor \frac{N}{2} \rfloor})) = \frac{\mathrm{SS}(\mathbf{D}_i^{*\top}\mathbf{D}_i^*) - N^2m}{m(m-1)}.$$

If $\mathbf{D}_i^* \in \mathcal{D}^{\pm}(N, 2^m, (0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$ then $[\mathbf{1}_N\,\mathbf{D}_i^*] \in \mathcal{D}^{\pm}(N, 2^{(m+1)}, (1, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$. Now,

$$\begin{aligned}
\mathrm{LB}_i(N, (0, k_1 \ldots, k_{\lfloor \frac{N}{2} \rfloor})) &= \frac{\mathrm{SS}(\mathbf{D}_i^{*\top}\mathbf{D}_i^*) - N^2m}{m(m-1)}, \\
\mathrm{LB}_i(N, (1, k_1, \ldots, k_{\lfloor \frac{N}{2} \rfloor})) &= \frac{\mathrm{SS}([\mathbf{1}_N\,\mathbf{D}_i^*]^\top[\mathbf{1}_N\,\mathbf{D}_i^*]) - N^2(m+1)}{(m+1)m}.
\end{aligned}$$

Hence, by Lemma 5

$$(m+1)m\,\mathrm{LB}_i(N, (1, k_1, \ldots, k_{\lfloor \frac{N}{2} \rfloor})) = m(m-1)\,\mathrm{LB}_i(N, (0, k_1, \ldots, k_{\lfloor \frac{N}{2} \rfloor})) + 2\sum_l k_l(N-2l)^2.$$

Now, the result follows from

$$\mathrm{LB}(N, (k_0, k_1, \ldots, k_{\lfloor \frac{N}{2} \rfloor})) = \max_{i \in \{1,2\}} \{\mathrm{LB}_i(N, (k_0, k_1, \ldots, k_{\lfloor \frac{N}{2} \rfloor}))\}$$

and

$$\begin{aligned}
&(m+1)m \max_{i \in \{1,2\}} \mathrm{LB}_i(N, (1, k_1, \ldots, k_{\lfloor \frac{N}{2} \rfloor})) = \\
&m(m-1) \max_{i \in \{1,2\}} \mathrm{LB}_i(N, (0, k_1, \ldots, k_{\lfloor \frac{N}{2} \rfloor})) + 2\sum_l k_l(N-2l)^2.
\end{aligned}$$

The definition of an SSD $\mathbf{D} \in \mathcal{D}^{\pm}(N, 2^m, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$ requires that it has no pair of *aliased* columns, i.e., columns $\mathbf{d}_i$ and $\mathbf{d}_j$ such that $\mathbf{d}_i^\top\mathbf{d}_j \in \{-N, N\}$. Hence, we must have $k_0 \leq 1$ for each SSD $\mathbf{D}$.

Theorem 5: An SSD $\mathbf{D}$ achieving $\mathrm{LB}(N, (0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$ exists if and only if an SSD $[\mathbf{1}_N\,\mathbf{D}]$ achieving $\mathrm{LB}(N, (1, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$ exists.

Proof: The result follows immediately from Lemmas 5 and 6.

By Theorem 5, WLOG, we can restrict our search for SSDs achieving $\mathrm{LB}(N, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$ to SSDs with $k_0 = 1$. Accordingly, we wrote a C program for the NOA_p for $p = 2, 4, 8$ algorithms together with the derived $\mathrm{E}(s^2)$ lower bound $\mathrm{LB}(N, (1, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$ to search for SSDs in $\mathcal{D}^{\pm}(N, 2^m, (1, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$ achieving this bound.

There is a one-to-one correspondence between all SSDs achieving $\mathrm{LB}(N, (1, 0, \ldots, 0, m-1))$ and all balanced SSDs achieving $\mathrm{LB}(N, (0, \ldots, 0, m-1))$ obtained by deleting the all ones column. The NOA_p for $p = 2, 4, 8$ algorithms could find a balanced $\mathrm{E}(s^2)$-optimal SSD for $N = 14, 16$ in all cases except the $N = 14, m = 16$ case (Bulutoglu and Ryan, 2008). For the cases with $N \in \{14, 16\}$ and $15 \leq m \leq 70$ for which $\mathrm{LB}(N, (0, \ldots, 0, m-1))$ is (not) equal to $\mathrm{E}(s^2)$-lower bound in Bulutoglu and Ryan (2008), we were (not) able to find an SSD in $\mathcal{D}^{\pm}(N, 2^m, (1, 0, \ldots, 0, m))$ achieving $\mathrm{LB}(N, (1, 0, \ldots, 0, m))$ except for the $N = 14, m = 17$ case (this case with the all 1's column corresponds to the $N = 14, m = 16$ case in Ryan and Bulutoglu (2007). These observations confirm the correctness of our C program.

For each of the number of rows N and number of columns m combinations in Table 1 we randomly generated 100 vectors $(1, k_1, \ldots, k_{\lfloor N/2 \rfloor})$ such that $k_0 = 1$ and $\sum_l k_l = m$. Then for each of these 100 vectors $(1, k_1, \ldots, k_{\lfloor N/2 \rfloor})$, we used the NOA_4 exchange algorithm (Bulutoglu and Ryan, 2008; Ryan and Bulutoglu, 2007) to search for an SSD $\mathbf{D} \in \mathcal{D}^{\pm}(N, 2^m, (1, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$ achieving $\mathrm{LB}(N, (1, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$. The complexity of running the NOA_p algorithm increases with p. However, for each random starting design, increased p increases the probability of converging to an SSD which has no aliased columns by definition. After experimenting with NOA_2, NOA_4, and NOA_8 we decided to use NOA_4 as a compromise between speed and avoidance of converging to a design with aliased columns. In Table 1, the column N reports N, the column m reports a range of m for which this experiment was conducted, and the column *numiter* reports the number of random starting designs that were used each time the NOA_4 algorithm was run. (We changed *numiter* only with N.) For each N, m combination such that m is within the reported range of m, the numbers of successes column of Table 1 reports the number of times out of 100 sampled vectors $(1, k_1, \ldots, k_{\lfloor N/2 \rfloor})$ an SSD achieving $\mathrm{LB}(N, (1, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$ was found. The number of successes in Table 1, which are in fact each a percent out of 100 sampled vectors $(1, k_1, \ldots, k_{\lfloor N/2 \rfloor})$, can only underestimate the true percentage of the sampled vectors $(1, k_1, \ldots, k_{\lfloor N/2 \rfloor})$ where our bound is achievable. Hence, for each vector $(1, k_1, \ldots, k_{\lfloor N/2 \rfloor})$, we needed a sufficient number of random starting designs to avoid significantly underestimated true percentages. We observed that for the balanced cases of $N = 10$ and $N = 14$, $numiter = 10^6$ and $numiter = 10^7$ were sufficient. (We were able to determine this, as for the balanced cases we know exactly when $\mathrm{LB}(N, (0, \ldots, 0, m))$ is achievable.) Hence, we set $numiter = 10^7$ for $11 \leq N \leq 14$ and $numiter = 10^6$ for $N \leq 10$. However, just because these values of *numiter* are sufficient in the balanced cases does not guarantee that they will be sufficient for the corresponding unbalanced cases.

Our Table 1 estimates do decrease with increased N. The output of our computational experiments also provided us with the iteration number at which an $\mathrm{E}(s^2)$-optimal SSD achieving $\mathrm{LB}(N, (1, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$ was found. We used this information to perform a statistical analysis to determine the significance of underestimation. Our statistical analysis suggests that the true percentage of the sampled vectors $(1, k_1, \ldots, k_{\lfloor N/2 \rfloor})$ where our bound is achievable $\mathrm{LB}(N, (1, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$

Table 1: Numbers of times an SSD achieving $\mathrm{LB}(N,(1,k_1,\ldots,k_{\lfloor N/2\rfloor}))$ **found for a randomly generated set of** 100 **vectors** $(1,k_1,\ldots,k_{\lfloor N/2\rfloor})$ **such that** $1+\sum_l k_l = m$

N	m	*numiter*	numbers of successes each out of 100
7	7-20	10^6	17 52 34 36 68 68 55 41 67 72 55 51 45 69
8	8-35	10^6	12 18 10 44 57 45 50 63 76 65 58 59 69 83 64 68 67 75 69 46 45 56 55 45 45 53 40 30
9	9-69	10^6	4 4 15 37 45 33 31 47 54 76 61 68 74 75 70 63 63 66 82 72 56 60 69 65 52 45 52 50 37 37 40 44 34 30 33 23 28 18 18 22 15 14 7 12 17 7 3 5 9 4 2 3 5 1 1 2 2 5 0 0 2
10	10-63	10^6	2 4 25 33 18 11 45 64 57 52 54 61 66 58 53 68 59 63 60 67 61 62 59 54 58 59 61 59 59 47 55 55 52 49 45 50 46 38 34 34 36 34 27 30 24 16 26 28 21 26 18 11 13 13
11	11-50	10^7	2 8 20 8 2 13 29 56 52 41 36 42 46 54 43 54 58 58 53 53 53 56 53 49 55 46 50 50 46 41 40 38 39 39 36 29 30 34 36 33
12	12-43	10^7	0 0 0 0 12 36 56 32 18 21 28 57 28 27 35 48 39 32 26 32 32 31 32 32 33 29 34 30 25 33 30 22
13	13-35	10^7	0 0 0 3 14 27 20 2 4 12 31 25 21 17 18 29 27 27 28 18 36 28 20
14	14-35	10^7	0 0 3 12 2 1 0 4 7 9 6 1 4 4 7 5 1 3 6 4 3 3

is underestimated significantly only in the $N \geq 13$ rows of Table 1.

5. Application

In this section we discuss two possible applications of our newly derived $\mathrm{E}(s^2)$ lower bound.

5.1. Application to searching for D-optimal designs

Let $\mathbf{D}$ be an $N \times m$ ($m \leq N$) matrix with entries from $\{\pm 1\}$ representing an N row, two-symbol and m column array. Let $\lambda_1, \lambda_2, \ldots, \lambda_m$ be all the non-zero eigenvalues of $\mathbf{D}^\top\mathbf{D}$, where $\mathrm{rank}(\mathbf{D}^\top\mathbf{D}) = m$. By the spectral decomposition theorem,

$$\begin{aligned}
\sum_i \lambda_i &= \mathrm{Tr}(\mathbf{D}^\top\mathbf{D}) = Nm,\\
\sum_i \lambda_i^2 &= \mathrm{Tr}((\mathbf{D}^\top\mathbf{D})^2) = \mathrm{Tr}((\mathbf{D}\mathbf{D}^\top)^2) = \mathrm{SS}(\mathbf{D}^\top\mathbf{D}) = \mathrm{SS}(\mathbf{D}\mathbf{D}^\top),\\
\mathrm{Det}(\mathbf{D}^\top\mathbf{D}) &= \prod_i \lambda_i.
\end{aligned}$$

If $\sum_i \lambda_i^2 = \theta$, then $\prod_i \lambda_i \leq d(\theta)$ for some $d(\theta) \geq 0$, where θ is some positive integer. This is equivalent to $-\log(\prod_i \lambda_i) \geq -\log(d(\theta))$ by the monotonicity of the $-\log(\cdot)$ function. To find such $d(\theta)$ that is as small as possible we consider the following smooth, non-convex nonlinear

programming (NLP) problem

$$\begin{aligned} \min \ & \sum_i -\log(\lambda_i) \\ \text{subject to: } & \sum_i \lambda_i = Nm, \\ & \sum_i \lambda_i^2 = \theta, \quad \lambda_i \geq 0. \end{aligned} \tag{10}$$

NLP (10) was solved analytically in Cheng (1978). Cheng (1978) showed that the minimum is attained at a point which has constant coordinates $\lambda^* = N$ when $\theta = N^2m$ or has two distinct coordinates $\lambda_1^* > \lambda_2^* > 0$ when $\theta > N^2m$, where λ_1^* has multiplicity n, λ_2^* has multiplicity $m-n$, and

$$\begin{aligned} \lambda_1^* &= N + \sqrt{\frac{(m-n)(\theta - N^2m)}{mn}}, \\ \lambda_2^* &= N - \sqrt{\frac{n(\theta - N^2m)}{(m-n)m}}. \end{aligned}$$

Let

$$d(\theta, N, m, n) = \left(N + \sqrt{\frac{(m-n)(\theta - N^2m)}{mn}}\right)^n \left(N - \sqrt{\frac{n(\theta - N^2m)}{(m-n)m}}\right)^{m-n}.$$

Cheng (1978) also showed that $d(\theta, N, m, n)$ is a strictly decreasing function of n. For $\theta > N^2m$, this result implies that

$$d(\theta) = d(\theta, N, m, 1) = \left(N + \sqrt{\frac{(m-1)(\theta - N^2m)}{m}}\right)\left(N - \sqrt{\frac{(\theta - N^2m)}{(m-1)m}}\right)^{m-1}$$

is a valid upper bound for $\mathrm{Det}(\mathbf{D}^\top\mathbf{D}) = \prod_i \lambda_i$. Then for fixed N and m, by differentiating $\log(d(\theta))$, we see that $d(\theta)$ is a strictly decreasing function of θ for $\theta > 0$. Hence, we get the following theorem.

Theorem 6: Let $\theta \in \mathbb{Z}$ be such that $\theta > N^2m$. Then

$$\mathrm{SS}(\mathbf{D}^\top\mathbf{D}) = \mathrm{SS}(\mathbf{D}\mathbf{D}^\top) = \sum_i \lambda_i^2 \geq \theta,$$

implies

$$\mathrm{Det}(\mathbf{D}^\top\mathbf{D}) = \prod_i \lambda_i \leq d(\theta).$$

The following example shows how Theorem 6 can be used to derive upper bounds for the $\mathrm{SS}(\mathbf{D}^*(\mathbf{D}^*)^\top)$, where $\mathbf{D}^*$ is a two-symbol, D-optimal design.

Example 4: For $N = 22$ and $m = 22$, the largest possible $\mathrm{Det}(\mathbf{D}^\top\mathbf{D})$ of $20^{12} \times (6400000)^2$ for a two-symbol design with entries from $\{\pm 1\}$ is given by Chasiotis *et al.* (2018). Then by Theorem 6, for a D-optimal design $\mathbf{D}^*$ with 22 rows and 22 columns, we must have

$$\mathrm{SS}((\mathbf{D}^*)^\top\mathbf{D}^*) = \mathrm{SS}(\mathbf{D}^*(\mathbf{D}^*)^\top) \leq 11,920.$$

This proves that a D-optimal design in this case cannot be balanced. Since for a balanced 22 row and 22 column design $\mathbf{D}$, $\mathrm{SS}(\mathbf{D}^\top\mathbf{D}) \geq 12,496$.

Next, we provide a connection between D-optimal designs and the $\mathrm{E}(s^2)$ lower bound that we derived in Section 2. Let $\mathbf{D} = (d_{ij})$ be a sought after D-optimal design with N rows and m columns and $\mathbf{d}_1^\top$ be the first row of $\mathbf{D}$. Let

$$\hat{\mathbf{D}} = \mathbf{D}\mathrm{diag}(\mathbf{d}_1),$$

where $\mathrm{diag}(\mathbf{d}_1)$ is the $m \times m$ diagonal matrix whose (j, j)'th entry is the j'th entry of $\mathbf{d}_1$. Now, $\hat{\mathbf{D}}^\top$ can be viewed as a not necessarily balanced SSD. Let $\mathrm{E}(s^2)$ be the $\mathrm{E}(s^2)$ value of $\hat{\mathbf{D}}^\top$. Then the $\mathrm{E}(s^2)$ values of $\hat{\mathbf{D}}^\top$ satisfies

$$\mathrm{SS}(\hat{\mathbf{D}}^\top\hat{\mathbf{D}}) = \mathrm{SS}(\hat{\mathbf{D}}\hat{\mathbf{D}}^\top) = \mathrm{E}(s^2)N(N-1) + Nm^2.$$

Hence we get the following corollary to Theorem 6.

Corollary 1: Let $\theta \in \mathbb{Z}$ be such that $\theta > N^2m$. Then

$$\mathrm{E}(s^2) \geq \frac{\theta - Nm^2}{N(N-1)}$$

implies

$$\mathrm{Det}(\mathbf{D}^\top\mathbf{D}) = \mathrm{Det}(\hat{\mathbf{D}}^\top\hat{\mathbf{D}}) \leq d(\theta).$$

Now for a given value of $d(\theta)$, when searching for a $\mathbf{D}$ with

$$\mathrm{Det}(\mathbf{D}^\top\mathbf{D}) > d(\theta),$$

the search can be restricted to $\hat{\mathbf{D}}^\top$ with balancedness structure $(1, k_1, \ldots, k_{\lfloor m/2 \rfloor})$ such that $\hat{\mathbf{D}}^\top \in \mathcal{D}^{\pm}(m, 2^N, (1, k_1, \ldots, k_{\lfloor m/2 \rfloor}))$ and

$$\mathrm{LB}(m, (1, k_1, \ldots, k_{\lfloor \frac{m}{2} \rfloor})) \leq \frac{\theta - Nm^2}{N(N-1)}. \tag{11}$$

The θ in (11) can be decreased by using the methods of Chasiotis *et al.* (2018). This restriction should decrease the search space significantly and improve algorithm performance for finding D-optimal designs in Brent *et al.* (2011).

Requiring an SSD to be minimax optimal has the benefit of reducing the search space and can be useful in the search for a D-optimal design. In fact, there is a 22 row, 22 column, D-optimal design that can be viewed as an unbalanced, $\mathrm{E}(s^2)$-optimal, minimax-optimal SSD achieving the naive $\mathrm{E}(s^2)$ lower bound (5), (see Chasiotis *et al.* 2018). However, finding even a balanced $\mathrm{E}(s^2)$-optimal and minimax-optimal SSD in general is a very difficult problem (Morales and Bulutoglu, 2018).

5.2. Application to finding upper bounds on the maximum number of columns

The following is an important theoretical problem in the SSD literature (Cheng and Tang, 2001).

Problem 1: For a given $0 \le t \le N$ find the maximum number of columns $B(N, t, (0, \ldots, 0, m))$ such that an SSD $\mathbf{D} \in \mathcal{D}^{\pm}(N, 2^m, (0, \ldots, 0, m))$ with $s_{\max} \le t$ exists. The generalization of this problem to unbalanced SSDs is determining $B(N, t, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$. Multiplying a subset of rows of an SSD $\mathbf{D}$ does not change the $s_{\max}$ of $\mathbf{D}$. Hence, by Lemma 1 in (Cheng and Tang, 2001), WLOG it suffices to find $B(N, t, (1, k'_1, \ldots, k'_{\lfloor N/2 \rfloor}))$, where $k'_i = 0$ for $1 \le i < (N - t)/2$.

By using $\mathrm{E}(s^2)$ lower bounds on balanced SSDs, Cheng and Tang (2001) found an upper bound on $B(N, t, (0, \ldots, 0, m))$. Our newly derived $\mathrm{E}(s^2)$ lower bounds for unbalanced SSDs can be used to generalize the upper bound on $B(N, t, (0, \ldots, 0, m))$ to an upper bound on $B(N, t, (1, k'_1, \ldots, k'_{\lfloor N/2 \rfloor}))$ with $k'_i = 0$ for $1 \le i < (N - t)/2$.

The following is an important theoretical problem in the frame theory literature (Szöllösi and Östergård, 2018).

Problem 2: For given $a \in (0, 1)$ and $d \in \mathbb{Z}^{\ge 0}$, find the maximum number of equiangular lines in $\mathbb{R}^d$ with pairwise angle $\arccos(a)$.

For an SSD $\mathbf{X} \in \mathcal{D}^{\pm}(N, 2^m, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$, let $s_{\min} = \min_{i<j} |s_{ij}|$. Let $C(N, t, d)$ be the maximum m such that an SSD $\mathbf{X} \in \mathcal{D}^{\pm}(N, 2^m, (k_0, k_1, \ldots, k_{\lfloor N/2 \rfloor}))$ with $t = s_{\max} = s_{\min}$ and $\mathrm{rank}(\mathbf{X}) = d \le N$ exists. Since multiplying a subset of rows of an SSD $\mathbf{D}$ does not change the $s_{\max}$ and $s_{\min}$ of $\mathbf{D}$, WLOG we can assume $\mathbf{X} \in \mathcal{D}^{\pm}(N, 2^m, (1, k'_1, \ldots, k'_{\lfloor N/2 \rfloor}))$, where $k'_{(N-t)/2} = m - 1$ and $k'_i = 0$ for $i \notin \{0, (N - t)/2\}$. Hence, it is plain to see that $C(N, t, d) \le B(N, t, (1, k'_1, \ldots, k'_{\lfloor N/2 \rfloor}))$.

The number $C(N, t, d)$ is a lower bound on the number of equiangular lines in $\mathbb{R}^d$ with pairwise angle $\arccos(t/N)$. Hence, determining $C(N, t, d)$ provides information on the solution of Problem 2 for $a = t/N$ and $B(N, t, (1, k'_1, \ldots, k'_{\lfloor N/2 \rfloor}))$, where $k'_{(N-t)/2} = m - 1$ and $k'_i = 0$ for $i \notin \{0, (N - t)/2\}$ bounds $C(N, t, d)$. We propose determining $C(N, t, d)$ for many N, t, d combinations by using the upper bound $B(N, t, (1, k'_1, \ldots, k'_{\lfloor N/2 \rfloor}))$ as needed as a future research project.

Acknowledgements

The research of Dr. Dursun A. Bulutoglu was supported by the Air Force Office of Scientific Research grant F4FGA08087J006.

The views expressed in this article are those of the authors and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.

References

Brent, R. P., Orrick W. P., Osborn, J. and Zimmermann, P. (2011). Maximal determinants and saturated d -optimal designs of orders 19 and 37. *Arxiv*. *http://arxiv.org/pdf/ 1112.4160.pdf*

Bulutoglu, D. A. and Ryan, K. J. (2008). $E(s^2)$-optimal supersaturated designs with good minimax properties when n is odd. *Journal of Statistical Planning and Inference*, **138**, 1754-1762.

Chasiotis, V., Kounias, S. and Farmakis, N. (2018). The d-optimal saturated designs of order 22. *Discrete Mathematics*, **341**, 380–387.

Cheng, C. S. (1978). Optimality of certain asymmetrical experimental designs. *Annals of Statistics*, **6**, 1239–1261.

Cheng, C. S. and Tang, B. (2001). Upper bounds on the number of columns in supersaturated designs. *Biometrika*, **88**, 1169–1174.

Das, A., Dey, A., Chan, L-Y. and Chatterjee, K. (2008). $E(s^2)$- optimal supersaturated designs. *Journal of Statistical Planning and Inference*, **138**, 3749–3757.

Ehlich, H. (1964). Determinantenabschätzung für binäre matrizen. *Mathematische Zeitschrift*, **83**, 123–132.

Georgiou, S. D. (2014). Supersaturated designs: A review of their construction and analysis. *Journal of Statistical Planning and Inference*, **144**, 92–109.

Koukouvinos, C. and Stylianou, S. (2008). On skew-hadamard matrices. *Discrete Mathematics*, **308**, 2723–2731.

Morales, L. B. and Bulutoglu, D. A. (2018). On $E(s^2)$-optimal and minimax-optimal supersaturated designs with 20 rows and 76 columns. *Journal of Combinatorial Designs*, **26**, 344–355.

Morales, L. B., Bulutoglu, D. A. and Arasu, K. T. (2019). The maximum number of columns in supersaturated designs with $s_max = 2$. *Journal of Combinatorial Designs*, **27**, 448–472.

Ryan, K. J. and Bulutoglu, D. A. (2007). $E(s^2)$-optimal supersaturated designs with good minimax properties. *Journal of Statistical Planning and Inference*, **137**, 2250–2262.

Suen, C. S. and Das, A. (2010). $E(s^2)$-optimal supersaturated designs with odd number of runs. *Journal of Statistical Planning and Inference*, **140**, 1398–1409.

Szöllösi, F. and Östergård, P. R. J. (2018). Enumeration of seidel matrices. *European Journal of Combinatorics*, **69**, 169–184.

Waldron, S. (2009). On the construction of equiangular frames from graphs. *Linear Algebra and its Applications*, **431**, 2228–2242.

Statistics and Applications {ISSN 2454–7395(online)}
Volume 19, No. 1, 2021 (New Series), pp 29–40

Randomized Block Designs, Balanced Incomplete Block Designs and Latin Square Designs with Neighbor Effects in the Presence of Covariates

Sobita Sapam[1], K. K. Singh Meitei[1] and Bikas K. Sinha[2]
[1] *Department of Statistics, Manipur University, Canchipur, India*
[2] *Indian Statistical Institute, Kolkata [Retired Faculty], India*

Received: 06 May, 2020; Revised: 06 July, 2020; Accepted: 09 July, 2020

Abstract

The research work undertaken in this paper is motivated by a real life scenario in the context of agricultural experiments. It is believed that the neighboring 'plots' in a Block Design or in a Latin Square Design [LSD] tend to influence each other in terms of the mean yield through the 'neighboring effects of the treatments' applied in these plots. Further to this, there are quantifiable and controllable covariates acting linearly in the mean model. We contemplate a linear ANCOVA model and study its analysis - with special emphasis on the question of estimability of the regression coefficient(s) involving the covariates. We focus on RBDs with $b = v = 4$, on an $SBIBD(7, 7, 4, 4, 2)$ and also on an LSD of order 4.

Key words: Randomized block designs; Balanced incomplete block designs; Latin square designs; Direct treatment effects; Neighbor treatment effects; Left neighbors; Right neighbors; Top neighbors; Bottom neighbors; Linear ANCOVA model; Covariates; Optimal covariate matrices.

1. Introduction

The key reference to this article is Springer Publication by Das *et al.* (2015) titled "Optimal Covariate Designs". Generally speaking, in the context of an experimental design with covariates, each experimental unit is supposed to have attached to it a number of quantifiable and measurable covariates. Assuming that there is a large pool of units, we have a choice for selection of the units with assigned covariate-values. Optimal covariate designs are the designs which provide optimal or most efficient estimation of the covariates' effects in terms of the parameters in an assumed linear model. The experimental set-up is quite general - starting with CRDs, RBDs, BIBDs, LSDs etc. The number of covariates need not be just one or two. Optimality problems center around characterization and constructions of designs *i.e.*, choice of experimental units with 'optimally assigned' covariate values in a given experimental set-up. The reader is referred to Das *et al.*(2015) for details. This area of research grew over the last 40 years or so.

Correponding Author: Bikas K. Sinha
Email: bikassinha1946@gmail.com

Covariate Models or ANCOVA Models are seen as a 'blend' of 'design model' and 'regression model'. In a block design set-up, writing y_{ij} for the observation in the experimental unit corresponding to i-th block and j-th treatment, we may write the model as

$$y_{ij} = \mu + \beta_i + \tau_j + \beta_1 x_{1;(i,j)} + \beta_2 x_{2;(i,j)} + \ldots + e_{ij},$$

where it is assumed that $x_{1;(i,j)}, x_{2;(i,j)}, \ldots$ are the covariate values attached to the unit labelled (i, j) with associated linear effects parameters $\beta_1, \beta_2, \ldots$. Our purpose is to identify and select those experimental units which collectively provide optimal estimation of the covariate parameters *i.e.*, of the β's. Note that the design set- up could be very much general in nature. However, unless there is a nice combinatorial structure of the underlying design [without the covariate effects], the problem, in its most general form, is untraceable. That is why only CRDs, RBDs, BIBDs, LSDs etc have been studied in the literature. The complexity of the problem cannot be undermined if there are a number of covariates. In general terms, for any number of covariates and any experimental design set- up, it transpires that $Var(\hat{\beta}) \geq \sigma^2 / \sum x^2_{(i,j)}$. It can be argued that, without any loss of generality, we may assume $-1 \leq x_{(i,j)} \leq 1$. This takes the variance bound to σ^2/n where n is the total number of observations. We need to examine the case of 'equality' and that too, for each of the covariates and there again, we need to attain 'equality' simultaneously for all the covariates parameters' estimates. Note that we are examining the status of a design only wrt the β-parameters, ignoring other fixed-effects parameters in the model. Anyway, there are too many issues involved and, without any further digression, we refer to Das *et al.* (2015).

Specifically, if we are dealing with an RBD involving b blocks and v treatments and if there are k covariates $(X_{(1)}, X_{(2)}, \ldots, X_{(k)})$, we will attain 'equality' in the variance bound simultaneously for all the covariates if and only if the following conditions are met :

$$(i) \sum_j x_{(u;(i,j))} = 0, \quad 1 \leq i \leq b; \quad (ii) \sum_i x_{(u;(i,j))} = 0, \quad 1 \leq j \leq v;$$

$$(iii) \sum_{1 \leq i \leq b} \sum_{1 \leq j \leq v} x_{u;(i,j)} x_{u^*;(i,j)} = nI(u, u^*); \quad 1 \leq u, u^* \leq k.$$

where, in the above, $I(..)$ is the usual indicator function and $n = bv$.

In this paper we will deal with an $RBD(b = 4, v = 4)$, a $BIBD(7, 7, 4, 4, 2)$ and an LSD of order 4. Moreover, we will adopt a model where, besides the block effects/row- column effects and treatment effects, we also have neighbor effects - designated as Left-Neighbor (LN)-Effects, Right-Neighbor (RN)-Effects etc. Naturally, we will require more conditions to be satisfied by the collection of the $x_{(u;i,j)}$'s. Note that (iii) requires that $x_{(u;(i,j))} = +1/-1$ for all choices of $_{(u;(i,j)})$'s. With this background, we will proceed to derive/present the results on optimal covariates designs in a model with N-Effects. In doing so, our target will be to cover maximum number of such covariates with most efficient estimation for each one. Once for all, we refer to systematic study of four- sided RN- and CN- effects as proposed and discussed in Varghese *et al.* (2014) for an explanation of neighbor effects. There are two

follow- up papers in this direction as well. [Sapam *et al.* (2019a, 2019b)]. We may mention another related paper by Jaggi *et al.* (2018).

2. RBD with b=v=4

We start with the following RBD in Table 1 wherein we also display the Left-sided and Right-sided Neighbor Effects, assuming a circular model. [Vide Kunert (1984)].

Table 1: *RBD* with b=v=4: First Choice

LN 4	1	2	3	4	RN 1
LN 1	2	3	4	1	RN 2
LN 2	3	4	1	2	RN 3
LN 3	4	1	2	3	RN 4

We assume the existence of a controllable and quantifiable covariate (X) attached to every plot in the block design. We denote by x_{ij} the value of the covariate attached to the plot labelled (i, j) which corresponds to plot number i in block number j; $i, j = 1, 2, 3, 4$.

Without any loss of generality, we further assume that $-1 \leq x_{ij} \leq 1$ for each of the covariate values.

Under the assumed linear model, it follows that $I(\beta) \leq \sum\sum x_{ij}^2 \leq bv = 16$, dropping the error variance σ^2 in the model. The case of 'equality' has been studied earlier in our papers in easier settings. We refer to Das *et al.* (2015) for details. However, the present setting is a bit complicated since there are block effects, (direct) treatment effects and both LN- and RN- Effects of the treatments. Consider the following $X_{(1)}$-matrix in Table 2 for one choice of the covariate values.

Table 2: Covariate matrix for *RBD* with v=b=4 in Table 1

$$X_{(1)} = \begin{array}{|cccc|}\hline 1 & -1 & 1 & -1 \\ \hline 1 & -1 & 1 & -1 \\ \hline 1 & -1 & 1 & -1 \\ \hline 1 & -1 & 1 & -1 \\ \hline \end{array}$$

It can be verified that this choice of the X-matrix provides equality in the above wrt information on β. As a matter of fact, the column vector of order 16×1 consisting of the covariate values is seen to be orthogonal to each of the $4+4+4+4 = 16$ vectors corresponding to 4 block effects parameters, 4 treatment effects parameters, 4 LN-Effects parameters and 4 RN-Effects parameters. It would be an interesting exercise to figure out how many such X-matrices can be made available which are (i) orthogonal to those listed in the above and (ii) themselves mutually orthogonal. Here are two others *i.e.*, $X_{(2)}$ and $X_{(3)}$ displayed in Table 3.

We now refer to Das *et al.* (2015) Monograph on 'Optimal Covariate Designs'. Specifically, subsection 3.2 lists 9 matrices, denoted as $W^{(1)}, W^{(2)}, \ldots, W^{(9)}$, in the context of an

Table 3: Covariate matrices for *RBD* with *v*=*b*=4 in Table 1

$$X_{(2)} = \begin{array}{|cccc|}\hline 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ \hline\end{array} \qquad X_{(3)} = \begin{array}{|cccc|}\hline 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ \hline\end{array}$$

RBD with parameters $b = v = 4$. It turns out that all these 9 matrices serve our purpose in the present context. As a matter of fact, we have already listed 3 of them [W^1, W^2, W^3] in the above - suitably rewritten to fit in our framework as $X_{(1)}$, $X_{(2)}$, $X_{(3)}$ in Table 2 and Table 3. The rest are shown in the Appendix - A.

Remark 1: It must be noted that not all block design structures are amenable to this kind of allocation of covariate values with desirable orthogonality properties. Take, for example, the following RBD in Table 4 with associated LN- and RN-Effects shown along the margins. We may try to convert W into X-matrix, hoping that it would serve the purpose ! We show it below in Table 5.

Table 4: *RBD* with *b*=*v*=4: Second Choice

LN 4	1	2	3	4	RN 1
LN 3	1	2	4	3	RN 1
LN 4	2	1	3	4	RN 2
LN 3	2	1	4	3	RN 2

Table 5: Non-conformative Covariate Matrix for *RBD* with *b*=*v*=4: Second Choice

1	−1	1	−1
−1	1	1	−1
−1	1	1	−1
1	−1	1	−1

It turns out that (i) block total of x-values is zero for each block;
(ii) treatment total of x-values is zero for each treatment.
However, orthogonality fails wrt LN- and RN-Effects. The message is clear. We have to study the structure of allocation of the treatments in the RBD and proceed accordingly. For the RBD in Table 4, we are able to establish that there are at the most 4 X-matrices - satisfying the desirable properties. We provide a proof of this statement as also display all the available X-matrices in the Appendix - B.

Remark 2: Every layout of an $RBD(b = v = 4)$ is special and has to be dealt with due

attention to its structure. Here we have one more in the 'affirmative' sense displayed in Table 6. It has at least one underlying X-matrix and we display one X-matrix in Table 7.

Table 6: Covariate Matrix for *RBD* with b=v=4: Third Choice

LN 4	1	2	3	4	RN 1
LN 3	2	1	4	3	RN 2
LN 2	3	4	1	2	RN 3
LN 1	4	3	2	1	RN 4

Table 7: *RBD* with b=v=4: Third Choice

-1	1	-1	1
-1	1	-1	1
-1	1	-1	1
-1	1	-1	1

Remark 3: It is tempting to conjecture that for any given layout of an RBD, there is at least one X-matrix available satisfying all the properties stipulated.

3. BIBD with b=v=7, r=k=4, λ=2

We borrow necessary results from Das *et al.* (2015), Chapter 4, Subsection 4.2. We take up the $SBIBD(7, 7, 4, 4, 2)$ and display the incidence matrix in a slightly modified form below in Table 8. We also show the LN- and RN-Effects in the same table, assuming a circular model. We now display the X-matrix of $((+1, -1))$'s in Table 9.

Table 8: *SBIBD* (7, 7, 4, 4, 2)

LN 7	1	4	6	7	RN 1
LN 7	1	2	5	7	RN 1
LN 2	1	6	3	2	RN 1
LN 7	2	3	4	7	RN 2
LN 3	1	5	4	3	RN 1
LN 6	2	4	5	6	RN 2
LN 7	3	5	6	7	RN 3

It is readily verified that this X-matrix is one desired solution to provide most efficient estimation of the β-coefficient even in the presence of LN- and RN-effects of the treatments. As a ready reckoner, we display below in Table 10, the LN-and RN-effects of the treatments, assuming a circular model. Note that the positions of the treatments within the blocks are important for assessing the properties of the X-matrix. It would be interesting to investigate if there are other such X-matrices and mutually orthogonal to the one just found.

Table 9: X-matrix for *SBIBD* (7,7,4,4,2)

1	0	0	−1	0	1	−1
−1	1	0	0	−1	0	1
1	−1	1	0	0	−1	0
0	1	−1	1	0	0	−1
−1	0	1	−1	1	0	0
0	−1	0	1	−1	1	0
0	0	−1	0	1	−1	1

Table 10: LN- and RN-Effects under a circular model

LN Tr 1	Coeff	RN Tr 1	Coeff.
4	−1	7	−1
2	1	7	1
6	−1	2	−1
5	1	3	1
*	*	*	*
LN Tr 2	Coeff	RN Tr 2	Coeff.
5	−1	1	−1
1	1	3	1
3	−1	7	−1
4	1	6	1
*	*	*	*
LN Tr 3	Coeff	RN Tr 3	Coeff.
2	−1	6	−1
4	1	2	1
1	−1	4	−1
5	1	7	1
*	*	*	*

LN Tr 4	Coeff	RN Tr 4	Coeff.
6	1	1	1
7	−1	3	−1
3	1	5	1
5	−1	2	−1
*	*	*	*
LN Tr 5	Coeff	RN Tr 5	Coeff.
7	1	2	1
4	−1	1	−1
6	1	4	1
6	−1	3	−1
*	*	*	*
LN Tr 6	Coeff	RN Tr 6	Coeff.
7	−1	4	−1
3	1	3	1
2	−1	5	−1
7	1	5	1
*	*	*	*
LN Tr 7	Coeff	RN Tr 7	Coeff.
6	1	6	1
5	−1	5	−1
2	1	4	1
3	−1	6	−1
*	*	*	*

4. Latin Square Design of Order 4

So far we have developed study of RBDs and BIBDs with covariates and in the presence of neighbor- effects. Now we focus on an LSD of order 4. We refer to Das *et al.* (2015), pages 155 - 159. In Example 8.2.3 (page 155), an LSD of order 4 has been laid out. We reproduce it here in Table 11 along with all the four-sided neighbor-effects : Left-sided Neighbor Effects (LN), Right-sided Neighbor Effects (RN), Top-sided Neighbor Effects (TN) and Down-sided Neighbor Effects (DN). We assume a circular model - covering all sides.

Since in an LSD of order 4, there are six (6) orthogonal linear error functions (*i.e.*, 6 error df), in the Example 8.2.3, six (6) orthogonal X-matrices have been shown. Vide the

Table 11: *LSD* of order 4 with 4-sided NEs

TN	4	3	2	1	Effects
LN-Effects					RN-Effects
4	1	2	3	4	1
3	2	1	4	3	2
2	3	4	1	2	3
1	4	3	2	1	4
DN	1	2	3	4	Effects

bottom part of the matrix shown in the expression for E^{LSD}. These represent optimal choices of six orthogonal covariate matrices for estimation of the same number of beta-coefficients. This, however, holds without the presence of any sort of neighbor effects. While we introduce the N-Effects on all sides (*i.e.*, in all directions), it follows that only four (4) of them are valid X-matrices. These are the 2nd, 4th, 5th and 6th X-matrices in the bottom part of the table for E^{LSD}. These are reproduced below for the sake of completeness in Table 12. Moreover, as in the case of the RBD in Table 4, we prove that for the LSD under consideration, there exist only 4 distinct and mutually orthogonal X-matrices as are found out and displayed in Table 12. This is taken up in Appendix - C.

Table 12: Optimal X-matrices

X(1)	1	−1	−1	1;	−1	1	1	−1;	1	−1	−1	1;	−1	1	1	−1
X(2)	1	−1	1	−1;	−1	1	−1	1;	−1	1	−1	1;	1	−1	1	−1
X(3)	1	−1	−1	1;	1	−1	−1	1;	−1	1	1	−1;	−1	1	1	−1
X(4)	1	1	−1	−1;	−1	−1	1	1;	1	1	−1	−1;	−1	−1	1	1

Remark 4: We must note that the choice of the specific form of the LSD is very crucial for existence of such X-matrices. For example, if we adopt the LSD shown in Table 13 [reproduced as L_2 on Page 29 of Das *et al.* (2015)], then we can find one X-matrix comfortably and it is shown in Table 14. However, our attempt to find one more did not succeed.

Remark 5: Even though we are discussing about LSDs of order 4, very general treatments of row-column designs are available in the literature. Vide, for example, Shah and Sinha (1996). The reader might like to study such general patterns in the light of Neighbor- Effects and covariates.

Table 13: *LSD* of order 4 from Das *et al.* (2015) Page 29 L_2

1	2	3	4
3	4	1	2
4	3	2	1
2	1	4	3

Table 14: X-matrix for *LSD* in Table 13

1	−1	−1	1
1	−1	−1	1
−1	1	1	−1
−1	1	1	−1

5. Conclusion

In this paper we have examined the existence of 'optimal covariates designs' in the presence of neighbor-effects. The designs considered are (i) $RBD(b = v = 4)$, (ii) $BIBD(b = v = 7, r = k = 4, \lambda = 2)$ and (iii) $LSD\ of\ Order\ 4$. The model adopted is linear in the general mean, block - effects / row-column effects, treatment effects and circularly located neighbor- efffects. The presence of covariates makes the analysis complicated unless their effects are optimally and orthogonally estimated. This study shows that at times we are in a position to achieve this by suitably allocating the covariates values in the experimental units. Even though the experimental set-ups are simple, the results are non-trivial and worth noting.

Acknowledgement

The first author thanks her Mentor Professor KK Singh Meitei for providing all facilities towards successfully pursuing her research in the broad area of DoE and also for arranging visits of Professor Sinha to the Manipur University for collaborative research. The authors are thankful to an anonymous referee for careful reading of the manuscript and useful suggestions.

The first author also acknowledges financial support from **DST Women Scientist Scheme-A, Project Sanction order No. SR/WOS-A/PM-98/2017(G)**.

References

Das, P., Dutta, G., Mandal, N. K. and Sinha, Bikas K. (2015). *Optimal Covariate Designs.* Springer-Verlag Text Book Series.

Jaggi, S., Pateria, D. K., Varghese, C., Varghese, E. and Bhowmik A. (2018). A note on circular neighbor balanced designs. *Communication in Statistics- Simulation and Computation*, **47(10)**, 2896-2905.

Kunert, J. (1984). Optimality of balanced uniform repeated measurements designs. *Annals of Statistics*, **12(3)**, 1006 - 1017.

Sapam, Sobita, Mandal, Nripes K. and Sinha, Bikas K. (2019a). Latin Square designs with neighbour effects. *The Journal of Indian Society of Agricultural Statistics*, **73(2)**, 91-98.

Sapam, Sobita, Mandal, Nripes K. and Sinha, Bikas K. (2019b). Latin Square designs with neighbor effects - part II, *Communications in Statistics - Theory and Methods.* Published online 2019. https://doi.org/10.1080/03610926.2019.1702694.

Shah, K. R. and Sinha, Bikas K. (1996). Row-column designs. In *Design and Analysis of Experiments*. 903-937, *Handbook of Statistics*, **Vol. 13**., North-Holland, Amsterdam.

Varghese, E., Jaggi, S. and Varghese, C. (2014). Neighbor- Balanced Row Column Designs. *Communications in Statistics - Theory and Methods*, **43(6)**, 1261-1276.

APPENDIX

A.I : Choice of Six Additional and Mutually Orthogonal Optimal Covariate Matrices for the RBD(b=v=4) in Table 1

Table 15: Covariate matrices for *RBD* with v=b=4 in Table 1

$X_{(4)}$=

1	−1	1	−1
−1	1	−1	1
−1	1	−1	1
1	−1	1	−1

$X_{(5)}$=

1	−1	−1	1
1	1	−1	−1
−1	1	1	−1
−1	−1	1	1

$X_{(6)}$=

1	1	−1	−1
−1	1	1	−1
1	1	−1	−1
−1	1	1	−1

$X_{(7)}$=

1	−1	1	−1
1	−1	1	−1
−1	1	−1	1
−1	1	−1	1

$X_{(8)}$=

1	−1	−1	1
−1	−1	1	1
1	−1	−1	1
−1	−1	1	1

$X_{(9)}$=

1	1	−1	−1
−1	1	1	−1
−1	−1	1	1
1	−1	−1	1

A.II : Verification of Orthogonality wrt LN- and RN-effects of each of the treatments

We take up the verification wrt $X_{(1)}$ below in Table 16.

The nature of incidence of the treatments as LN- and RN-Effects is very special. That is clearly visible in Table 16. The conditions relating to orthogonality wrt these N-Effects are the same as orthogonality wrt (direct) treatment effects which is true. Therefore, all the X-matrices satisfy the stipulated conditions of orthogonality.

Table 16: Coefficients of LN- and RN- Effects for *RBD* in Table 1 corresponding to the covariate matrix $X_{(1)}$

Blocks	Tr. 1 as LNE	as LNE coeff	Tr. 1 as RNE	as RNE coeff
1	2	−1	4	−1
2	2	1	4	1
3	2	−1	4	−1
4	2	1	4	1
Total		0		0
Blocks	Tr. 2 as LNE	as LNE coeff	Tr. 2 RNE	as RNE coeff
1	3	1	1	1
2	3	−1	1	−1
3	3	1	1	1
4	3	−1	1	−1
Total		0		0
Blocks	Tr. 3 as LNE	as LNE coeff	Tr. 3 RNE	as RNE coeff
1	4	−1	2	−1
2	4	1	2	1
3	4	−1	2	−1
4	4	1	2	1
Total		0		0
Blocks	Tr. 4 as LNE	as LNE coeff	Tr. 4 RNE	as RNE coeff
1	1	1	3	1
2	1	−1	3	−1
3	1	1	3	1
4	1	−1	3	−1
Total		0		0

B. X- matrices for RBD : Second Choice

We have displayed four mutually orthogonal covariate matrices for the $RBD(b = v = 4)$: *Second Choice* in the Table 17. We now establish that no further X-matrices exist in this context. Let us start with a general form of an X-matrix given in Table 18.

Table 17: Four covariate matrices for *RBD* v=b=4: Second Choice

$X_{(1)}$=

1	−1	1	−1
1	−1	1	−1
1	−1	1	−1
1	−1	1	−1

$X_{(2)}$=

1	−1	1	−1
−1	1	−1	1
−1	1	−1	1
1	−1	1	−1

$X_{(3)}$=

1	−1	−1	1
−1	1	1	−1
−1	1	1	−1
1	−1	−1	1

$X_{(4)}$=

1	1	−1	−1
−1	−1	1	1
−1	−1	1	1
1	1	−1	−1

We realize that there are too many restrictions on the elements of X. It may be noted that WOLG, we may assume $a = 1$. The restrictions are listed below in Table 19. By examining

Table 18: General form of a covariate matrix X for *RBD* (*v*=*b*=4): Second Choice

a	b	c	d
e	f	g	h
i	j	k	l
m	n	o	p

the triplet (b, e, f) and all the $2^3 = 8$ combinations along with $a = 1$, we can argue that the following are the only feasible combinations in this context.
$(a, b, e, f) = (1, 1, -1, -1), X_{(4)} Matrix$;
$(a, b, e, f) = (1, -1, 1, -1), X_{(1)} Matrix$;
$(a, b, e, f) = (1, -1, -1, 1), X_{(2)}$ *and* $X_{(3)}$ *Matrices.*
Hence the stated claim is established.

Table 19: Restrictions on the elements of X

Sl. No.	Restriction
Tr1	a+e+j+n = 0
Tr2	b+f+i+m = 0
Tr3	c+h+k+p = 0
Tr4	d+g+l+o = 0
Bl1	a+b+c+d = 0
Bl2	e+f+g+h = 0
Bl3	i+j+k+l = 0
Bl4	m+n+o+p = 0
LN1	b+f+k+o = 0
LN2	c+g+j+h = 0
LN3	d+e+l+m = 0
LN4	a+h+i+p = 0
RN1	d+h+i+m = 0
RN2	a+e+l+p = 0
RN3	b+g+j+o=0
RN4	c+f+k+n=0

C : Existence of four mutually orthogonal X-matrices for the LSD in Table 11
We refer to Table 11 for the particular LSD of order 4 and also to Table 18 for a general structure of an X-matrix. We now incorporate the conditions for optimality.

(a) Consideration of Treatment Effects :

$$a + f + k + p = 0 \quad (1); \quad b + e + l + o = 0 \quad (2);$$
$$c + h + i + n = 0 \quad (3); \quad d + g + j + m = 0 \quad (4).$$

(b) Consideration of Row Effects :

$$a + b + c + d = 0 \quad (5); \quad e + f + g + h = 0 \quad (6);$$

$$i+j+k+l=0 \quad (7); \quad m+n+o+p=0 \quad (8).$$

(c) Consideration of Column Effects :

$$a+e+i+m=0 \quad (9); \quad b+f+j+n=0 \quad (10);$$

$$c+g+k+o=0 \quad (11); \quad d+h+l+p=0 \quad (12).$$

(d) Consideration of Left-Neighbor Effects :

$$b+g+l+m=0 \quad (13); \quad c+f+i+p=0 \quad (14);$$

$$d+e+j+o=0 \quad (15); \quad a+h+k+n=0 \quad (16).$$

(e) Consideration of Right-Neighbor Effects :

$$b+g+l+m=0 \quad (17); \quad c+f+i+p=0 \quad (18);$$

$$e+j+o+d=0 \quad (19); \quad a+h+k+n=0 \quad (20).$$

(f) Consideration of Top-Neighbor Effects

$$b+g+l+m=0 \quad (21); \quad a+h+k+n=0 \quad (22);$$

$$e+j+o+d=0 \quad (23); \quad c+f+i+p=0 \quad (24).$$

(g) Consideration on Down-Neighbor Effects

$$e+j+o+d=0 \quad (25); \quad c+f+i+p=0 \quad (26);$$

$$b+g+l+m=0 \quad (27); \quad a+h+k+n=0 \quad (28).$$

From the above, we find that the 4 equation sets, *viz.*, those arising out of LN-sum, RN-sum, TN-sum and DN-sum, each of 4 equations, are the same. So, we consider only the 4 equation sets, *viz.*, those arising from Treatment- sum, Row- sum, Column- sum and LN-sum. If there exists a solution of these equations with solution space $[1,-1]$, an X-matrix exists. As in the case of RBD set-up, WOLG, we set $a=1$ and examine all the 8 combinations corresponding to choices of (b,e,f). The results are stated below.

Case 1. $b=e=f=1:$ *no solution*;
Case 2. $b=-1,\ e=f=1:$ *no solution*;
Case 3. $b=f=-1, e=1:$ *one solution viz.*, $X(1)$;
Case 4. $e=-1, b=f=1:$ *no solution*;
Case 5. $f=-1, b=e=1:$ *no solution*;
Case 6. $b=e=-1, f=1:$ $X(3)$ *and* $X(2)$ *are the two solutions*;
Case 7. $e=f=-1, b=1:$ *one solution viz.*, $X(4)$;
Case 8. $b=e=f=-1:$ *no solution.*

Hence the claim is justified.

Statistics and Applications {ISSN 2454-7395(online)}
Volume 19, No. 1 (New Series), pp 41–60

Bulk Behaviour of Skew-Symmetric Patterned Random Matrices

Arup Bose[1] and Soumendu Sundar Mukherjee[2]
[1]*Statistics and Mathematics Unit, Indian Statistical Institute, 203 B. T. Road, Kolkata 700108, India*
[2]*Interdisciplinary Statistical Research Unit, Indian Statistical Institute, 203 B. T. Road, Kolkata 700108, India*

Received: 29 March 2020; Revised: 09 July 2020; Accepted: 10 July 2020

Abstract

Limiting Spectral Distributions (LSDs) of real symmetric patterned matrices have been well-studied. In this article, we consider skew-symmetric/anti-symmetric patterned random matrices and establish the LSDs of several common matrices.

For the skew-symmetric Wigner, skew-symmetric Toeplitz and the skew-symmetric Circulant, the LSDs (on the imaginary axis) are the same as those in the symmetric cases.

However, for the skew-symmetric Hankel and the skew-symmetric Reverse Circulant, we obtain new LSDs. We also show the existence of LSDs for the triangular versions of these matrices.

We then introduce a related modification of the symmetric matrices by changing the sign of the lower triangle part of the matrices. In this case, the modified Wigner, modified Hankel and the modified Reverse Circulant have the same LSDs as their usual symmetric counterparts while new LSDs are obtained for the modified Toeplitz and the modified Symmetric Circulant.

Key words: Patterned random matrices; Limiting spectral distribution; Skew-symmetric Circulant, Hankel, Reverse Circulant, Toeplitz, Wigner matrices; Semi-circular law; Link function; Catalan word; Symmetric word.

AMS Subject Classifications: 60B20, 60B10

I knew Professor Aloke Dey for more than two decades. Aloke-da was a caring elder brother as well as a close family friend. I took his advice on numerous occasions and he was a pillar of support. A quintessential gentleman, he always maintained the most cordial relation with everyone, irrespective of which side of an issue the person was. His demise has left a large void and I miss him sorely. Fond memories survive.

Arup Bose

Corresponding Author: Arup Bose
Email: bosearu@gmail.com

1. Introduction

Suppose A_n is an $n \times n$ matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$. The empirical spectral measure μ_n of A_n is the random measure

$$\mu_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i}, \tag{1}$$

where δ_x is the Dirac delta measure at x. The corresponding random probability distribution (on $\mathbb{R}$ or $\mathbb{R}^2$, depending on whether the eigenvalues are real or complex) is known as the *Empirical Spectral Distribution* (ESD) and is denoted by F^{A_n}.

The sequence $\{F^{A_n}\}$ is said to converge (weakly) almost surely to a non-random distribution function F if, outside a null set, as $n \to \infty$, $F^{A_n}(\cdot) \to F(\cdot)$ at all continuity points of F. F is known as the *Limiting Spectral Distribution* (LSD).

There has been a lot of recent work on obtaining the LSDs of large dimensional patterned random matrices. These matrices may be defined as follows (Bose and Sen (2008)). Let $(a_i)_{i\geq 1}$ be a sequence of random variables, called an *input sequence*. Let $\mathbb{Z}$ be the set of all integers and $\mathbb{Z}_+$ be the set of all positive integers. Let

$$L_n : \{1, 2, \ldots n\}^2 \to \mathbb{Z} \text{ (or } \mathbb{Z}^2), \ n \geq 1, \tag{2}$$

be a sequence of functions. We shall write $L_n = L$ and call it the *link* function. By a slight abuse of notation, we shall write $\mathbb{Z}_+^2$ as the common domain of $\{L_n\}_{n\geq 1}$. Matrices of the form

$$A_n = n^{-1/2}((a_{L(i,j)}))_{1\leq i,j\leq n} \tag{3}$$

are called *patterned matrices*. If $L(i, j) = L(j, i)$ for all i, j, then the matrix is symmetric. We shall denote the LSD of $\{n^{-1/2}A_n\}$, if it exists, by $\mathcal{L}_A$.

The real symmetric patterned matrices that have received particular attention in the literature are the Wigner, Toeplitz, Hankel, Reverse Circulant and the Symmetric Circulant matrices. Their link functions are given in Table 1.

Table 1: Some common symmetric patterned matrices and their link functions.

Matrix	Notation	Link function
Wigner	W_n	$L_W(i,j) = (\min\{i,j\}, \max\{i,j\})$
Toeplitz	T_n	$L_T(i,j) = \lvert i-j \rvert$
Hankel	H_n	$L_H(i,j) = i+j$
Symmetric Circulant	SC_n	$L_{SC}(i,j) = \frac{n}{2} - \lvert \frac{n}{2} - \lvert i-j \rvert \rvert$
Reverse Circulant	RC_n	$L_{RC}(i,j) = (i+j) (\text{mod } n)$

While the LSDs of the Wigner, Reverse Circulant and the Symmetric Circulant are known explicitly, very little is known about the LSDs of the Hankel and the Toeplitz (see,

Table 2: Skew-symmetric patterned matrices and their LSDs.

Matrix	Notation (M)	LSD of iM
Skew-symmetric Wigner	$\widetilde{W}_n$	Same as W_n
Skew-symmetric Toeplitz	$\widetilde{T}_n$	Same as T_n
Skew-symmetric Hankel	$\widetilde{H}_n$	New LSD
Skew-symmetric Circulant	$\widetilde{SC}_n$	Same as SC_n
Skew-symmetric Reverse Circulant	$\widetilde{RC}_n$	New LSD

e.g., Bose (2018)). Existence of LSD is also known for the upper triangular versions of these matrices, though the nature of these limits is not known.

In this article, we study the existence of the LSDs of skew-symmetric/anti-symmetric patterned matrices. Recall that a matrix S is called skew-symmetric if $S = -S^\top$. In the Physics literature, the term "anti-symmetric" is more common. Technically, if S is a skew-symmetric matrix, then iS is called an anti-symmetric matrix, where i is the imaginary unit. Note that iS is Hermitian. Anti-symmetric Gaussian matrices appeared in the classic work of Mehta (2004) who, among other things, gave an expression for the joint distribution of their eigenvalues. Singular values of skew-symmetric Gaussian Wigner matrices are useful in Statistics too, *e.g.*, in the paired comparisons model (see Kuriki (1993, 2010)). Recently, Dumitriu and Forrester (2010) obtained tridiagonal realizations of anti-symmetric Gaussian β-ensembles.

We first establish the existence of the LSDs of several real skew-symmetric patterned random matrices and identify the limits in some cases. For the skew-symmetric Wigner, skew-symmetric Toeplitz and the skew-symmetric Circulant, the LSDs (on the imaginary axis) are the same as those in the symmetric cases. However, for the skew-symmetric Hankel and the skew-symmetric Reverse Circulant, we obtain new LSDs (see Figure 1). See Table 2 for a summary. We also show the existence of the LSDs for the triangular versions of these matrices that were introduced in Basu *et al.* (2012). While the LSDs are known for the Hermitian versions of some of these matrices, we show that the limits for the skew-symmetric versions may be derived from the proofs for symmetric matrices using simple arguments.

We also introduce a related modification of the symmetric matrices by changing the sign of the lower triangle part below the main anti-diagonal. In this case, the modified Wigner, the modified Hankel and the modified Reverse Circulant have the same LSDs as their symmetric counterparts whereas new LSDs are obtained for the modified Toeplitz and the modified Symmetric Circulant (see Figure 2). See Table 3 for a summary.

2. Preliminaries

We shall use the method of moments to establish the existence of LSDs. For any matrix A, let $\beta_h(A)$ denote the h-th moment of the ESD of A. We quote the following lemma from Bose (2018) which is easy to prove.

Table 3: Modified patterned matrices and their LSDs.

Matrix	Notation	LSD
Modified Wigner	$\widehat{W}_n$	Same as W_n
Modified Toeplitz	$\widehat{T}_n$	New LSD
Modified Hankel	$\widehat{H}_n$	Same as H_n
Modified Symmetric Circulant	$\widehat{SC}_n$	New LSD
Modified Reverse Circulant	$\widehat{RC}_n$	Same as RC_n

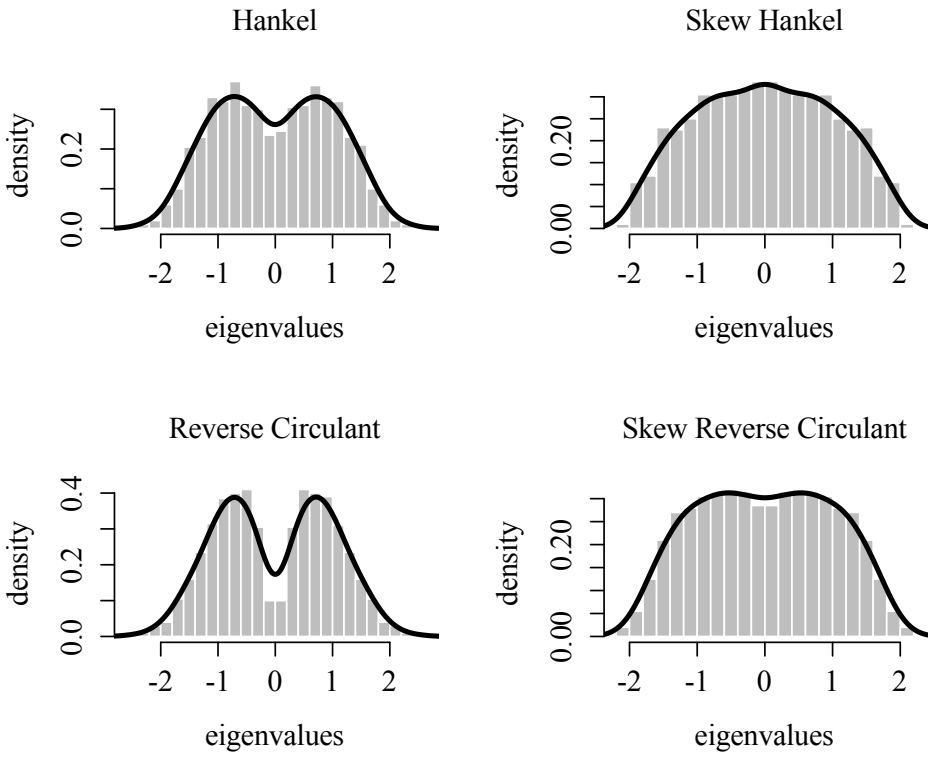


Figure 1: Histograms and kernel density estimates of the spectra of $n^{-1/2}H_n$, $n^{-1/2}i\widetilde{H}_n$, $n^{-1/2}RC_n$ and $n^{-1/2}i\widetilde{RC}_n$ with $n = 1000$ and $\mathcal{N}(0,1)$ entries.

Lemma 1: Let $\{A_n\}$ be a sequence of random matrices with all real eigenvalues. Suppose there exists a sequence $\{\beta_h\}$ such that

(i) for every $h \geq 1$, $\mathbb{E}(\beta_h(A_n)) \to \beta_h$,

(ii) $\sum_{n=1}^{\infty} \mathbb{E}[\beta_h(A_n) - \mathbb{E}(\beta_h(A_n))]^4 < \infty$ for every $h \geq 1$ and

(iii) the sequence $\{\beta_h\}$ satisfies Carleman's condition, $\sum \beta_{2h}^{-1/2h} = \infty$.

Then the LSD of F^{A_n} exists and equals F with moments $\{\beta_h\}$.

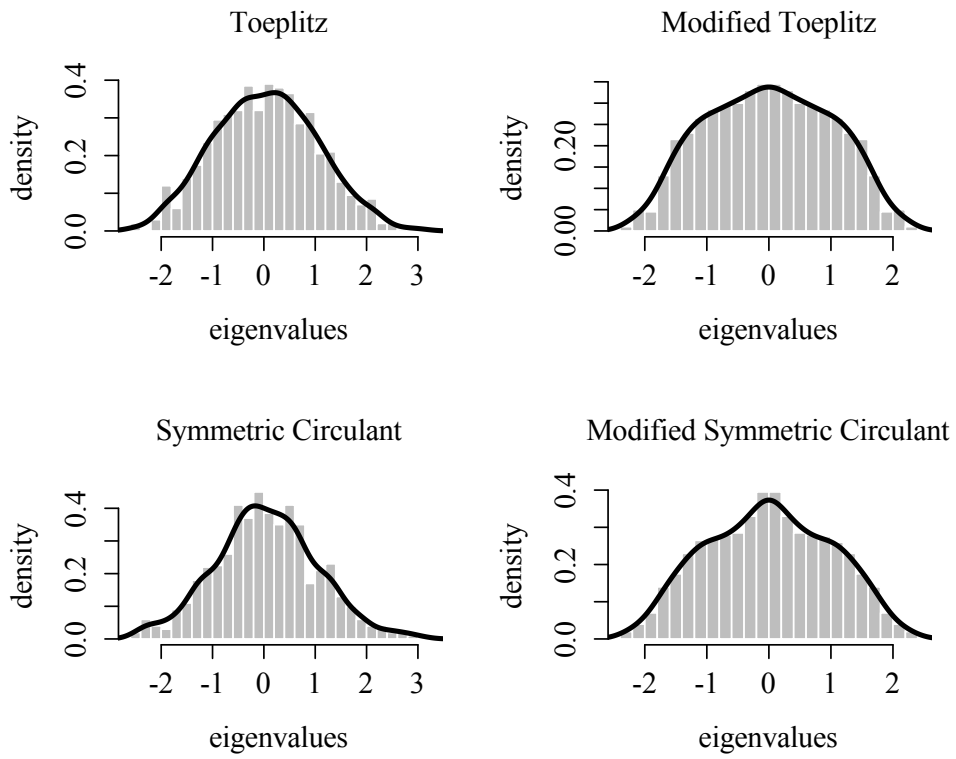


Figure 2: Histograms and kernel density estimates of the spectra of $n^{-1/2}T_n$, $n^{-1/2}\widehat{T}_n$, $n^{-1/2}SC_n$ and $n^{-1/2}\widehat{SC}_n$ with $n = 1000$ and $\mathcal{N}(0,1)$ entries.

To prove the existence of any LSD, we shall make use of the general notation and theory developed in Bose and Sen (2008) for patterned matrices (see also Bose (2018)). First observe that all the link functions in Table 1 satisfy the so called Property B: the total number of times any particular variable appears in any row is uniformly bounded. Moreover, the product of the total number of different variables in the matrix and the maximum number of times any variable appears in the matrix is $O(n^2)$. These two facts imply that the general theory applies to the link functions in Table 1.

We shall consider the following sets of assumptions on the input sequence.

(A1). $(a_i)_{i\geq 1}$ are independent and uniformly bounded with mean 0, and variance 1.

(A2). $(a_i)_{i\geq 1}$ are i.i.d. with mean 0 and variance 1.

(A3). $(a_i)_{i\geq 1}$ are independent with mean 0, variance 1, and uniformly bounded moments of all orders.

Note that Assumption (A1) implies Assumption (A3). Traditionally, LSD results are stated under Assumption (A1) while Assumption (A3) is appropriate for studying the joint convergence of more than one sequence of matrices. It turns out that, for the matrices under our consideration, if LSDs exist under Assumption (A1), then the same LSDs continue to

hold under Assumptions (A2) or (A3). Thus in our proofs, without loss of any generality, Assumption (A1) is assumed to hold. Below we give a brief outline of the reasoning. The reader may consult Bose (2018) for detailed justifications in the similar context of symmetric patterned matrices.

(i) When the entries satisfy Assumption (A1), the main idea is to show that the expected moments of the ESD of A_n converge and these limit moments determine a unique distribution. Moreover, these limit moments depend only on the pattern and not on the specific distribution of the entries. We thus call this limit *universal*.

(ii) If the entries of the matrix under consideration satisfy Assumption (A2), then one considers the same matrix but where the entries are truncated suitably and standardized to have mean 0 and variance 1. This matrix satisfies Assumption (A1) and hence has the same (universal) limit. Then one shows that the original matrix and the modified matrix are close in a suitable metric as $n \to \infty$. This leads us to conclude that the same universal limit persists under Assumption (A2).

(iii) Finally, suppose that the entries satisfy Assumption (A3). Then we compute the moments of the ESD again. Using the "uniformly bounded moments" assumption and Property B of the link function, it can be shown that the third or higher order moments of the variables do not influence the LSD (somewhat like the central limit theorem, for example), and we have the same limit as obtained under Assumption (A1).

The *Moment-Trace Formula* plays a key role in this approach. A function

$$\pi : \{0, 1, \cdots, h\} \to \{1, 2, \cdots, n\}$$

with $\pi(0) = \pi(h)$ is called a *circuit* of length h. The dependence of a circuit on h and n is suppressed. Then, for any $n \times n$ square matrix $A = ((a_{L(i,j)}))$, we have

$$\beta_h(A) = \frac{1}{n} \operatorname{tr}(A^h) = \frac{1}{n} \sum_{\pi \text{ circuit of length } h} a_\pi,$$

where

$$a_\pi := a_{L(\pi(0),\pi(1))} a_{L(\pi(1),\pi(2))} \cdots a_{L(\pi(h-1),\pi(h))}.$$

If $L(\pi(i-1), \pi(i)) = L(\pi(j-1), \pi(j))$, with $i < j$, we shall use the notation (i, j) to denote such a match of the L-values. From the general theory, it follows that circuits where there are only pair-matches are relevant when computing limits of moments.

Two circuits π_1 and π_2 are equivalent if and only if their L-values respectively match at the same locations, *i.e.* if, for all i, j,

$$L(\pi_1(i-1), \pi_1(i)) = L(\pi_1(j-1), \pi_1(j)) \Leftrightarrow L(\pi_2(i-1), \pi_2(i)) = L(\pi_2(j-1), \pi_2(j)).$$

Any equivalence class can be indexed by a partition of $\{1, 2, \cdots, h\}$. We label these partitions by *words* w of length h of letters where the first occurrence of each letter is in alphabetical order. For example, if $h = 4$, then the partition $\{\{1, 3\}, \{2, 4\}\}$ is represented

by the word *abab*. This identifies all circuits π for which $L(\pi(0),\pi(1)) = L(\pi(2),\pi(3))$ and $L(\pi(1),\pi(2)) = L(\pi(3),\pi(1))$. Let $w[i]$ denote the i-th entry of w. The equivalence class corresponding to w is

$$\Pi(w) := \{\pi \mid w[i] = w[j] \Leftrightarrow L(\pi(i-1),\pi(i)) = L(\pi(j-1),\pi(j))\}.$$

By varying w, we obtain all the equivalence classes. It is important to note that, for any fixed h, even as $n \to \infty$, the number of words (equivalence classes) remains finite but the number of circuits in any given $\Pi(w)$ may grow indefinitely. Henceforth we shall denote the set of all words of length h by $\mathcal{A}_h$.

Notions of matches carry over to words. A word is *pair-matched* if every letter appears exactly twice in that word. The set of all pair-matched words of length $2k$ is denoted by $\mathcal{W}_{2k}$. For technical reasons, it is often easier to deal with a class larger than $\Pi(w)$:

$$\Pi^*(w) := \{\pi \mid w[i] = w[j] \Rightarrow L(\pi(i-1),\pi(i)) = L(\pi(j-1),\pi(j))\}.$$

Any i (or $\pi(i)$ by abuse of notation) is a *vertex*. It is *generating* if either $i = 0$ or $w[i]$ is the first occurrence of a letter. Otherwise, it is called non-generating. For example, if $w = abbcab$, then $\pi(0),\pi(1),\pi(2),\pi(4)$ are generating and $\pi(3),\pi(5),\pi(6)$ are non-generating. The set of generating vertices (indices) is denoted by S. By Property B, a circuit is completely determined, up to finitely many choices, by its generating vertices.

From the general theory for symmetric random matrices it follows that the LSD exists if, for each $w \in \mathcal{W}_{2k}$, the following limit exists:

$$p(w) = \lim_n n^{-(k+1)} \#\Pi^*(w).$$

3. A Unified Framework for Real Skew-symmetric Matrices

If A is an $n \times n$ skew-symmetric matrix, then all its eigenvalues $\{\lambda_j\}$ are purely imaginary (and has one zero eigenvalue when n is odd), and every eigenvalue occurs in conjugate pairs. As discussed in the introduction, the Hermitian matrix iA will then have real spectrum. Consider the ESD of iA on $\mathbb{R}$:

$$F^{iA}(x) = \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}_{\{i\lambda_j \le x\}}.$$

Note that F^{iA} is a symmetric (about zero) distribution. Therefore, in order to apply the moment method, it suffices to deal with only the even moments. Note that

$$\begin{aligned}\beta_{2k}(iA) &= \int x^{2k}\, dF^{iA}(x) \\ &= \frac{1}{n}\sum_{j=1}^{n}(i\lambda_j)^{2k} = (-1)^k \frac{1}{n}\sum_{j=1}^{n}\lambda_j^{2k} = (-1)^k \frac{1}{n}\operatorname{tr}(A^{2k}).\end{aligned}$$

Let $\{A_n\}$ be a sequence of $n \times n$ patterned random matrices with the symmetric link function L. Let

$$s_{ij} = (1-\delta_{ij})(-1)^{\mathbf{1}_{\{i>j\}}},$$

where δ_{ij} is the Kronecker-delta. Let $S_n = ((s_{ij}))$ be the $n \times n$ matrix

$$S_n = \begin{pmatrix} 0 & 1 & \dots & 1 \\ -1 & 0 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \dots & 0 \end{pmatrix}_{n\times n}.$$

Then we can construct $\widetilde{A}_n$, the skew-symmetric version of A_n by

$$\widetilde{A}_n = S_n \odot A_n,$$

where $\odot$ denotes the Schur-Hadamard/entrywise product.

We shall assume without loss of generality that (A1) holds. The moment-trace formula for $i\widetilde{A}_n$ may be written as

$$\beta_{2k}(n^{-1/2} i\widetilde{A}_n) = (-1)^k \frac{1}{n^{1+k}} \sum_{\pi \text{ circuit of length } 2k} s_\pi a_\pi.$$

Therefore

$$\mathbb{E}\beta_{2k}(n^{-1/2} i\widetilde{A}_n) = (-1)^k \frac{1}{n^{1+k}} \sum_{\pi \text{ circuit of length } 2k} s_\pi \mathbb{E} a_\pi.$$

Using the concept of words, we may rewrite the above equality as

$$\mathbb{E}\beta_{2k}(n^{-1/2} i\widetilde{A}_n) = (-1)^k \frac{1}{n^{1+k}} \sum_{w\in\mathcal{A}_{2k}} \sum_{\pi\in\Pi(w)} s_\pi \mathbb{E} a_\pi.$$

Suppose L satisfies Property B. Let $C^L_{h,3+}$ denote the set of L-matched h-circuits on $\{1, \cdots, n\}$ with at least one edge of order ≥ 3. Then Lemma 1(a) of Bose and Sen (2008) says that there is a constant C depending on L and h such that

$$\#C^L_{h,3+} \leq C n^{\lfloor (h+1)/2 \rfloor}.$$

Combining this with the observation that $|s_\pi| \leq 1$ it is easy to see that

$$\lim_n \frac{1}{n^{1+k}} \sum_{\pi\in C^L_{2k,3+}} s_\pi \mathbb{E} a_\pi = 0.$$

Therefore

$$\lim_n \mathbb{E}\beta_{2k}(n^{-1/2} i\widetilde{A}_n) = (-1)^k \lim_n \frac{1}{n^{1+k}} \sum_{w\in\mathcal{W}_{2k}} \sum_{\pi\in\Pi(w)} s_\pi \mathbb{E} a_\pi.$$

Since, by our assumptions, $\mathbb{E}a_\pi = 1$ for any pair-matched circuit π, the above expression reduces to

$$\lim_n \mathbb{E}\beta_{2k}(n^{-1/2} i\widetilde{A}_n) = (-1)^k \sum_{w\in\mathcal{W}_{2k}} \lim_n \frac{1}{n^{1+k}} \sum_{\pi\in\Pi(w)} s_\pi, \tag{4}$$

provided the limits on the right-hand side exist. In fact, since $\Pi^*(w) \setminus \Pi(w) \subseteq C^L_{2k,3+}$, one has

$$\lim_n \frac{1}{n^{1+k}} \sum_{\pi \in \Pi(w)} s_\pi = \lim_n \frac{1}{n^{1+k}} \sum_{\pi \in \Pi^*(w)} s_\pi,$$

and thus one can write

$$\lim_n \mathbb{E}\beta_{2k}(n^{-1/2} i\widetilde{A}_n) = (-1)^k \sum_{w \in \mathcal{W}_{2k}} \lim_n \frac{1}{n^{1+k}} \sum_{\pi \in \Pi^*(w)} s_\pi, \tag{5}$$

provided the limits exist for each w. If we define

$$p_{\widetilde{A}}(w) := (-1)^k \lim_n \frac{1}{n^{1+k}} \sum_{\pi \in \Pi(w)} s_\pi,$$

then (5) becomes

$$\lim_n \mathbb{E}\beta_{2k}(n^{-1/2} i\widetilde{A}_n) = \sum_{w \in \mathcal{W}_{2k}} p_{\widetilde{A}}(w). \tag{6}$$

In this context, we recall the analogous expression for symmetric matrices A_n from Bose and Sen (2008):

$$\lim_n \mathbb{E}\beta_{2k}(n^{-1/2} A_n) = \sum_{w \in \mathcal{W}_{2k}} p_A(w),$$

where

$$p_A(w) := \lim_n \frac{1}{n^{1+k}} \#\Pi(w) = \lim_n \frac{1}{n^{1+k}} \#\Pi^*(w)$$

is assumed to exist for each $w \in \mathcal{W}_{2k}$.

It is not difficult to show that if the limits exist in (5), then Condition (iii) of Lemma 1 follows (see Theorem 3 of Bose and Sen (2008) for the argument in the symmetric case; in the skew-symmetric case too, one can use their argument verbatim because $|s_\pi| \leq 1$). In fact, the limiting moments are *sub-Gaussian*, *i.e.* the even moments are dominated by the even moments of some Gaussian distribution. The verification of Condition (ii) is also easy since

$$\prod_{j=1}^4 \mathbb{E}(s_{\pi_j} a_{\pi_j} - \mathbb{E}s_{\pi_j} a_{\pi_j}) = s_{\pi_1} s_{\pi_2} s_{\pi_3} s_{\pi_4} \prod_{j=1}^4 \mathbb{E}(a_{\pi_j} - \mathbb{E}a_{\pi_j})$$

and the arguments given in the proof of Lemma 2 of Bose and Sen (2008) apply with minor modifications.

In the next section, we shall consider several skew-symmetric patterned matrices and show that Condition (i) of Lemma 1 holds by arguing that the limits on the right-hand side of (4) holds in each case.

4. Some Specific Matrices

First note that

$$s_\pi = (-1)^{\sum_{j=1}^{2k} \mathbf{1}_{\{\pi(j-1) > \pi(j)\}}} \prod_{j=1}^{2k} (1 - \delta_{\pi(j-1),\pi(j)}).$$

It is convenient to use some graph theoretic terminology to deal with the above expression. Consider the complete directed graph DK_n on $V = \{1, \cdots n\}$. Note that π defines a directed circuit of length $2k$ on this graph. Call the numerical value of each vertex its *level*. Associate with each π a marking-vector $(\epsilon_1, \cdots, \epsilon_{2k})$, where

$$\epsilon_j = (-1)^{\mathbf{1}_{\{\pi(j-1)>\pi(j)\}}}(1 - \delta_{\pi(j-1),\pi(j)}).$$

Note that if a traveler moves along the circuit π, starting from $\pi(0)$, and marks each move $\pi(j-1) \rightsquigarrow \pi(j)$ by ϵ_j, then moving to a higher (respectively lower) level corresponds to a mark of 1 (respectively -1) and remaining at the same level corresponds to marking with 0. Then

$$s_\pi = \prod_{j=1}^{2k} \epsilon_j.$$

Note that a circuit π contains a *loop* if and only if $s_\pi = 0$.

We first tackle the skew-symmetric Wigner matrix $n^{-1/2}\widetilde{W}_n$. To do so recall the concept of *Catalan* words from Bose (2018). A Catalan word of length 2 is just a double letter aa. In general, a Catalan word of length $2k$, $k > 1$, is a word $w \in \mathcal{W}_{2k}$ containing a double letter such that if one deletes the double letter the reduced word becomes a Catalan word of length $2k-2$. For example, $abba$, $aabbcc$, $abccbdda$ are Catalan words whereas $abab$, $abccab$, $abcddcab$ are not. The set of all Catalan words of length $2k$ will be denoted by $\mathcal{C}_{2k}$. It is known that

$$\#\mathcal{C}_{2k} = \frac{1}{k+1}\binom{2k}{k},$$

the ubiquitous Catalan number from Combinatorics. It is known that $\#\mathcal{C}_{2k}$ also equals the $2k$-th moment of the semi-circular law, the LSD of the Wigner matrix.

Theorem 1: If the input sequence satisfies (A1) or (A2) or (A3), then the LSD of $n^{-1/2}i\widetilde{W}_n$ is the semi-circular law.

Proof: It is well known (see, *e.g.*, Bose (2018)) that, for the symmetric Wigner matrix, only Catalan words contribute in the limit. In fact, one has

$$p_W(w) = \lim_n \frac{1}{n^{1+k}}\#\Pi^*(w) = \begin{cases} 0 & \text{if } w \notin \mathcal{C}_{2k}, \\ 1 & \text{if } w \in \mathcal{C}_{2k}. \end{cases}$$

From this and the fact that $|s_\pi| \leq 1$ it follows that

$$|p_{\widetilde{W}}(w)| \begin{cases} = 0 & \text{if } w \in \mathcal{W}_{2k} \setminus \mathcal{C}_{2k}, \\ \leq 1 & \text{if } w \in \mathcal{C}_{2k}. \end{cases}$$

We shall prove that if w is a Catalan word, then $p_{\widetilde{W}}(w)$ exists and equals 1. Then (5) would imply that

$$\lim_n \mathbb{E}\beta_{2k}(n^{-1/2}i\widetilde{W}_n) = \#\mathcal{C}_{2k},$$

establishing the semi-circular limit for the ESD of $\{n^{-1/2}\widetilde{W}_n\}$.

We first observe that if we replace the diagonal entries by 0, then the LSD does not change. It follows from this observation that *circuits with loops do not have any contribution to* $p_{\widetilde{W}}(w)$. It now suffices for our purpose to prove that if $w \in \mathcal{C}_{2k}$ and $\pi \in \Pi^*(w)$, then

$$s_\pi = \begin{cases} (-1)^k & \text{if } \pi \text{ is loopless,} \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

To prove this, suppose that a double letter appears at the i-th and the $(i+1)$-th positions. Consider a loopless $\pi \in \Pi^*(w)$. Since, $w[i] = w[i+1]$, we must have

$$L_W(\pi(i-1), \pi(i)) = L_W(\pi(i), \pi(i+1)).$$

Since π is loopless, it follows that we must have $\pi(i-1) = \pi(i+1) \neq \pi(i)$. There are two possibilities: either $\pi(i-1) < \pi(i)$ or $\pi(i-1) > \pi(i)$. In the first case, $\epsilon_i = 1$ and $\epsilon_{i+1} = -1$, while, in the second case, $\epsilon_i = -1$ and $\epsilon_{i+1} = -1$. In either case, we have

$$\epsilon_i \epsilon_{i+1} = -1.$$

Now delete the double letter and think of π as a circuit of length $2k-2$ by identifying the vertices $(i-1)$ and $(i+1)$ as identical and deleting the vertex i. The resulting word w' is still Catalan and the resulting circuit π' is loopless and lies in $\Pi^*(w')$. Apply the above procedure again. Clearly, we will need k iterations of this procedure to empty the word w and each such iteration contributes one -1, which proves (7) and hence the theorem. □

Remark 1: Basu *et al.* (2012) considered upper/lower triangular versions of the Wigner, W_n^Δ. Its LSD $\mathcal{L}_{W^\Delta}$ is different from the semi-circular law, but its free convolution with itself is the semi-circular law. It follows from the proof of Theorem 1 and their moment calculations that the LSD of $i\widetilde{W}^\Delta$ is again $\mathcal{L}_{W^\Delta}$.

The existence of the LSD of the symmetric Toeplitz matrix T_n was first established by Hammond and Miller (2005) and Bryc *et al.* (2006). The properties of the limit law $\mathcal{L}_T$ are not well understood. We now consider the skew-symmetric Toeplitz $\widetilde{T}_n$.

Theorem 2: If the input sequence satisfies (A1) or (A2) or (A3), then the LSD of $n^{-1/2} i\widetilde{T}_n$ is $\mathcal{L}_T$, the LSD of the symmetric Toeplitz.

Proof: Let $w \in \mathcal{W}_{2k}$ and $s(i) := \pi(i) - \pi(i-1)$. Define

$$\Pi^{**}(w) := \{\pi \mid w[i] = w[j] \Rightarrow s(i) + s(j) = 0\}.$$

Then Bose and Sen (2008) show that

$$p_T(w) = \lim_n \frac{1}{n^{1+k}} \#\Pi^*(w) = \lim_n \frac{1}{n^{1+k}} \#\Pi^{**}(w). \tag{8}$$

As in the Wigner case, circuits with loops do not contribute and to establish our goal it suffices to prove that if $w \in \mathcal{W}_{2k}$ and $\pi \in \Pi^{**}(w)$, then

$$s_\pi = \begin{cases} (-1)^k & \text{if } \pi \text{ is loopless,} \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

The proof of this is much easier than the Wigner case as all the difficulty is relegated to the proof of (8). Consider a loopless circuit $\pi \in \Pi^{**}(w)$. Note that $w[i] = w[j]$ implies that $s(i) + s(j) = 0$ and since π is loopless, we have

$$s(i)s(j) = -s(j)^2 < 0.$$

This immediately implies that

$$\epsilon_i \epsilon_j = (-1)^{\mathbf{1}_{\{s(i)<0\}} + \mathbf{1}_{\{s(j)<0\}}} = -1.$$

Since w is pair-matched, there are exactly k matches from each of which comes one -1. This establishes (9) and completes the proof. □

Remark 2: Basu *et al.* (2012) considered upper/lower triangular versions of the Toeplitz, T_n^{Δ}. They proved the existence of the LSD but it could not be identified. It follows from the proof of Theorem 2 and their moment calculations that the LSD of $i\widetilde{T}^{\Delta}$ is again $\mathcal{L}_{T^{\Delta}}$, exactly paralleling the Wigner case.

The Symmetric Circulant matrix SC_n and the Palindromic Toeplitz matrix PT_n have the standard Gaussian distribution $\mathcal{N}(0,1)$ as their LSD (see Bose (2018)). We now consider the skew-symmetric versions $\widetilde{SC}_n$ and $\widetilde{PT}_n$.

Theorem 3: If the input sequence satisfies (A1) or (A2) or (A3), then the LSDs of $n^{-1/2}i\widetilde{SC}_n$ and $n^{-1/2}i\widetilde{PT}_n$ are the same as the LSDs of their symmetric counterparts, *i.e.* the standard Gaussian distribution.

Proof: We first tackle $\widetilde{SC}_n$. From Bose and Sen (2008), it is known that, for any $w \in \mathcal{W}_{2k}$, if one defines

$$\Pi'(w) := \{\pi \mid w[i] = w[j] \Rightarrow s(i) + s(j) = 0, \pm n\},$$

then one actually has

$$p_{SC}(w) = \lim_n \frac{1}{n^{1+k}} \#\Pi^*(w) = \lim_n \frac{1}{n^{1+k}} \#\Pi'(w) = 1.$$

Once again, circuits with loops have no role to play and to prove the desired result it suffices to prove that if $w \in \mathcal{W}_{2k}$ and $\pi \in \Pi'(w)$, then

$$s_\pi = \begin{cases} (-1)^k & \text{if } \pi \text{ is loopless,} \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

Due to the similarity with the Toeplitz link function, the proof of the above is similar to that in the Toeplitz case. Let π be a loopless circuit from $\Pi'(w)$. Suppose that $w[i] = w[j]$. Then we have $s(i) + s(j) = 0, \pm n$. We treat each of these three cases separately:

1. $s(i) + s(j) = 0$. This is the same as the Toeplitz case and we conclude that $\epsilon_i \epsilon_j = -1$.

2. $s(i) + s(j) = n$. Note that $s(i) = n - s(j)$ and since π is loopless,

$$|s(j)| = |\pi(j) - \pi(j-1)| \leq n - 1.$$

Therefore $s(i) = n - s(j) > 0$. By symmetry, $s(j) > 0$. Therefore, in this case, $\epsilon_i \epsilon_j = 1$.

3. $s(i) + s(j) = -n$. Note that $s(i) = -(n + s(j))$, and therefore $s(i)$, and by symmetry $s(j)$, are both negative ceding $\epsilon_i \epsilon_j = 1$.

Therefore, combining the above cases,

$$s_\pi = (-1)^{k - e_\pi},$$

where e_π is the number of matches (i, j) where $s(i) + s(j) = \pm n$. It suffices to show that e_π is even. But note that

$$\sum_{i=1}^{2k} s(i) = \pi(2k) - \pi(0) = 0,$$

which cannot occur unless e_π is even. This establishes (10) and completes the proof for $\widetilde{SC}_n$.

To prove the same for $\widetilde{PT}_n$ we take the approach of Bose and Sen (2008). We need the following version of the well known interlacing inequality. We omit its proof.

Suppose A is a real skew-symmetric matrix with eigenvalues $i\lambda_j$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Let B be the $(n-1) \times (n-1)$ principal submatrix of A with eigenvalues $i\mu_k$ with $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_{n-1}$. Then one has

$$\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \geq \cdots \geq \mu_{n-1} \geq \lambda_n,$$

i.e. the imaginary parts of the eigenvalues of B are interlaced between the imaginary parts of the eigenvalues of A.

As a consequence

$$||F^A - F^B||_\infty \leq \frac{1}{n}. \tag{11}$$

Now note that the $n \times n$ principal submatrix of $\widetilde{SC}_{n+1}$ is $\widetilde{PT}_n$. Therefore, from (11), we can conclude that $n^{-1/2} i\widetilde{PT}_n$ also has the standard Gaussian law as its LSD. □

Remark 3: Basu *et al.* (2012) considered the upper/lower triangular versions of the symmetric Circulant, SC_n^Δ. They proved the existence of the LSD but it could not be identified. It follows from the proof of Theorem 3 and their moment calculations that the LSD of $i\widetilde{SC}^\Delta$ is again $\mathcal{L}_{SC^\Delta}$.

The skew-symmetric matrices considered so far have the same LSD (on the imaginary axis) as their corresponding symmetric versions. However, simulations suggest that the LSDs of $n^{-1/2} i\widetilde{H}_n$ and $n^{-1/2} i\widetilde{RC}_n$ exist and are different from those of $n^{-1/2} H_n$ and $n^{-1/2} RC_n$ respectively. See Figure 1. We now establish this rigorously.

In this context, *symmetric* words play the key role. A word $w \in \mathcal{W}_{2k}$ is called symmetric if each letter in w occurs once each in an odd and an even position. For example, the word *aabb* is symmetric and the word *abab* is not. We shall denote the set of symmetric words of length $2k$ by $\mathcal{S}_{2k}$. All Catalan words are symmetric. An example of a non-Catalan symmetric word is *abcabc*. It is easy to prove that

$$\#\mathcal{S}_{2k} = k!.$$

Theorem 4: If the input sequence satisfies (A1) or (A2) or (A3), then the LSDs of $n^{-1/2}i\widetilde{H}_n$ and $n^{-1/2}i\widetilde{RC}_n$ exist and are different from the LSDs of $n^{-1/2}H_n$ and $n^{-1/2}RC_n$ respectively.

Proof: We first consider the skew-symmetric Hankel. First suppose $w \in \mathcal{C}_{2k}$. It is known that then $p_H(w) = 1$. By an argument similar to that given in the proof of Theorem 1 one can show that $p_{\widetilde{H}}(w) = 1$.

Now suppose that w is not symmetric. It is known that then $p_H(w) = 0$. Since, $|s_\pi| \leq 1$, it follows that $p_{\widetilde{H}}(w)$ also vanishes.

More generally, for any pair-matched word w, the limit $p_{\widetilde{H}}(w)$ can be shown to exist using the same Riemann approximation technique that is used in the Hankel case (see, for example, Bose and Sen (2008)). We omit the details.

We now show that this LSD is not the same as in the symmetric Hankel case. Since $|s_\pi| \leq 1$, it is clear that the limit is *sub-Hankel*, *i.e.* $\lim_n \beta_{2k}(n^{-1/2}i\widetilde{H}_n) \leq \lim_n \beta_{2k}(n^{-1/2}H_n)$ for all $k \geq 1$. It is thus enough to show that $\lim_n \beta_{2k}(n^{-1/2}i\widetilde{H}_n) < \lim_n \beta_{2k}(n^{-1/2}H_n)$ for some $k \geq 1$. Since Catalan words contribute 1 to both of these and non-symmetric words do not contribute at all, we need to look at non-Catalan symmetric words. The first such word is $w = abcabc$. We shall show that $p_{\widetilde{H}}(abcabc) < \frac{1}{2} = p_H(abcabc)$.

So let us consider the word $w = abcabc$ and its four generating vertices, *viz.*, $\pi(0)$, $\pi(1)$, $\pi(2)$, $\pi(3)$. Writing $\nu_i = \pi(i)/n$ and expressing the $\frac{1}{n^4}\#\Pi^*(w)$ as a Riemann sum, we know from Bose and Sen (2008) that, for the Hankel matrix,

$$p_H(w) = \int_{I^4} \mathbf{1}_{\{0<\nu_0+\nu_1-\nu_3<1,\, 0<\nu_2+\nu_3-\nu_0<1\}} d\nu_3 d\nu_2 d\nu_1 d\nu_0,$$

where I^4 is the unit 4-cube. Let P be the subset of I^4 where the integrand above is positive. For the skew-symmetric case, however, there are many $\pi \in \Pi^*(w)$ such that $s_\pi = -1$, which means that there are lots of cancellations. More formally, for any $\pi \in \Pi^*(w)$, we have

$$\nu_4 = \nu_0 + \nu_1 - \nu_3,$$
$$\nu_5 = \nu_2 + \nu_3 - \nu_0.$$

If we define

$$g(\nu) = s_\pi = (-1)^{\sum_{j=1}^{2k} \mathbf{1}_{\{\nu_{j-1}<\nu_j\}}},$$

then by resorting to the Riemann approximation technique it is easy to see that

$$p_{\widetilde{H}}(w) = (-1)^3 \int_{I^4} g(\nu) \mathbf{1}_{\{0<\nu_0+\nu_1-\nu_3<1,\, 0<\nu_2+\nu_3-\nu_0<1\}} d\nu_3 d\nu_2 d\nu_1 d\nu_0.$$

We shall show that on a subset of P of positive Lebesgue measure, $g(\nu) = 1$. Consider the set $U = P \cap \{(\nu_0, \nu_1, \nu_2, \nu_3) \mid 0 < \nu_0 < \nu_1 < \nu_2 < \nu_3 < 1\} \subseteq I^4$. We claim that on U, one has $g(\nu) = 1$. To see this, note that we automatically have $\nu_j - \nu_{j-1} > 0$ for $j = 1, 2, 3$. Moreover,

$$\begin{aligned} \nu_4 - \nu_3 &= \nu_1 + \nu_0 - 2\nu_3 < 0, \\ \nu_5 - \nu_4 &= (\nu_2 - \nu_1) + 2(\nu_3 - \nu_0) > 0, \\ \nu_6 - \nu_5 &= 2\nu_0 - \nu_2 - \nu_3 < 0. \end{aligned}$$

Therefore, on U, we have $g(\nu) = (-1)^{1+1+1+(-1)+1+(-1)} = 1$. It now suffices to show that

$$\int_U \mathbf{1}_{\{0<\nu_0+\nu_1-\nu_3<1,\, 0<\nu_2+\nu_3-\nu_0<1\}} d\nu_3 d\nu_2 d\nu_1 d\nu_0 > 0.$$

With some easy manipulations with the constraints it is easy to show that

$$\begin{aligned} \int_U \mathbf{1}_{\{0<\nu_0+\nu_1-\nu_3<1,\, 0<\nu_2+\nu_3-\nu_0<1\}} d\nu &\geq \int_{\frac{1}{3}}^{\frac{1}{2}} \int_{\nu_0}^{\frac{1}{2}} \int_{1-\nu_1}^{\frac{1+\nu_0}{2}} \int_{\nu_2}^{1+\nu_0-\nu_2} d\nu_3 d\nu_2 d\nu_1 d\nu_0 \\ &= \frac{19}{62208} > 0. \end{aligned}$$

This completes the proof for the skew-symmetric Hankel.

Now consider the skew-symmetric Reverse Circulant. By following the arguments in the Hankel case, it is easy to see that each word limit exists, thereby proving the existence of the LSD. Moreover, it is known that, for the Reverse Circulant, $p_{RC}(w) = 1$ if w is symmetric and 0 otherwise. In the present case, $p_{\widetilde{RC}}(w) \leq 1$ for all symmetric words and the non-symmetric words continue to contribute zero. It is also easy to show that if $w \in \mathcal{C}_{2k}$, then $p_{\widetilde{RC}}(w) = p_{RC}(w) = 1$. Thus, as before, it remains to seek out a symmetric non-Catalan word w such that $p(w) < 1$. Once again, we may look at $w = abcabc$ and prove this. Due to the similarity with the Hankel case, we skip the details. □

5. A Related Class of Symmetric Matrices

We have seen that skew-symmetry does not change the LSDs of the Wigner, Toeplitz and the Symmetric Circulant, whereas it changes the LSDs of the Hankel and the Reverse Circulant. We now investigate this issue a little more.

Let M_n be the $n \times n$ symmetric matrix whose upper and lower triangle entries are respectively $+1$ and -1, the anti-diagonal consisting of 0's. Then $M_n = ((m_{ij}))$ where

$$m_{ij} = \begin{cases} 1 & \text{if } i + j < n + 1, \\ 0 & \text{if } i + j = n + 1, \\ -1 & \text{if } i + j > n + 1. \end{cases}$$

We show that an LSD exists for the Schur-Hadamard product of M_n with each of the above five matrices. For a patterned matrix A_n, we denote by $\widehat{A}_n$ its modified version $M_n \odot A_n$.

Note that, for the Wigner and the Hankel cases, the Schur-Hadamard product is also of the same type (with a modified input sequence where the signs have changed for some elements of the sequence)–the fact that the anti-diagonal is zero does not affect the LSDs. Hence their LSDs remain unchanged due to the universality of LSDs with respect to the input variables as long as they satisfy Assumptions (A1) or (A2) or (A3). As we shall see, the LSD remains unchanged for the modified Reverse Circulant matrix too.

Note that $n^{-1/2}\widehat{T}_n$ and $n^{-1/2}\widehat{SC}_n$ are not Toeplitz and Symmetric Circulant matrices. We show that LSDs exist for both and are different from $\mathcal{L}_T$ and $\mathcal{N}(0,1)$ respectively. See Figure 2 for simulations.

Similar to the skew-symmetric case, define

$$\epsilon_i = (1 - \mathbf{1}_{\{\pi(i-1)+\pi(i)=n+1\}})(-1)^{\mathbf{1}_{\{\pi(i-1)+\pi(i)>n+1\}}},$$

and

$$m_\pi = \prod_{i=1}^{n} \epsilon_i.$$

Then we have the following analogue of (6):

$$\lim \mathbb{E}\beta_{2k}(n^{-1/2}\widehat{A}_n) = \sum_{w\in\mathcal{W}_{2k}} p_{\widehat{A}}(w), \tag{12}$$

where

$$p_{\widehat{A}}(w) := \lim_n \frac{1}{n^{1+k}} \sum_{\pi\in\Pi(w)} m_\pi = \lim_n \frac{1}{n^{1+k}} \sum_{\pi\in\Pi^*(w)} m_\pi$$

is assumed to exist for each $w \in \mathcal{W}_{2k}$. First we consider the LSD of $n^{-1/2}\widehat{RC}_n$.

Theorem 5: If the input sequence satisfies (A1) or (A2) or (A3), then the LSD of $n^{-1/2}\widehat{RC}_n$ is the same as the LSD of $n^{-1/2}RC_n$, *i.e.* $\mathcal{L}_{RC}$.

Proof: To prove this theorem, note that, by (12), it is enough to prove that $m_\pi = 1$ for each $\pi \in \Pi^*(w)$, where $w \in \mathcal{W}_{2k}$. Define

$$t(i) = \pi(i-1) + \pi(i) \text{ and } u(i) = t(i) - (n+1).$$

Call a circuit π *good* if $m_\pi \neq 0$. It is enough to consider only such circuits.

If $w[i] = w[j]$, then we have

$$t(i) \equiv t(j) \pmod{n},$$

which implies that $u(i) \equiv u(j) \pmod{n}$. Now note that

$$-(n-1) = 2 - (n+1) \leq u(i) \leq n + n - (n+1) = n - 1,$$

and hence

$$|u(i) - u(j)| \leq 2(n-1).$$

So, we must have

$$u(i) - u(j) = 0, \pm n.$$

Observe that

1. If $u(i) - u(j) = 0$, then $\epsilon_i = \epsilon_j$, which yields $\epsilon_i \epsilon_j = 1$.
2. If $u(i) - u(j) = n$, then $u(i) = n + u(j) > 0$, and $u(j) = u(i) - n < 0$, as $|u(l)| \leq n - 1$ for any l. So, in this case, $\epsilon_i \epsilon_j = -1$.
3. If $u(i) - u(j) = -n$, then, again, $\epsilon_i \epsilon_j = -1$ by interchanging the roles of i and j in the previous argument.

As a consequence

$$m_\pi = (-1)^{e_\pi},$$

where e_π is the number of matches (i, j) in π for which $u(i) - u(j) = t(i) - t(j) = \pm n$. Let further e_π^+ be the number of matches (i, j) in π for which $t(i) - t(j) = n$ and $e_\pi^- = e_\pi - e_\pi^+$. First notice that

$$\sum_{i=1}^{2k} t(i) = 2 \sum_{i=1}^{2k} \pi(i).$$

The same sum can be written as

$$\sum_{(i,j)\,\text{match}} (t(i) + t(j)).$$

Notice then that

$$\begin{aligned} \sum_{(i,j)\,\text{match}} (t(i) + t(j)) &= \sum_{(i,j)\,\text{match}} (t(i) - t(j)) + 2 \sum_{(i,j)\,\text{match}} t(j) \\ &= (e_\pi^+ - e_\pi^-)n + 2 \sum_{(i,j)\,\text{match}} t(j) \\ &= ne_\pi - 2ne_\pi^- + 2 \sum_{(i,j)\,\text{match}} t(i). \end{aligned}$$

It follows from the above considerations that ne_π is always even. Now suppose that n is odd. It then follows that e_π is even and therefore $m_\pi = 1$. The case with n even seems to be more complicated. It is not clear why e_π has to be even. We shall use a little trick to bypass the need to pinpoint the parity of e_π in this case. Define, for $w \in \mathcal{W}_{2k}$,

$$\begin{aligned} q_n(w) &:= \frac{1}{n^{1+k}} \sum_{\pi \in \Pi^*(w)} m_\pi, \\ p_n(w) &:= \frac{1}{n^{1+k}} \#\Pi^*(w). \end{aligned}$$

Then it is known from Bose and Sen (2008) that

$$p_n(w) = p_{RC}(w) + o(1),$$

which implies, since $|q_n(w)| \leq |p_n(w)|$, that

$$|q_n(w)| = O(1). \tag{13}$$

We have already proved that (as we have proved that $m_\pi = 1$ for n odd)

$$q_{2n+1}(w) = p_{RC}(w) + o(1). \tag{14}$$

In the following lemma, we shall write $\Pi_n^*(w)$ instead of $\Pi^*(w)$ to explicitly denote the dependence on n.

Lemma 2: We have

$$\#\Pi_{n+1}^*(w) - \#\Pi_n^*(w) = o(n^{1+k}).$$

Proof: We have

$$p_n(w) = \frac{1}{n^{1+k}}\#\Pi_n^*(w) = p(w) + o(1),$$

which can be rewritten as

$$\#\Pi_n^*(w) = p(w)n^{1+k} + o(n^{1+k}).$$

As a consequence

$$\#\Pi_{n+1}^*(w) - \#\Pi_n^*(w) = p(w)((n+1)^{1+k} - n^{1+k}) + o(n^{1+k}),$$

from which the lemma follows since the first term is $O(n^k)$. □ We need another lemma.

Lemma 3: We have

$$q_{n+1}(w) - q_n(w) = o(1).$$

Proof: We have, using the triangle inequality,

$$\begin{aligned}
&|q_{n+1}(w) - q_n(w)| \\
&\quad = \left| \frac{1}{(n+1)^{1+k}} \sum_{\pi \in \Pi_{n+1}^*(w)} m_\pi - \frac{1}{n^{1+k}} \sum_{\pi \in \Pi_n^*(w)} m_\pi \right| \\
&\quad = \left| \frac{1}{(n+1)^{1+k}} \sum_{\pi \in \Pi_n^*(w)} m_\pi + \frac{1}{(n+1)^{1+k}} \sum_{\pi \in \Pi_{n+1}^*(w) \setminus \Pi_n^*(w)} m_\pi - \frac{1}{n^{1+k}} \sum_{\pi \in \Pi_n^*(w)} m_\pi \right| \\
&\quad \le \left| \frac{1}{(n+1)^{1+k}} \sum_{\pi \in \Pi_n^*(w)} m_\pi - \frac{1}{n^{1+k}} \sum_{\pi \in \Pi_n^*(w)} m_\pi \right| + \left| \frac{1}{(n+1)^{1+k}} \sum_{\pi \in \Pi_{n+1}^*(w) \setminus \Pi_n^*(w)} m_\pi \right| \\
&\quad =: \text{(I)} + \text{(II)}.
\end{aligned}$$

Using (13), we get

$$\text{(I)} \le \left| \left(\frac{n}{n+1} \right)^{1+k} - 1 \right| \times |q_n(w)| = o(1) \times O(1) = o(1).$$

On the other hand, by Lemma 2, we have

$$\text{(II)} \le \frac{1}{n^{1+k}} \#(\Pi_{n+1}^*(w) \setminus \Pi_n^*(w)) = o(1).$$

Together, the above two estimates imply the lemma. □ Coming back to the original problem, because of Lemma 3 and (14), we can write

$$\begin{aligned} q_{2n+2}(w) &= q_{2n+1}(w) + o(1) \\ &= p_{RC}(w) + o(1). \end{aligned}$$

This establishes, irrespective of the parity of n, that

$$q_n(w) = p_{RC}(w) + o(1),$$

which completes the proof of the theorem. □ Finally, we give the result on the LSDs of $n^{-1/2}\widehat{T}_n$ and $n^{-1/2}\widehat{SC}_n$.

Theorem 6: If the input sequence satisfies (A1) or (A2) or (A3), then the LSDs of $n^{-1/2}\widehat{T}_n$ and $n^{-1/2}\widehat{SC}_n$ exist and are different from the LSDs of $n^{-1/2}T_n$ and $n^{-1/2}SC_n$ respectively.

Proof: We shall outline the proof only for $n^{-1/2}\widehat{T}_n$. The proof for $n^{-1/2}\widehat{SC}_n$ is similar and is omitted.

Once again, the existence of the LSD, say $\mathcal{L}_{\widehat{T}}$, may be proven using the Riemann approximation technique. We show that $\mathcal{L}_{\widehat{T}}$ does not equal $\mathcal{L}_T$. As in the proof of Theorem 1 we can show that, for each Catalan word w, $p_{\widehat{T}}(w) = 1 = p_T(w)$. Thus we need to look at a non-Catalan pair-matched word. The first such word is $w = abab$. We shall show that $p_{\widehat{T}}(abab) \neq p_T(abab) = 2/3$, which would conclude proof. Using the Riemann approximation argument, it is easy to show that

$$p_{\widehat{T}}(w) = \int_{I^3} (-1)^{\sum_{i=1}^{4} \mathbf{1}_{\{\nu_i + \nu_{i-1} > 1\}}} \mathbf{1}_{\{0 \leq \nu_0 - \nu_1 + \nu_2 \leq 1\}} d\nu_2 d\nu_1 d\nu_0,$$

where $\nu_3 = \nu_0 - \nu_1 + \nu_2$ and $\nu_4 = \nu_0$. Now, similar to the skew-symmetric Hankel case, one can show that on a subset of positive Lebesgue measure the integrand above is negative. In fact, a calculation in `Mathematica` reveals that $p_{\widehat{T}}(abab) = 2/9$. This proves the theorem completely. □

Acknowledgments

We thank the anonymous referee for a careful reading of the manuscript. Her/his suggestions have led to an improved clarity in our presentation. Most of the work in this paper appeared in Chapter 2 of Mukherjee (2014), the Masters dissertation of SSM under the guidance of AB. Research of AB is supported by a J. C. Bose National Fellowship, Department of Science and Technology, Government of India. SSM is supported by an INSPIRE Faculty Fellowship, Department of Science and Technology, Government of India.

References

Basu, R., Bose, A., Ganguly, S. and Hazra, R. S. (2012). Spectral properties of random triangular matrices. *Random Matrices. Theory and Applications*, **1(3)**,1250003.

Bose, A. (2018). *Patterned Random Matrices*. CRC Press.

Bose, A. and Sen, A. (2008). Another look at the moment method for large dimensional random matrices. *Electronic Journal of Probability*, **13(21)**, 588–628.

Bryc, W., Dembo, A. and Jiang, T. (2006). Spectral measure of large random Hankel, Markov and Toeplitz matrices. *The Annals of Probability*, **34(1)**, 1–38.

Dumitriu, I. and Forrester, P. J. (2010). Tridiagonal realization of the antisymmetric Gaussian β-ensemble. *Journal of Mathematical Physics*, **51(9)**, 093302.

Hammond, C. and Miller, S. J. (2005). Distribution of eigenvalues for the ensemble of real symmetric Toeplitz matrices. *Journal of Theoretical Probability*, **18(3)**, 537–566.

Kuriki, S. (1993). Orthogonally invariant estimation of the skew-symmetric normal mean matrix. *Annals of the Institute of Statistical Mathematics*, **45(4)**, 731–739.

Kuriki, S. (2010). Distributions of the largest singular values of skew-symmetric random matrices and their applications to paired comparisons. *Communications in Statistics—Theory and Methods*, **39(8-9)**, 1522–1535.

Mehta, M. L. (2004). *Random Matrices*. Academic Press.

Mukherjee, S. S. (2014). *Limiting Spectra of Random Matrices*. Masters dissertation. Indian Statistical Institute, Kolkata.

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 61-65

R-optimal Designs for Linear Haar-Wavelet Regression Models

Kashinath Chatterjee
Department of Statistics, Visva Bharati University, Santiniketan, West Bengal, India

Received: 18 July 2020; Revised: 14 August 2020; Accepted: 19 August 2020

Abstract

This paper considers the R-optimal design problem for a linear Haar-wavelet regression model. It is proved that the proposed designs are R-optimal by means of the equivalence theorem.

Key words: Haar-wavelet; R-optimal designs; Equivalence theorem.

AMS Subject Classifications: 62K05, 62J05

1. Introduction

An extensive literature review reveals that the wavelet models are gradually becoming popular from theoretical and application point of view. Mention may be made in this regard to Herzberg and Traves (1994), Oyet and Wiens (2000), Oh, Naveau and Lee (2001), Oyet (2002), Xie (2002), Tian and Herzberg (2006, 2007) and Maronge *et al.* (2017) for the novel study of various linear and approximately linear wavelet models and their optimal designs.

Tian and Herzberg (2007) rightly pointed out that wavelets can be considered as a basis for representing square integrable functions in different scales in the same way as polynomials, trigonometric functions, rational functions can be. Haar (1910) pioneered the notion of wavelet system on the real line $\mathbb{R}$. The system is an orthogonal basis in $L_2(\mathbb{R})$ generated by the Haar scaling function $\phi(x)$ and the Haar primary wavelet $\psi(x)$, where

$$\phi(x) = \begin{cases} 1, & 0 \le x < 1, \\ 0, & \text{otherwise}, \end{cases} \qquad \psi(x) = \begin{cases} 1, & 0 \le x < \frac{1}{2}, \\ -1, & \frac{1}{2} \le x < 1, \\ 0, & \text{otherwise}. \end{cases}$$

It is to be noted that Haar-wavelets are piecewise constant functions on the real line $\mathbb{R}$ and can take only three values. Moreover, Haar-wavelets, like the well-known Wash functions (Rao 1983), form an orthogonal and complete set of functions representing discretized functions and piecewise constant functions. Tian and Herzberg (2007) investigated the linear Haar-wavelet models and obtained the D-, A- and E-optimal designs.

The R-optimality criterion was introduced by Dette (1997) and it minimizes the volume of the rectangular confidence region for the regression parameters based on the Bonferroni t-intervals. The R-optimal design problem has been investigated in linear models [see He and Yue (2019) for

Correponding Author: Kashinath Chatterjee
Email: kashinathchatterjee@gmail.com

recent reference]. Our aim here is to consider the R-optimal designs for the linear Haar-wavelet regression model.

The rest of this paper is organized as follows. In Section 2, we introduce the notation and preliminaries. Section 3 provides the main result of this paper with an example.

2. Model and Preliminaries

Consider the linear Haar-wavelet model of order m

$$E[y(x)] = \beta_0 + \sum_{j=0}^{m} \sum_{k=0}^{2^j-1} \beta_{jk}\psi_{jk}(x), \quad x \in \mathcal{X} = [0,1], \tag{1}$$

with unknown parameters $\beta_0, \beta_{00}, \cdots, \beta_{m,2^m-1}$, where

$$\psi_{jk}(x) = 2^{j/2}\psi(2^j x - k) = \begin{cases} 2^{j/2}, & \frac{k}{2^j} \le x < \frac{k}{2^j} + \frac{1}{2^{j+1}}, \\ -2^{j/2}, & \frac{k}{2^j} + \frac{1}{2^{j+1}} \le x < \frac{k}{2^j} + \frac{1}{2^j} \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

for $j \in \{0, 1, \cdots, m\},\ k \in \{0, 1, \cdots, 2^j - 1\}$.

Throughout the paper we consider approximate designs of the form

$$\xi = \begin{Bmatrix} x_1 & \cdots & x_n \\ w_1 & \cdots & w_n \end{Bmatrix}, \quad x_i \in \mathcal{X}, \quad 0 < w_i < 1, \quad \sum_{i=1}^{n} w_i = 1.$$

Denote the set of all approximate designs with non-singular information matrix on $\mathcal{X}$ by Ξ. For the model (1) the information matrix of $\xi \in \Xi$ is

$$M(\xi) = \int_{\mathcal{X}} \boldsymbol{f}(x)\boldsymbol{f}^T(x)d\xi(x), \tag{3}$$

where $\boldsymbol{f}(x) = (1, \psi_{00}(x), \cdots, \psi_{m,2^m-1}(x))^T$.

The following definition, due to Dette (1997), provides the R-optimality criterion for a design belonging to Ξ.

Definition 1: A design $\xi^* \in \Xi$ is called R-optimal for the model (1) if it minimizes

$$\Psi(\xi) = \prod_{i=1}^{p} \left(M^{-1}(\xi)\right)_{ii} = \prod_{i=1}^{p} e_i^T M^{-1}(\xi) e_i \tag{4}$$

over Ξ, where p is the dimension of the regression vector $\boldsymbol{f}(x)$ and e_i denotes the ith unit vector in $\mathbb{R}^p$.

The following equivalence theorem provides an important tool for the determination of R-optimal designs which has been proved by Dette (1997).

Theorem 1: For the model (1) let

$$\phi(\mathbf{x},\xi) = \boldsymbol{f}^T(\mathbf{x})M^{-1}(\xi)\left(\sum_{i=1}^{p}\frac{\boldsymbol{e}_i\boldsymbol{e}_i^T}{\boldsymbol{e}_i^T M^{-1}(\xi)\boldsymbol{e}_i}\right)M^{-1}(\xi)\boldsymbol{f}(\mathbf{x}). \tag{5}$$

Then a design $\xi^* \in \Xi$ is R-optimal if and only if

$$\sup_{\mathbf{x}\in\mathcal{X}}\phi(\mathbf{x},\xi^*) = p. \tag{6}$$

Moreover, the supremum is achieved at the support points of ξ^*.

3. R-optimal Designs

The following theorem provides R-optimal designs for the linear Haar-wavelet regression model (1).

Theorem 2: For the model (1), let x_i be arbitrary point in

$$\mathcal{X}_i = \left[\frac{i-1}{2^{m+1}}, \frac{i}{2^{m+1}}\right), \quad i = 1, \cdots, 2^{m+1}.$$

Then the design ξ^* of the form

$$\xi^* = \begin{pmatrix} x_1 & \cdots & x_{2^{m+1}} \\ \frac{1}{2^{m+1}} & \cdots & \frac{1}{2^{m+1}} \end{pmatrix} \tag{7}$$

is R-optimal.

Proof: It is to be noted that $\psi_{jk}(x)$'s are step functions and for any $x \in \mathcal{X}_i$, $(i = 1, \cdots, 2^{m+1})$

$$\psi_{jk}(x) \;=\; \psi_{jk}(\mu_i)$$

and

$$\boldsymbol{f}(x) \;=\; \boldsymbol{f}(\mu_i),$$

where $\mu_i = (i-1)2^{-(m+1)}$. For any design $\xi^* \in \Xi$ of the form (7), we have

$$M(\xi^*) \;=\; \frac{1}{2^{m+1}}\sum_{i=1}^{2^{m+1}}\boldsymbol{f}(\mu_i)\boldsymbol{f}^T(\mu_i) \;=\; \frac{1}{2^{m_2+1}}\left(M_{s_1,s_2}\right)_{1\le s_1,s_2\le 2^{m+1}},$$

where $M_{00} = 2^{m+1}$ and for $s_r = 2^{j_r} + k_r$ with $j_r \in \{0, 1, \cdots, m\}$, $k_r \in \{0, 1, \cdots, 2^{j_r} - 1\}$ and $r = 1, 2$,

$$M_{0s_2} \;=\; \sum_{i=1}^{2^{m+1}}\psi_{j_2k_2}(\mu_i), \quad M_{s_10} \;=\; \sum_{i=1}^{2^{m+1}}\psi_{j_1k_1}(\mu_i), \quad M_{s_1s_2} \;=\; \sum_{i=1}^{2^{m+1}}\psi_{j_1k_1}(\mu_i)\psi_{j_2k_2}(\mu_i).$$

Now, we can get

$$\begin{aligned} M_{s_1 0} &= \sum_{i=1}^{2^{m+1}} \psi_{j_1 k_1}(\mu_i) = 2^{j_1/2} \sum_{i=1}^{2^{m+1}} \psi(2^{j_1}\mu_i - k_1) = 2^{j_1/2} \sum_{i=1}^{2^{m+1}} \psi((i-1)/2^{m-j_1+1} - k_1) \\ &= 2^{j_1/2} \sum_{i=2^{m-j_1+1}k_1+1}^{2^{m-j_1+1}(k_1+1)} \psi((i-1)/2^{m-j_1+1} - k_1) = 0, \end{aligned}$$

and similarly, $M_{0 s_2} = 0$. Moreover,

$$\begin{aligned} M_{s_1 s_1} &= \sum_{i=1}^{2^{m+1}} (\psi_{j_1 k_1}(\mu_i))^2 = 2^{j_1} \sum_{i=1}^{2^{m+1}} (\psi(2^{j_1}\mu_i - k_1))^2 = 2^{j_1} \sum_{i=1}^{2^{m+1}} (\psi((i-1)/2^{m-j_1+1} - k_1))^2 \\ &= 2^{j_1} \sum_{i=2^{m-j_1+1}k_1+1}^{2^{m-j_1+1}(k_1+1)} (\psi((i-1)/2^{m-j_1+1} - k_1))^2 = 2^{m+1}, \end{aligned}$$

and

$$\begin{aligned} M_{s_1 s_2} &= \sum_{i=1}^{2^{m+1}} \psi_{j_1 k_1}(\mu_i)\psi_{j_2 k_2}(\mu_i) = 2^{(j_1+j_2)/2} \sum_{i=1}^{2^{m+1}} \psi(2^{j_1}\mu_i - k_1)\psi(2^{j_2}\mu_i - k_2) \\ &= 2^{(j_1+j_2)/2} \sum_{i=1}^{2^{m+1}} \psi((i-1)/2^{m-j_1+1} - k_1)\psi((i-1)/2^{m-j_2+1} - k_2) = 0. \end{aligned}$$

Therefore, it is clear that $M(\xi^*) = I_{2^{m+1}}$, where I_n is the $n \times n$ identity matrix. It follows that, for any $x \in [0, 1]$,

$$\begin{aligned} \phi(x, \xi^*) &= \boldsymbol{f}^T(x) M^{-1}(\xi^*) \left(\sum_{i=1}^{p} \frac{\boldsymbol{e}_i \boldsymbol{e}_i^T}{\boldsymbol{e}_i^T M^{-1}(\xi^*) \boldsymbol{e}_i} \right) M^{-1}(\xi^*) \boldsymbol{f}(x) \\ &= \boldsymbol{f}^T(x) \boldsymbol{f}(x) = 1 + \sum_{j=0}^{m} \sum_{k=0}^{2^j - 1} \psi_{jk}^2(x) = 1 + \sum_{j=0}^{m_1} 2^j = 2^{m+1}, \end{aligned} \tag{8}$$

and then ξ^* is R-optimal from Theorem 1.

Example: Consider the R-optimal design for the Haar-wavelet model (1) of order $m = 2$, *i.e.*,

$$\begin{aligned} E[y(x)] = \beta_0 + \beta_{00}\psi_{00}(x) + \beta_{10}\psi_{10}(x) + \beta_{11}\psi_{11}(x) + \beta_{20}\psi_{20}(x) \\ + \beta_{21}\psi_{21}(x) + \beta_{22}\psi_{22}(x) + \beta_{23}\psi_{23}(x), \end{aligned} \tag{9}$$

In this case, the design region $\mathcal{X} = [0, 1]$ is divided into $2^{m+1} = 8$ sub-intervals

$$\mathcal{X}_i = \left[\frac{i-1}{8}, \frac{i}{8} \right), \quad i = 1, \cdots, 8.$$

From Theorem 2, a R-optimal design for the model (9) is as follows:

$$\xi^* = \begin{pmatrix} 0/8 & 1/8 & \cdots & 7/8 \\ 1/8 & 1/8 & \cdots & 1/8 \end{pmatrix}. \tag{10}$$

References

Dette, H. (1997). Designing experiments with respect to "standardized" optimality criteria. *Journal of the Royal Statistical Society*, **B 59**, 97-110.

Haar, A. (1990). Zur theorie der orthogonalen Funktionensysteme. *Mathematische Annalen*, **69**, 331-371.

Hardle, W., Kerkyacharian, G., Picard, D. and Tsybajov, A. (1998). *Wavelets, Approximation and Statistical Applications*, Springer, Berlin, New York.

He, L. and Yue, R-X. (2019). *R*-optimality criterion for regression models with asymmetric errors. *Journal of Statistical Planning and Inference,* **199**, 318-326.

Herzberg, A. M. and Traves, W. N. (1994). An optimal experimental design for the Haar regression model. *The Canadian Journal of Statistics*, **22**, 357-364.

Kerkyacharian, G. and Picard, D. (1996). Estimating nonquadratic functionals of a density using Haar wavelets. *Annals of Statistics*, **24**, 485-507.

Maronge, J. M., Zhai, Y., Wiens, D. P. and Fang, Z. D. (2017). Optimal designs for spline wavelet regression models. *Journal of Statistical Planning and Inference*, **184**, 94-104.

Oh, H-S., Naveau, P. and Lee, G. (2001). Polynomial boundary treatment for wavelet regression. *Biometrika*, **88**, 291-298.

Oyet, A. J. (2002). Minimax A- and D-optimal integer-valued wavelet designs for estimation. *The Canadian Journal of Statistics*, **30**, 301-306.

Oyet, A. J. and Wiens, D. P. (2000). Robust designs for wavelet approximations of regression models. *Journal of Nonparametric Statistics*, **12**, 837-859.

Rao, G. P. (1983). *Piecewise Constant Orthogonal Functions and their Application to Systems and Control*. Springer, Berlin Heidelberg, New York

Tian, Y. G. and Herzberg, A. M. (2006). A-minimax and D-minimax robust optimal designs for approximately linear Haar wavelet models. *Computational Statistics & Data Analysis*, **50**, 2942-2951.

Tian, Y. G. and Herzberg, A. M. (2007). Estimation and optimal designs for linear Haar wavelet models. *Metrika*, **65**, 311-324.

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021(New Series), pp 67–76

Nonparametric Estimation of Linear Multiplier for Processes Driven by Mixed Fractional Brownian Motion

B. L. S. Prakasa Rao
CR RAO Advanced Institute of Mathematics
Statistics and Computer Science, Hyderabad, India

Received: 01 August 2020; Revised: 30 September 2020; Accepted: 04 October 2020

Abstract

We study the problem of nonparametric estimation of linear multiplier function $\theta(t)$ for processes satisfying stochastic differential equations of the type

$$dX_t = \theta(t)X_t dt + \epsilon\, d\tilde{W}_t^H, X_0 = x_0, 0 \le t \le T$$

where $\{\tilde{W}_t^H, t \ge 0\}$ is a mixed fractional Brownian motion with known Hurst index H and study the asymptotic behaviour of the estimator as $\epsilon \to 0$.

Key words: Nonparametric estimation; Linear multiplier; Mixed Fractional Brownian motion.

AMS Subject Classifications: 62K05

1. Introduction

Professor Aloke Dey and I were colleagues for several years at the Indian Statistical Institute, Delhi Centre until I left due to my superannuation in the year 2004. Prof. Dey's expertise was in the area of optimal designs and my area of interest is in inference for stochastic processes. Even though our areas of research are completely different, we appreciated each others works and had a high regard for each other. I missed his association after I moved to Hyderabad. We did meet once or twice during the last sixteen years after I left New Delhi. I would like to thank Professor Vinod Gupta for inviting me to submit an article for the special issue of this journal dedicated to the memory of Professor Aloke Dey and pay my homage to a great statistician.

Statistical inference for fractional diffusion type processes satisfying stochastic differential equations driven by fractional Brownian motion have been studied earlier and a comprehensive survey of various methods is given in Mishura (2008) and Prakasa Rao (2010). There has been a recent interest to study similar problems for stochastic processes driven by a mixed fractional Brownian motion (mfBm). Existence and uniqueness for solutions of stochastic differential equations driven by a mfBm are investigated in Mishura and Shevchhenko (2012) and Shevchenko (2014) among others. Maximum likelihood estimation for estimation of drift parameter in a linear stochastic differential equations driven by a mfBm

Corresponding Author: B. L. S. Prakasa Rao
Email: blsprao@gmail.com

is investigated in Prakasa Rao (2018). The method of instrumental variable estimation for such parametric models is investigated in Prakasa Rao (2017). Some applications of such models in finance are presented in Prakasa Rao (2015 a,b). For related work on parametric inference for processes driven by mfBm, see Marushkevych (2016), Rudomino-Dusyatska (2003), Song and Liu (2014), Mishra and Prakasa Rao (2017), Prakasa Rao (2009) and Miao (2010) among others. Nonparametric estimation of the trend coefficient in models governed by stochastic differential equations driven by a mixed fractional Brownian motion is investigated in Prakasa Rao (2019).

We now discuss the problem of estimating the function $\theta(t), 0 \leq t \leq T$ (linear multiplier) based on the observations of a process $\{X_t, 0 \leq t \leq T\}$ satisfying the stochastic differential equation

$$dX_t = \theta(t)X_t dt + \epsilon\, d\tilde{W}_t^H, X_0 = x_0, 0 \leq t \leq T$$

where $\{\tilde{W}_t^H, t \geq 0\}$ is a mixed fractional Brownian motion (mfBm) and study the properties of the estimator as $\epsilon \to 0$.

2. Mixed Fractional Brownian Motion

We will now summarize some properties of stochastic processes which are solutions of stochastic differential equations driven by a mixed fractional Brownian motion for completeness.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ be a stochastic basis satisfying the usual conditions. The natural filtration of a stochastic process is understood as the P-completion of the filtration generated by this process. Let $\{W_t, t \geq 0\}$ be a standard Wiener process and $W^H = \{W_t^H, t \geq 0\}$ be an independent normalized fractional Brownian motion with Hurst parameter $H \in (0, 1)$, that is, a Gaussian process with continuous sample paths such that $W_0^H = 0, E(W_t^H) = 0$ and

$$E(W_s^H W_t^H) = \frac{1}{2}[s^{2H} + t^{2H} - |s - t|^{2H}], t \geq 0, s \geq 0. \tag{1}$$

Let

$$\tilde{W}_t^H = W_t + W_t^H, t \geq 0.$$

The process $\{\tilde{W}_t^H, t \geq 0\}$ is called the mixed fractional Brownian motion with Hurst index H. We assume here after that Hurst index H is known. Following the results in Cheridito (2001), it is known that the process $\tilde{W}^H$ is a semimartingale in its own filtration if and only if either $H = 1/2$ or $H \in (\frac{3}{4}, 1]$.

Let us consider a stochastic process $X = \{X_t, t \geq 0\}$ defined by the stochastic integral equation

$$X_t = \int_0^t C(s)ds + \tilde{W}_t^H, t \geq 0 \tag{2}$$

where the process $C = \{C(t), t \geq 0\}$ is an $(\mathcal{F}_t)$-adapted process. For convenience, we write the above integral equation in the form of a stochastic differential equation

$$dX_t = C(t)dt + d\tilde{W}_t^H, t \geq 0 \tag{3}$$

driven by the mixed fractional Brownian motion $\tilde{W}^H$. Following the recent works by Cai *et al.* (2016) and Chigansky and Kleptsyna (2015), one can construct an integral transformation that transforms the mixed fractional Brownian motion $\tilde{W}^H$ into a martingale M^H. Let $g_H(s,t)$ be the solution of the integro-differential equation

$$g_H(s,t) + H\frac{d}{ds}\int_0^t g_H(r,t)|s-r|^{2H-1}sign(s-r)dr = 1, 0 < s < t. \tag{4}$$

Cai *et al.* (2016) proved that the process

$$M_t^H = \int_0^t g_H(s,t)d\tilde{W}_s^H, t \geq 0 \tag{5}$$

is a Gaussian martingale with quadratic variation

$$< M^H >_t = \int_0^t g_H(s,t)ds, t \geq 0 \tag{6}$$

Furthermore the natural filtration of the martingale M^H coincides with that of the mixed fractional Brownian motion $\tilde{W}^H$. It is clear that the quadratic variation $< M^H >_t$ is differentiable with respect to t. Let $\beta(t)$ denote the derivative of the function $< M^H >_t$ with respect to t. Suppose that, for the martingale M^H defined by the equation (6), the sample paths of the process $\{C(t), t \geq 0\}$ are smooth enough in the sense that the process

$$Q_H(t) = \frac{d}{d < M^H >_t}\int_0^t g_H(s,t)C(s)ds, t \geq 0 \tag{7}$$

is well defined. Define the process

$$Z_t = \int_0^t g_H(s,t)dX_s, t \geq 0. \tag{8}$$

As a consequence of the results in Cai *et al.* (2016), it follows that the process Z is a fundamental semimartingale associated with the process X in the following sense.

Theorem 1: Let $g_H(s,t)$ be the solution of the equation (4). Define the process Z as given in the equation (8). Then the following relations hold.

(i) The process Z is a semimartingale with the decomposition

$$Z_t = \int_0^t Q_H(s)d < M^H >_s + M_t^H, t \geq 0 \tag{9}$$

where M^H is the martingale defined by the equation (5). (ii) The process X admits the representation

$$X_t = \int_0^t \hat{g}_H(s,t)dZ_s, t \geq 0 \tag{10}$$

where

$$\hat{g}_H(s,t) = 1 - \frac{d}{d < M^H >_s} \int_0^t g_H(r,s)dr. \tag{11}$$

(iii) The natural filtrations $(\mathcal{X}_t)$ and $(\mathcal{Z}_t)$ of the processes X and Z respectively coincide.

Applying the Corollary 2.9 in Cai *et al.* (2016), it follows that the probability measures μ_X and $\mu_{\tilde{W}^H}$ generated by the processes X and $\tilde{W}^H$ on an interval $[0,T]$ are absolutely continuous with respect to each other and the Radon-Nikodym derivative is given by

$$\frac{d\mu_X}{d\mu_{\tilde{W}^H}} = \exp[\int_0^T Q_H(s)dZ_s - \frac{1}{2}\int_0^T [Q_H(s)]^2 d < M^H >_s] \tag{12}$$

which is also the likelihood function based on the observation $\{X_s, 0 \leq s \leq T.\}$ Since the filtrations generated by the processes $X$and Z are the same, the information contained in the families of σ-algebras $(\mathcal{X}_t)$ and $(\mathcal{Z}_t)$ is the same and hence the problem of the estimation of the parameters involved based on the observation $\{X_s, 0 \leq s \leq T\}$ and $\{Z_s, 0 \leq s \leq T\}$ are equivalent.

3. Preliminaries

Let $\tilde{W}^H = \{W_t^H, t \geq o\}$ be a mixed fractional Brownian motion with known Hurst parameter $H \in (1/2, 1)$. Consider the problem of estimating the function $\theta(t), 0 \leq t \leq T$ (linear multiplier) from the observations $\{X_t, 0 \leq t \leq T\}$ of process satisfying the stochastic differential equation

$$dX_t = \theta(t)X_t dt + \epsilon\, d\tilde{W}_t^H, X_0 = x_0, 0 \leq t \leq T \tag{13}$$

and study the properties of the estimator as $\epsilon \to 0$. Consider the differential equation in the limiting system of (13), that is , for $\epsilon = 0$, given by

$$dx_t = \theta(t)x_t dt, x_0, 0 \leq t \leq T. \tag{14}$$

Observe that

$$x_t = x_0 \exp\{\int_0^t \theta(s)ds).$$

We assume that the following condition holds:

(A_1): The trend coefficient $\theta(t)$ over the interval $[0,T]$ is bounded by a constant L.

The condition (A_1) will ensure the existence and uniqueness of the solution of the equation (13).

Lemma 1: Let the condition (A_1) hold and $\{X_t, 0 \leq t \leq T\}$ and $\{x_t, 0 \leq t \leq T\}$ be the solutions of the equations (13) and (14) respectively. Then, with probability one,

$$(a)|X_t - x_t| < e^{Lt}\epsilon \sup_{0 \leq s \leq t} |\tilde{W}_s^H| \tag{15}$$

and

$$(b) \sup_{0\le t\le T} E(X_t - x_t)^2 \le 4e^{2LT}\epsilon^2(T^{2H} + T). \tag{16}$$

Proof of (a): Let $u_t = |X_t - x_t|$. Then by (A_1); we have,

$$\begin{aligned} u_t &\le \int_0^t |\theta(v)(X_v - x_v)|dv + \epsilon|\tilde{W}_t^H| \\ &\le L\int_0^t u_v dv + \epsilon \sup_{0\le s\le t} |\tilde{W}_s^H|. \end{aligned} \tag{17}$$

Applying the Gronwall's lemma (cf. Lemma 1.12, Kutoyants (1994), p.26), it follows that

$$u_t \le \epsilon \sup_{0\le s\le t} |\tilde{W}_s^H| e^{Lt}. \tag{18}$$

Proof of (b): From the equation (15), we have

$$\begin{aligned} E(X_t - x_t)^2 &\le e^{2Lt}\epsilon^2 E[(\sup|\tilde{W}_s^H|)^2] \\ &\le 4e^{2Lt}\epsilon^2(t^{2H} + t) \end{aligned} \tag{19}$$

from the fact that the mixed fractional Brownian motion W^H is a sum of a Wiener process and fractional Brownian motion and from the maximal inequalities for a Wiener process and a fractional Brownian motion (cf. Muneya and Shieh (2009), Prakasa Rao (2014)). Hence

$$\sup_{0\le t\le T} E(X_t - x_t)^2 \le 4e^{2LT}\epsilon^2(T^{2H} + T). \tag{20}$$

This completes the proof of the lemma.

Define

$$\begin{aligned} Q_{H,\theta}^*(t) &= \frac{d}{d < M^H >_t} \int_0^t g_H(t,s)\theta(s)x(s)ds \\ &= \frac{d}{d < M^H >_t} \int_0^t g_H(t,s)\theta(s)[x_0 \exp(\int_0^s \theta(u)du)]ds \end{aligned} \tag{21}$$

by using the equation (14). Here after, we consider the problem of nonparametric estimation of the function $Q_{H,\theta}^*(t)$ instead of the function $\theta(t)$. We assume that the function $\theta(.)$ belongs to a class of functions Θ uniformly bounded by a constant L and the following condition holds:
(A_2): Differentiation under the integral sign is valid in the equation (21) and the function $\beta(t)Q_{H,\theta}^*(t)$is Lipschitz of order γ in the sense that

$$|\beta(t)Q_{H,\theta}^*(t) - \beta(s)Q_{H,\theta}^*(s)| \le C|t - s|^\gamma$$

for some constant $C > 0$ and $\gamma > 0$ uniformly for $\theta(.) \in \Theta$.

Instead of estimating the function $\theta(.)$, we consider the problem of estimating the function $Q^*_{H,\theta}(.)$ defined via the equation (21). This is justified by the observation that the processes $\{X_t, 0 \leq t \leq T\}$ governed by the stochastic differential equation (13) and the corresponding related process $\{Z_t, 0 \leq t \leq T\}$ as defined by (8) have the same filtrations by the results in Cai *et al.* (2016).

Consider the kernel type estimator defined by

$$\begin{aligned} \widehat{Q}_{H,\theta}(t) &= \frac{1}{h_\epsilon}\int_0^T G\left(\frac{s-t}{h_\epsilon}\right) dZ_s \\ &= \frac{1}{h_\epsilon}\int_0^T G\left(\frac{s-t}{h_\epsilon}\right)\left(Q_{H,\theta}(s) d < M^H >_s + \epsilon\, dM_s^H\right) \\ &= \frac{1}{h_\epsilon}\int_0^T G\left(\frac{s-t}{h_\epsilon}\right)\left(Q_{H,\theta}(s)\beta(s)ds + \epsilon\, dM_s^H\right) \end{aligned} \tag{22}$$

by using the equation (9) where $G(u)$ is a bounded function with finite support $[A, B]$ satisfying the condition

(A_3):$G(u) = 0$ for $u < A, u > B$, $\int_A^B |G(u)|du < \infty$ and $\int_A^B G(u)du = 1$;

Consider a normalizing function $h_\epsilon \to 0$ as $\epsilon \to 0$. In addition, suppose that $\epsilon^2 h_\epsilon^{-3/2} \to 0$ as $\epsilon \to 0$.

4. Main Results

Theorem 2 : Suppose the conditions $(A_1), (A_2)$ and (A_3) are satisfied. Then the estimator $\widehat{Q}_{H,\theta}(t)$ is uniformly consistent, that is,

$$\lim_{\epsilon \to 0} \sup_{\theta(.) \in \Theta} \sup_{0 \leq t \leq T} E_\theta(|\widehat{Q}_{H,\theta}(t) - \beta(t)Q^*_{H,\theta}(t)|^2) = 0. \tag{23}$$

Proof: From (9), we have,

(24)

$$\begin{aligned} & E_\theta|\widehat{Q}_{H,\theta}(t) - \beta(t)Q^*_{H,\theta}(t)|^2 \\ &= E\left|\frac{1}{h_\epsilon}\int_0^T G\left(\frac{s-t}{h_\epsilon}\right)(Q_{H,\theta}(s)\beta(s)ds + \epsilon dM_s^H) - \beta(t)Q^*_{H,\theta}(t)\right|^2 \\ &= E_\theta|\frac{1}{h_\epsilon}\int_0^T G\left(\frac{s-t}{h_\epsilon}\right)(Q_{H,\theta}(s) - Q^*_{H,\theta}(s))\beta(s)ds \\ &\quad + \frac{1}{h_\epsilon}\int_0^T G\left(\frac{s-t}{h_\epsilon}\right)(Q^*_{H,\theta}(s)\beta(s) - Q^*_{H,\theta}(t)\beta(t))ds \\ &\quad + \frac{\epsilon}{h_\epsilon}\int_0^T G\left(\frac{s-t}{h_\epsilon}\right) dM_s^H|^2 \end{aligned}$$

$$
\begin{aligned}
&= E_\theta[I_1 + I_2 + I_3]^2 \text{ (denoting the three integrals as } I_1, I_2 \text{ and } I_3 \text{ respectively)} \\
&\leq 3\, E(I_1^2) + 3\, E(I_2^2) + 3\, E(I_3^2).
\end{aligned}
$$

Now

$$
\begin{aligned}
3\, E_\theta[I_1^2] &= 3\, E_\theta \left| \frac{1}{h_\epsilon} \int_0^T G\left(\frac{s-t}{h_\epsilon}\right) (Q_{H,\theta}(t) - Q^*_{H,\theta}(s))\beta(s) ds \right|^2 \qquad (25) \\
&\leq \frac{3}{h_\epsilon^2} [\int_0^T G^2\left(\frac{s-t}{h_\epsilon}\right) ds][E \int_0^T \beta^2(s)(Q_{H,\theta}(s) - Q^*_{H,\theta}(s))^2 ds].
\end{aligned}
$$

Note that

$$
\begin{aligned}
&E_\theta \int_0^T \beta^2(s)(Q_{H,\theta}(s) - Q^*_{H,\theta}(s))^2 d < M^H >_s \qquad (26) \\
&= \int_0^T \beta^2(s) E_\theta \left[\frac{d}{d < M^H >_s} \int_0^s g_H(s,v)\theta(v)(X(v) - x(v)) dv \right]^2 d < M^H >_s \\
&\leq C_1 \int_0^T E_\theta \left[\int_0^s \frac{\partial g_H(s,v)}{\partial s} \theta(v)(X(v) - x(v)) dv \right]^2 ds \\
&\leq C_2 \int_0^T \{ \int_0^s \left(\frac{\partial g_H(s,v)}{\partial s} \right)^2 \theta^2(v) dv \int_0^s E(X(v) - x(v))^2 dv \} ds
\end{aligned}
$$

for some positive constant C_2 depending on T and H. Furthermore

$$E_\theta(X_v - x_v)^2 \leq 4e^{2Lv}\epsilon^2(v^{2H} + v) \text{ (by Lemma 1).}$$

Hence, from the equation (26) and the condition (A_3), we get that

$$
\begin{aligned}
3E_\theta[I_1^2] &\leq C \frac{1}{h_\epsilon^2} \left\{ \int_{-\infty}^{\infty} G^2(\frac{s-t}{h_\epsilon}) \beta(s) ds \right\} \epsilon^2 h_\epsilon \qquad (27) \\
&\qquad \times \int_0^T \beta^2(s) \left\{ \int_0^s e^{2Lv}(v^{2H} + v) dv \right\} \left\{ \int_0^s \left(\frac{\partial g_H(s,v)}{\partial s} \right)^2 dv \right\} ds \\
&\leq C_3 \epsilon^2 h_\epsilon^{-1}
\end{aligned}
$$

for some positive constant C_3 depending on T and H and the last term tends to zero as $\epsilon \to 0$.

In addition,

$$
\begin{aligned}
I_2^2 &= 3 \left\{ \frac{1}{h_\epsilon} \int_0^T G\left(\frac{s-t}{h_\epsilon}\right) (Q^*_{H,\theta}(s) ds - Q^*_{H,\theta}(t)) d < M^H >_s \right\}^2 \qquad (28) \\
&= \left\{ \frac{1}{h_\epsilon} \int_0^T G\left(\frac{s-t}{h_\epsilon}\right) (Q^*_{H,\theta}(s)\beta(s) - Q^*_{H,\theta}(t)\beta(t)) ds \right\}^2 \\
&= 3 \left\{ \int_{-\infty}^{\infty} G(u)(Q^*_{H,\theta}(t + h_\epsilon u)\beta(t + h_\epsilon u) - Q^*_{H,\theta}(t)\beta(t)) du \right\}^2 \quad (\text{by}(A_2))
\end{aligned}
$$

$$\begin{aligned} &\leq C_4 \left\{ \int_{-\infty}^{\infty} G(u) |h_\epsilon u|^\gamma du \right\}^2 \quad (\text{by}(A_2)) \\ &\leq C_4 h_\epsilon^{2\gamma} \left(\int_{-\infty}^{\infty} G(u) |u|^\gamma du \right)^2 \\ &\leq C_5 h_\epsilon^{2\gamma} \quad \text{by}(A_3)) \end{aligned}$$

for some positive constant C_5 depending on T and $H$and the last term tends to zero as $\epsilon \to 0$. Furthermore

$$\begin{aligned} I_3^2 &= \frac{3\epsilon^2}{h_\epsilon^2} E \left(\int_0^T G\left(\frac{s-t}{h_\epsilon} \right) dM_s^H \right)^2 \\ &= \frac{3\epsilon^2}{h_\epsilon^2} \int_0^T G^2 \left(\frac{s-t}{h_\epsilon} \right) \beta(s) ds \\ &\leq \frac{3\epsilon^2}{h_\epsilon^2} \left\{ \int_0^T G^2 \left(\frac{s-t}{h_\epsilon} \right) ds \int_0^T \beta^2(s) ds \right\}^{\frac{1}{2}} \\ &\leq C_6 \frac{3\epsilon^2}{h_\epsilon^2} \left\{ h_\epsilon (\int_{-\infty}^{\infty} G^2(u)\ du) \right\}^{\frac{1}{2}} \\ &\leq C_7 \epsilon^2 h_\epsilon^{-3/2}. \end{aligned} \tag{29}$$

for some positive constants C_7 depending on T and H. The result follows from the equations (27), (28) and (29).

Corollary 1: Under the conditions $(A_1), (A_2)$ and (A_3),

$$\lim_{\epsilon \to 0} \sup_{\theta(.) \in \Theta} E \left\{ \widehat{Q}_{H,\theta}(t) - \beta(t) Q^*_{H,\theta}(t) \right\}^2 \overline{\epsilon}^{\frac{8\gamma}{4\gamma+3}} < \infty.$$

Proof: From the inequalities derived in (27), (28) and (29), we get that there exist positive constants D_1, D_2 and D_3 depending on T and H such that

$$\sup_{\theta(.) \in \Theta} E \left\{ \widehat{Q}_{H,\theta}(t) - \beta(t) Q^*_{H,\theta}(t) \right\}^2 \leq D_1 \epsilon^2 h_\epsilon^{-1} + D_2 h_\epsilon^{2\gamma} + D_3 \epsilon^2 h_\epsilon^{-\frac{3}{2}}. \tag{30}$$

Let $h_\epsilon = \epsilon^\beta, 0 < \beta < \frac{4}{3}$. Then the condition $h_\epsilon^{2\gamma} = \epsilon^2 h_\epsilon^{-3/2}$ leads to the choice $\beta = \frac{4}{4\gamma+3}$ and we get an optimum bound in (30) and hence

$$\lim_{\epsilon \to 0} \sup_{\theta(.) \in \Theta} E \left[\widehat{Q}_{H,\theta}(t) - \beta(t) Q^*_{H,\theta}(t) \right]^2 \epsilon^{-\frac{8\gamma}{4\gamma+3}} \leq C \tag{31}$$

for some positive constant C which implies the result.

Acknowledgements

The author thanks the referee for the incisive comments. This work was supported by the Indian National Science Academy (INSA) under the scheme "INSA Senior Scientist" at the CR RAO Advanced Institute of Mathematics, Statistics and Computer Science, Hyderabad 500046, India.

References

Cai, C., Chigansky, P. and Kleptsyna, M. (2016). Mixed Gaussian processes. *Annals of Probability*, **44**, 3032–3075.

Cheridito, P. (2001). Mixed fractional Brownian motion. *Bernoulli*, **7**, 913–934.

Chigansky, P. and Kleptsyna, M. (2015). Statistical analysis of the mixed fractional Ornstein-Uhlenbeck process. arXiv:1507.04194.

Kutoyants, Y. A.(1994) *Identification of Dynamical Systems with Small Noise*. Kluwer, Dordrecht.

Marushkevych, Dmytro. (2016). Large deviations for drift parameter estimator of mixed fractional Ornstein-Uhlenbeck process. *Modern Stochastics : Theory and Applications*, **3**, 107–117.

Miao, Y. (2010). Minimum L_1-norm estimation for mixed Ornstein-Uhlenbeck type process. *Acta Mathematica Vietnamica*, **35**, 379–386.

Mishura, Y. (2008).*Stochastic Calculus for Fractional Brownian Motion and Related Processes*. Springer, Berlin.

Mishura, Y. and Shevchenko, G. (2012). Existence and uniqueness of the solution of stochastic differential equation involving Wiener process and fractional Brownian motion with Hurst index $H > 1/2$. *Computers and Mathematics with Applications*, **64**, 3217–3227.

Mishra, M. N. and Prakasa Rao, B. L. S. (2017). Large deviation probabilities for maximum likelihood estimator and Bayes estimator of a parameter for mixed fractional Ornstein-Uhlenbeck type process. *Bulletin of Informatics and Cybernetics*, **49**, 67–80.

Muneya, M. and Shieh, N. -R. (2009). On the exponentials of fractional Ornstein-Uhlenbeck processes. *Electron Journal of Probability*, **14**, 594–611.

Prakasa Rao, B. L. S. (2009). Estimation for stochastic differential equations driven by mixed fractional Brownian motion. *Calcutta Statistical Association Bulletin*, **61**, 143–153.

Prakasa Rao, B. L. S. (2010). *Statistical Inference for Fractional Diffusion Processes*. Wiley, Chichester.

Prakasa Rao, B. L. S. (2014). Maximal inequalities for fractional Brownian motion: an overview. *Stochastic Analysis and Applications*, **32**, 450–479.

Prakasa Rao, B. L. S. (2015a). Option pricing for processes driven by mixed fractional Brownian motion with superimposed jumps. *Probability in the Engineering and Information Sciences*, **29**, 589–596.

Prakasa Rao, B. L. S. (2015b). Pricing geometric Asian power options under mixed fractional Brownian motion environment. *Physica A*, **446**, 92–99.

Prakasa Rao, B. L. S. (2017). Instrumental variable estimation for a linear stochastic differential equation driven by a mixed fractional Brownian motion. *Stochastic Analysis and Applications*, **35**, 943–953.

Prakasa Rao, B. L. S. (2018). Parameter estimation for linear stochastic differential equations driven by mixed fractional Brownian motion, *Stochastic Analysis and Applications*. **36**, 767–781.

Prakasa Rao, B. L. S. (2019). Nonparametric estimation of trend for stochastic differential equations driven by mixed fractional Brownian motion. *Stochastic Analysis and Applications*, **37**, 271–280.

Rudomino-Dusyatska, N. (2003). Properties of maximum likelihood estimates in diffusion and fractional Brownian models. *Theory of Probability and Mathematical Statistics*, **68**, 139–146.

Song, N. and Liu, Z. (2014). Parameter estimation for stochastic differential equations driven by mixed fractional Brownian motion. *Abstract and Applied Analysis* : 2014 Article ID 942307, 6 pp.

Shevchenko, G. (2014). Mixed stochastic delay differential equations. *Theory of Probability and Mathematical Statistics*, **89**, 181–195.

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 77–93

Tests for Equality of Hazard Quantile Functions

Pooja Soni[1], **Isha Dewan**[2] **and Kanchan Jain**[3]
[1] *University Business School,Panjab University, Chandigarh*
[2] *Indian Statistical Institute, New Delhi*
[3] *Department of Statistics, Panjab University, Chandigarh*

Received: 01 August 2020; Revised: 07 October 2020; Accepted: 12 October 2020

Abstract

In this paper, we propose two tests for testing equality of hazard quantile functions of two populations. The test statistics are based on estimators of the quantile density function. Limiting distribution of both these test statistics has been derived. The power of the new tests is computed through simulations for uncensored and censored observations. The new tests are compared with two existing tests available in literature. Procedures have been illustrated on real data.

Key words: Quantile function; Quantile density function; Kernel density estimator; Hazard rate; Hazard quantile function.

Reminiscences

Isha had known Prof. Aloke Dey for almost 36 years. She could walk into his office anytime to discuss statistics, official and even personal problems. He was a good listener, had a great sense of humor and was kind of a quick reference point on government rules and regulations. Three of us have fond memories of our association with Prof. Aloke Dey - a gentle person - went too soon.

1. Introduction

In survival analysis, the hazard rate is a basic reliability measure. It is studied as failure rate in reliability, force of mortality in demography or actuarial science, intensity function in stochastic processes and age specific failure rate in epidemiology. Sometimes, interest may be in comparing the hazard rates of two populations. Chikkagoudar and Shuster (1974) proposed the locally most powerful test for testing equality of hazard rates of two populations. Kochar (1979) provided distribution free test based on U-statistics and Kochar (1981) proposed a test based on linear function of order statistics for testing equality of hazard rates. For the same problem, Cheng (1985) proposed a test based on ranks.

Quantile based approach is popular now a days. The reliability analysis based on quantiles provides an alternate methodology for statistical analysis when cumulative distribution function (cdf) or probability density function (pdf) is not available in a closed form. Examples of such distributions are Generalised Lambda distribution (GLD) (Karian and Dudewicz

Correponding Author: Isha Dewan
Email: isha@isid.ac.in

(2000)), Skew logistic distribution (Gilchrist (2000)) and Davies distribution (Hankin and Lee (2006)). Maladan and Sankaran (2020) proposed a new family of distributions by using transformation in context of quantiles.

Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ be two independent random samples from two populations with distribution functions $F(x)$ and $G(x)$, survival functions $\bar{F}(x)$ and $\bar{G}(x)$, pdfs $f(x)$ and $g(x)$, hazard rate functions $h_1(x) = \dfrac{f(x)}{\bar{F}(x)}$ and $h_2(x) = \dfrac{g(x)}{\bar{G}(x)}$, respectively. The quantile function for the first population is denoted by $Q_1(u)$ and defined as

$$Q_1(u) = F^{-1}(u) = \inf\{x : F(x) \geq u\},\ 0 < u < 1. \tag{1}$$

From (1), it is seen that

$$F(Q_1(u)) = u. \tag{2}$$

Differentiating (2), we get the quantile density function for the first population as

$$q_1(u) = \frac{d}{du} Q_1(u) = \frac{1}{f(Q_1(u))}. \tag{3}$$

Note that the quantile density function as defined in (3) is not a density function in the usual sense but is reciprocal of density function at corresponding quantile function. Nair and Sankaran (2009a) presented various reliability measures *viz*, hazard rate, mean residual life function, variance residual life function and percentile residual life function in terms of quantiles. The hazard quantile function for the first population is given by

$$H_1(u) = h_1(Q_1(u)) = \frac{f(Q_1(u))}{\bar{F}(Q_1(u))} = ((1-u)q_1(u))^{-1}. \tag{4}$$

Hazard quantile function is the hazard function at the corresponding quantile function. The quantile function, quantile density function and hazard quantile function for second population are denoted by $Q_2(u)$, $q_2(u)$ and $H_2(u)$, respectively.

Many ageing concepts *viz* increasing failure rate, increasing failure rate average, new better than used and new better than used in expectation have been defined in terms of quantiles by Kumar and Nair (2011). Nair and Sankaran (2009b) studied estimation of the hazard quantile function based on right censored data. Peng and Fine (2007) provided tests for equality of cause specific hazard rates for competing risk data based on quantiles. Fan *et al.* (2020) proposed smooth kernel type estimator of quantile function for right-censored competing risks data.

We wish to test the null hypothesis of equality of hazard rate functions of two independent populations, that is

$$\begin{aligned} &H_0 : h_1(x) = h_2(x) \quad \text{for all } x \\ &\text{against the alternative} \\ &H_A : h_1(x) \leq h_2(x) \quad \text{for all } x \\ &\text{with strict inequality in (5) with a positive probability.} \end{aligned} \tag{5}$$

Kochar (1979) showed that for increasing failure rate distributions, location-scale ordering of distribution functions leads to ordering of their corresponding hazard rates.

Above testing problem can be equivalently written in terms of hazard quantile functions as follows

$$\begin{aligned} &H_0 : H_1(u) = H_2(u) \quad \text{for all } \ 0 < u < 1 \\ &\text{against the alternative} \\ &H_A : H_1(u) \leq H_2(u) \quad \text{for all } \ 0 < u < 1 \\ &\text{with strict inequality in (6) with a positive probability.} \end{aligned} \tag{6}$$

From (4), it is noted that for all $0 < u < 1$

$$\begin{aligned} H_1(u) = H_2(u) \quad &\text{iff} \quad q_1(u) = q_2(u), \\ H_1(u) \leq H_2(u) \quad &\text{iff} \quad q_1(u) \geq q_2(u). \end{aligned} \tag{7}$$

Hence, from (6) and (7), it is clear that testing for equality of hazard rates is equivalent to testing for equality of quantile density functions. Hence, we will propose tests for testing

$$\begin{aligned} &H_0 : q_1(u) = q_2(u) \quad \text{for all } \ 0 < u < 1 \\ &\text{against the alternative} \\ &H_A : q_1(u) \geq q_2(u) \quad \text{for all } \ 0 < u < 1 \\ &\text{with strict inequality in (8) with a positive probability.} \end{aligned} \tag{8}$$

In Section 2, we discuss few preliminaries that are needed to define and study the properties of test statistics. Two examples are given where distribution functions can not be expressed in closed forms but quantile functions have nice forms. We also discuss the estimator of quantile density function proposed by Soni *et al.* (2012). In Section 3, two test statistics - a supremum type and an integral type have been proposed for testing the equality of hazard quantile functions against the alternative that they are ordered. The statistics are based on estimators of quantile density functions due to Soni *et al.* (2012). Asymptotic distribution of two test statistics is discussed. The tests can be used when observations are uncensored or censored. Simulations are carried out in Section 4 for comparing power of the proposed tests with those suggested by Kochar (1979) and Cheng (1985). In Section 5, a real data set is considered to illustrate the utility of the tests proposed by us. The proofs of the Theorems and three Tables showing power comparisons are given in the Appendix.

2. Preliminaries

Two examples for which distribution function can not be written in a closed form but quantile function has a closed form, are discussed in Section 2.1. Estimator of quantile density function proposed by Soni *et al.* (2012) is discussed in Section 2.2.

2.1. Examples

(i) **Davies Distribution (Davies$(C, \lambda_1, \lambda_2)$)** with $C > 0,\ \lambda_1 > 0,\ \lambda_2 > 0$ was given by Hankin and Lee (2006). The quantile function, the quantile density function and the hazard quantile function for $0 < u < 1$ are

$$Q_D(u, C, \lambda_1, \lambda_2) = \frac{Cu^{\lambda_1}}{(1-u)^{\lambda_2}}, \tag{9}$$

$$q_D(u, C, \lambda_1, \lambda_2) = \frac{Cu^{\lambda_1 - 1}(\lambda_1(1-u) + \lambda_2 u)}{(1-u)^{\lambda_2 + 1}}, \tag{10}$$

$$H_D(u, C, \lambda_1, \lambda_2) = \frac{(1-u)^{\lambda_2}}{Cu^{\lambda_1 - 1}(\lambda_1(1-u) + \lambda_2 u)}. \tag{11}$$

(ii) **The Generalized Lambda Distribution** (GLD$(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$) was introduced by Ramberg and Schmeiser (1974) and further discussed by Karian and Dudewicz (2000). The quantile function, the quantile density function and the hazard quantile function for $0 < u < 1$ are given below:

$$Q_{GL}(u, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \lambda_1 + \frac{(u^{\lambda_3} - (1-u)^{\lambda_4})}{\lambda_2}, \tag{12}$$

$$q_{GL}(u, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \frac{\lambda_3 u^{\lambda_3 - 1} + \lambda_4 (1-u)^{\lambda_4 - 1}}{\lambda_2}, \tag{13}$$

$$H_{GL}(u, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \left((1-u)\frac{(\lambda_3 u^{\lambda_3 - 1} + \lambda_4 (1-u)^{\lambda_4 - 1})}{\lambda_2}\right)^{-1}. \tag{14}$$

The parameters $\lambda_1, \lambda_2, \lambda_3,$ and λ_4 can assume real values, but some restrictions on these parameters have been imposed for defining a valid distribution. The possible eight regions of parameter values for which GLD is a valid distribution have been listed in Karian and Dudewicz (2000). Table 1 gives two sets of choices of parameters λ_2, λ_3 and λ_4 of GLD with λ_1 taking any real value. These choices ensure that observations always have support on the positive real line.

Table 1: Considered regions and corresponding supports of GLD

	Regions	Supports
1.	$\lambda_2 > 0,\ \lambda_3 > 1,\ \lambda_4 > 0$	$\left(\lambda_1 - \frac{1}{\lambda_2}, \lambda_1 + \frac{1}{\lambda_2}\right)$
2.	$\lambda_2 < 0,\ \lambda_3 > 1,\ \lambda_4 < -1$	$\left(\lambda_1 - \frac{1}{\lambda_2}, \infty\right)$

Here λ_1 controls the left tail, λ_2 controls the right tail and C is the scale parameter.

These distributions will be used for simulation studies in Section 4.

2.2. Quantile density estimator

Estimators of the quantile density function were proposed by Parzen (1979), Csörgo (1981), Falk (1986), Jones (1992), Cheng and Parzen (1997) and Soni *et al.* (2012). The wavelet based estimator of quantile density function was proposed by Chesneau *et al.* (2016). This estimator behaved well in tails.

The estimator of $q_1(u)$ given by Soni *et al.* (2012), based on random sample $X_1, X_2, \ldots, X_n$ from $F(x)$ is

$$\hat{q}_1(u) = \frac{1}{h(n)} \int_0^1 \frac{K(\frac{t-u}{h(n)})}{f_n(\hat{Q}_1(t))} dt \tag{15}$$

where $f_n(x)$ is a kernel density estimator of $f(x)$ with $h(n)$ as bandwidth.

$\hat{Q}_1(u) = \inf\{x : F_n(x) \geq u\}, 0 < u < 1$ is the empirical estimator of the quantile function $Q(u)$ based on empirical distribution function $F_n(x)$. The kernel $K(.)$ is a density function satisfying regularity conditions (Prakasa Rao (1983))

Estimator proposed by Soni *et al.* (2012) performs better than those given by Jones (1992) in terms of mean square error. Soni *et al.* (2012) proved the following results for fixed u, where $0 < u < 1$:

(R1) $\hat{q}_1(u)$ is a consistent estimator of $q_1(u)$,

(R2) as $n \to \infty$, $\dfrac{\sqrt{n}h(n)(\hat{q}_1(u) - q_1(u))}{\sigma_{1n}(u)}$ is asymptotically normal with mean zero and variance 1, where

$\sigma_{1n}^2(u) = E(\int_0^1 dK_n^*(u,t)F_n(\hat{Q}_1(t)))^2$ with $K_n^*(u,t) = K(\frac{t-u}{h(n)})q_1(t)$.

Let $\hat{q}_2(u)$ denote the corresponding estimator of $q_2(u)$ based on a random sample $Y_1, Y_2, \ldots, Y_n$ from $G(x)$.

3. Test Statistics and Asymptotic Distribution

We propose two test statistics for testing H_0 against H_A. Let $\hat{q}_1(u)$ and $\hat{q}_2(u)$ be consistent estimators of $q_1(u)$ and $q_2(u)$ as discussed in Section 2.2. The difference $\hat{q}_1(u) - \hat{q}_2(u)$ is an empirical measure of departure from the null hypothesis. This difference is expected to be zero under the null hypothesis and non-negative under the alternative hypothesis.

First proposed test statistic T_1 is Kolmogorov-Smirnov type distance between $\hat{q}_1(u)$ and $\hat{q}_2(u)$ and is given as

$$T_1 = \sup_{0<u<1} (\hat{q}_1(u) - \hat{q}_2(u)). \tag{16}$$

Second proposed test statistic T_2 is Cramer-von Mises type difference, as given below

$$T_2 = \int_0^1 (\hat{q}_1(u) - \hat{q}_2(u)) d\Big(\frac{\hat{Q}_1(u) + \hat{Q}_2(u)}{2}\Big). \tag{17}$$

Test based on T_1, T_2 will reject H_0 in favour of H_A for large values of normalised versions of the statistics T_1 and T_2, respectively.

Next we consider a lemma needed to derive the asymptotic distributions of T_1 and T_2 under the null hypothesis.

D and $D[0,1]$ are equipped with the uniform norm $||.||$ and the product norm respectively. In the following lemma, weak convergence of the process $S_n(u)$ is established on D, where

$$S_n(u) = \{\sqrt{n}h(n)(\hat{q}_1(u) - q_1(u)), \sqrt{n}h(n)(\hat{q}_2(u) - q_2(u))\}.$$

Lemma 1: Let $B_1(q_1(u))$ and $B_2(q_1(u))$ be Brownian bridge processes with zero means. Then $S_n(u)$ converges in D to a 2-dimensional Gaussian Process $\{B_1(q_1(u)), B_2(q_2(u))\}$ as $n \to \infty$.

Proof: See the Appendix. □

The above lemma helps us in determining the asymptotic distribution of T_1 as established in Theorem 1 given below.

Theorem 1: Under H_0, as $n \to \infty$, $\sqrt{n}h(n)T_1$ converges in distribution to $\sup_{0<u<1} (B_1(q_1(u)) - B_2(q_2(u)))$.

Proof: See the Appendix. □

Remark 1: A slight modification can be made to the test statistic T_1 as discussed below.

Suppose under H_0, $B_1(q_1(u)) - B_2(q_2(u)) = g(u)$, where $g(u)$ is difference of two Brownian process. Then variance of random variable $g(u)$ is given by

$$Var(g(u)) = Var(B_1(q_1(u))) + Var(B_2(q_2(u))) = \sigma_g^2(u) \cdots (\text{say}).$$

If $\{W(t) : t \geq 0\}$ is a standard Brownian motion (Wiener process), then

$$g(u) \to W(\sigma_g^2(u)).$$

Under H_0, this gives for $0 < u < 1$ and $n \to \infty$,

$$\sqrt{n}h(n)(\hat{q}_1(u) - \hat{q}_2(u)) \to W(\sigma_g^2(u))$$

$$\Rightarrow \frac{\sqrt{n}h(n)(\hat{q}_1(u) - \hat{q}_2(u))}{\sigma_g(T)} \to W(\frac{\sigma_g^2(u)}{\sigma_g^2(T)})$$

where $\sigma_g^2(T) = \max_t[\sigma_g^2(t)]$ for $T \in (0,1)$. Note that $\frac{\sigma_g^2(u)}{\sigma_g^2(T)} \in (0,1)$.

Let $\hat{\sigma}_g(T)$ be a consistent estimator of $\sigma_g(T)$ and define

$$T_{1n}^* = \frac{\sqrt{n}h(n)(\sup_{0<u<1}(\hat{q}_1(u) - \hat{q}_2(u)))}{\hat{\sigma}_g(T)} = \frac{\sqrt{n}h(n)T_1}{\hat{\sigma}_g(T)}. \tag{18}$$

Theorem 2: Under H_0 ,

$$\lim_{n\to\infty} P[T_{1n}^* > b] = P[\sup_{0<u<1} W(u) > b] = 2(1 - \Phi(b)) \tag{19}$$

where $\Phi(b)$ is the cdf of Standard Normal distribution at b.

Proof: The proof follows from Section 7.4 of Durrett (2019). □

In the next theorem, we find the null asymptotic distribution of

$$T_{2n}^* = \sqrt{n}h(n)\int_0^1 (\hat{q}_1(u) - \hat{q}_2(u))d(\frac{\hat{Q}_1(u) + \hat{Q}_2(u)}{2}) = \sqrt{n}h(n)T_2. \tag{20}$$

Theorem 3: Under H_0, T_{2n}^* converges in distribution to a normal random variable with mean zero and variance σ^2 as $n \to \infty$ where

$$\sigma^2 = Var(\int (B_1(q_1(u)) - B_2(q_2(u)))d[\frac{Q_1(u) + Q_2(u)}{2}]).$$

Proof: The proof follows using Hadamard differentiability and functional delta method (Ref. van der Vaart and Wellner (1996); Theorem 3.9.4). For details, see the Appendix. □

In the sequel, T_{1n}^* will be referred to as the supremum statistic and T_{2n}^* as the integral statistic.

4. Simulations

A simulation study has been carried out to verify the asymptotic distribution of test statistics under H_0 and to compute size and power of the standardized versions of statistics T_{1n}^* and T_{2n}^*. The data are generated from GLD, Davies and exponential distributions with sample size $n = 25, 50, 100$. For censored data, censoring distribution is chosen so as to ensure 20% censoring. The chosen bandwidths are $0.15, 0.19, 0.25$ (Soni *et al.* (2012)) for GLD and exponential distributions and 0.85 for Davies distribution. Variances of T_{1n}^* and T_{2n}^*, are estimated by taking 5000 bootstrap samples from the underlying distribution and then T_{1n}^* and T_{2n}^* are calculated for each sample. The kernels used for estimation of quantile density functions are

(i) Triangular: $K(u) = (1 - |u|)I(|u| \leq 1)$ and

(ii) Epanechnikov: $K(u) = .75(1-u^2)I(|u| \leq 1)$ (the optimal kernel (Prakasa Rao (1983))).

4.1. Asymptotic distribution

Simulations are used to verify asymptotic distribution of the proposed statistics under H_0. Test Statistics have been calculated by considering GLD(1,1,2,1) distribution. Both graphical and testing procedures have been employed to test the normality. Figures 1 and 2 show Q-Q plots of standardized versions of T^*_{1n} and T^*_{2n} for $n = 25$ and these plots indicate normality of the statistics.

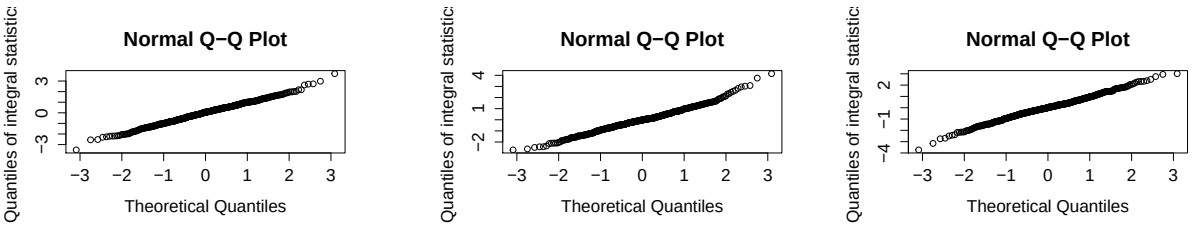


Figure 1: Q-Q plot of Integral statistic for $n = 25$, $h(n) = 0.15, 0.19, 0.25$

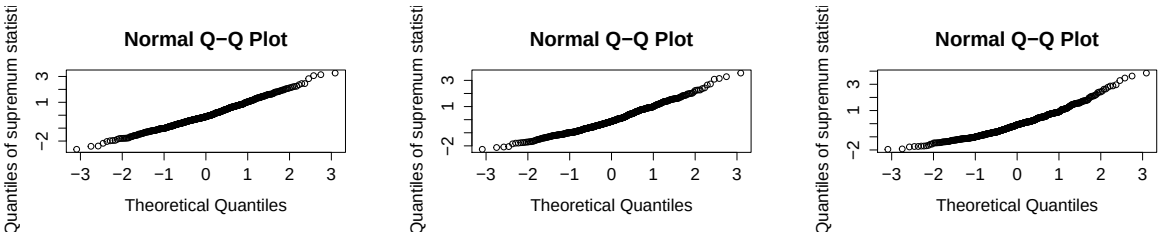


Figure 2: Q-Q plot of Supremum statistic for $n = 25$, $h(n) = 0.15, 0.19, 0.25$

Kolmogorov-Smirnov goodness of fit statistic is used to test the hypothesis that the simulated distributions of two test statistics are asymptotically normal. Table 2 shows p - values of Kolmogorov-Smirnov test statistic for $n = 25$ and bandwidth $h(n) = 0.15, 0.19, 0.25$.

Table 2: p - values of Kolmogorov-Smirnov Test

$h(n)$	Int Statistic	Sup Statistic
0.15	0.257	0.559
0.19	0.612	0.978
0.25	0.934	0.257

Hence from Q-Q plots (Figures 1 and 2) and Kolmogorov-Smirnov goodness of fit test, we conclude that standardized versions of both T^*_{1n} and T^*_{2n} follow Standard Normal distribution for $n \geq 25$. In the following subsection, we compute size and power of supremum and integral statistics when observations are uncensored and censored.

4.2. Calculation of estimates of size and power of tests

Power of the tests based on supremum and integral statistics and those given by Kochar (1979) and Cheng (1985) have been computed for GLD distribution with parameters in the regions listed in Table 1.

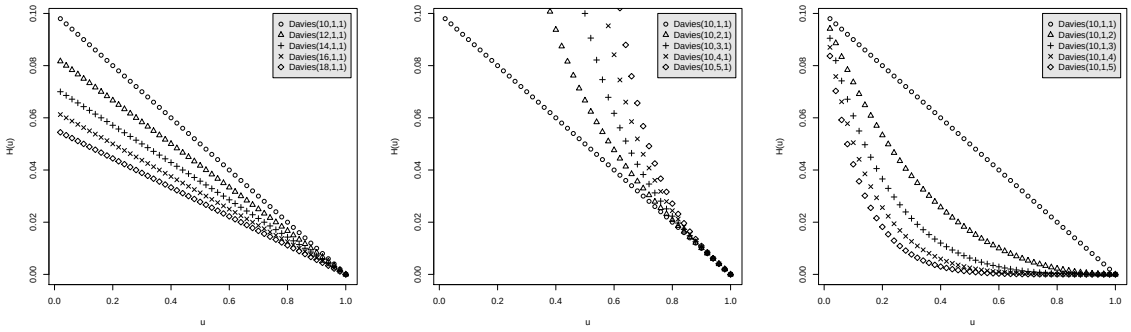
Table 3 depicts size of all tests for a sample of size 25, for uncensored data. For calculating size of the tests, considered distribution is GLD (1,1,2,1).

Table 3: Size of all tests

		$h(n)$		
n	Statistics	0.15	0.19	0.25
25	Sup	0.030	0.050	0.049
	Int	0.048	0.050	0.051
	Kochar	0.044	0.044	0.044
	Cheng	0.050	0.050	0.050

For the calculation of power of the tests, we first consider Davies distribution with quantile, quantile density and hazard quantile function as mentioned in (9), (10), and (11) respectively. The selection of parameters, which will lead to the ordering of hazard quantile functions is explained below through Figure 3(a)-3(c).

Figure 3 (a) shows the hazard quantile functions for Davies (10,1,1), Davies (10,2,1), Davies (10,3,1), Davies (10,4,1) and Davies (10,5,1), that is, λ_1 is changing. Figure 3(b) plots the hazard quantile functions for Davies (10,1,1), Davies (10,1,2), Davies (10,1,3), Davies (10,1,4) and Davies (10,1,5), that is, λ_2 varies. Figure 3(c) displays the hazard quantile functions for Davies (10,1,1), Davies (12,1,1), Davies (14,1,1), Davies (16,1,1) and Davies(18,1,1), that is, scale parameter C is varied.



(a) varying C **(b) varying λ_1** **(c) varying λ_2**

Figure 3: Hazard quantile functions for Davies distribution

Figures 3(a)-3(c) lead to the conclusions that

(i) for $C \leq C^*$, that is, when scale parameters are ordered,
$H_D(u, C^*, \lambda_1, \lambda_2) \leq H_D(u, C, \lambda_1, \lambda_2)$;

(ii) if $\lambda_1 \leq \lambda_1^*$, that is, shape parameters are ordered, then
$H_D(u, C, \lambda_1, \lambda_2) \leq H_D(u, C, \lambda_1^*, \lambda_2)$;

(iii) when $\lambda_2 \leq \lambda_2^*$,
$H_D(u, C, \lambda_1, \lambda_2^*) \leq H_D(u, C, \lambda_1, \lambda_2)$.

Tables 4 and 5 give the power of four statistics for comparing the hazard quantile functions of two Davies distributions for $h(n) = 0.85$. In these tables, shape parameter λ_2 of Davies distribution is varied and in Table 5, departure in shape parameter λ_2 is reduced. In body of Tables 4 and 5, the first (second) value corresponds to power when Triangular (Epanechnikov) kernel is used for the estimation of quantile density function.

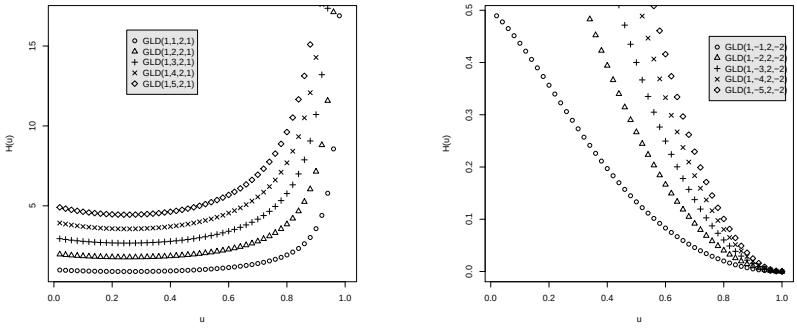
Table 4: Power comparison - Davies(10,1,1) *vs* Davies(10,1,2)(Uncensored case)

		n		
$h(n)$	Statistics	25	50	100
.85	Sup	0.186(0.260)	0.476(0.594)	0.886(0.902)
	Int	0.676(0.636)	0.838(0.848)	0.980(0.978)
	Kochar	0.508	0.786	0.972
	Cheng	0.566	0.754	0.938

Table 5: Power comparison - Davies(10,1,1) *vs* Davies(10,1,1.5)(Uncensored case)

		n		
$h(n)$	Statistics	25	50	100
.85	Sup	0.154(0.157)	0.212(0.238)	0.574(0.596)
	Int	0.286(0.490)	0.466(0.492)	0.646(0.696)
	Kochar	0.276	0.354	0.672
	Cheng	0.364	0.422	0.634

The next distribution of interest is GLD, with quantile, quantile density and hazard quantile function as mentioned in (12), (13), and (14) respectively. Selection of parameters of GLD, required for ordering of hazard quantile functions is explained through Figures 4(a) and 4(b). Figure 4(a) plots the hazard quantile functions of GLD (1,1,2,1), GLD (1,2,2,1), GLD (1,3,2,1), GLD (1,4,2,1) and GLD (1,5,2,1) and Figure 4(b) displays the hazard quantile functions of GLD (1,-1,2,-2), GLD (1,-2,2,-2), GLD (1,-3,2,-2), GLD (1,-4,2,-2) and GLD (1,-5,2,-2). Note that in both the figures, only scale parameter has been changed and all other parameters are same.



(a) Region 1 **(b) Region 2**

Figure 4: Hazard quantile functions for GLD in regions 1 and 2

For $0 < u < 1$, Figures 4(a) and 4(b) depict the following:

(i) In Region 1, for $\lambda_2 \leq \lambda_2^*$, it is observed that

$$H_{GL}(u, \lambda_1, \lambda_2, \lambda_3, \lambda_4) \leq H_{GL}(u, \lambda_1, \lambda_2^*, \lambda_3, \lambda_4);$$

(ii) In Region 2, for $\lambda_2 \leq \lambda_2^*$, it is seen that

$$H_{GL}(u, \lambda_1, \lambda_2^*, \lambda_3, \lambda_4) \leq H_{GL}(u, \lambda_1, \lambda_2, \lambda_3, \lambda_4).$$

Tables 6 and 7 (given in Appendix A.2) give the power for GLD in the Region 1 (Table 1) for Triangular and Epanechnikov kernels in censored as well uncensored case.

We consider GLD(1,1,2,1), GLD(1,2,2,1) in Table 6 and GLD(1,1,2,1), GLD(1,1.2,2,1) in Table 7 wherein departure in scale parameter λ_2 is reduced. The censoring variables have been generated from uniform distribution such that percentage of censoring in both cases is 20 and values in bold font are for censored case.

Table 8 (Appendix A.2) gives the power of our proposed test statistics for testing the equality of hazard quantile functions of EXP(1) and EXP(2) in censored as well as uncensored case. The censoring variables are distributed as EXP(.25) and EXP(0.5) respectively which ensure 20 percentage of censoring. In Tables 6-8, values in parentheses correspond to Epanechnikov kernel.

On the basis of values in Tables 3-5 and 6-8 (given in Appendix A.2), it can be concluded that

(i) for all test statistics and $n \geq 25$, size of tests ≤ 0.05 (level of significance);

(ii) power is not affected by choice of kernel considered;

(iii) power increases with an increase in sample size in uncensored as well as censored cases;

(iv) when observations are from GLD, both the proposed test statistics give higher power than Cheng's and Kochar's test statistics;

(v) when the parameters of GLD are in Region 1 (Table 1), integral statistic has more power than supremum statistic in uncensored case. Integral statistic is able to discriminate small departure in scale parameter λ_2;

(vi) when observations are from GLD, integral statistic is performing better than the supremum statistic in censored case;

(vii) when observations follow Davies distribution, the integral statistic has more power than all other test statistics;

(viii) when observations follow exponential distribution, supremum statistic performs better than integral statistic in censored case;

(ix) Cheng's and Kochar's statistics have more power than newly proposed test statistics when the underlying distribution is exponential.

5. Real Data

Data set of 101 patients with advanced acute myelogenous leukemia reported to International Bone Marrow Transplant Registry is considered (Source: Klein and Moeschberger (1997)). Fifty one of these patients had received an autologous bone marrow transplant in which high doses of chemotherapy and their own bone marrow were reinfused to replace their destroyed immune system. Fifty patients had an allogoneic bone marrow transplant where marrow from an HLA (Histocompatibility Leukocyte Antigen) matched sibling was used to replenish their immune systems. An important issue in bone marrow transplantation is the comparison of hazard quantile functions for these two methods. We compare hazard quantile functions of two techniques through their quantile density functions. Since test statistics proposed by us are for equal sample sizes, we randomly remove one observation from first sample.

Plots of quantile density functions are given in Figure 5. Solid line indicates estimate of quantile density function for auto transplant data and dotted one shows an estimate of quantile density function for allo transplant data. This figure shows that two quantile density functions are ordered. For supremum and integral statistics, p-values are 0.01 and 0.03 respectively. This leads to rejection of null hypothesis at 5 percent level of significance. Hence, it can be concluded that auto transplant technique is more effective than allogenic transplant technique.

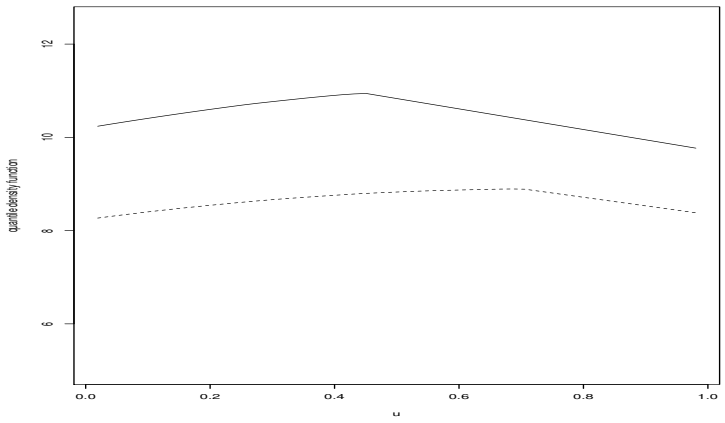


Figure 5: Graph of quantile density functions of autologous(auto) transplant and allogenic(allo) transplant

6. Conclusion

In this paper, we propose two tests based on consistent estimators of quantile density functions for testing equality of two hazard functions or equivalently, the hazard quantile functions, against the alternative that they are ordered. The tests have limiting normal distributions. Numerical studies show that all the tests attain their size. The supremum and the integral tests have better power than the tests proposed by Kochar and Cheng for some alternatives. However, it should be noted that tests by Kochar, Cheng and others can not be used when the observations are censored. But both the tests proposed in this paper can be used for censored data as well. Our tests perform well for families of distributions when closed form of distribution function is not available but explicit form of the quantile function is known.

Acknowledgement

Authors thank the referee for valuable comments which have led to improved version of the paper.

References

Cheng, K. F. (1985). Tests for the equality of failure rates. *Biometrika*, **72(1)** 211-215.

Cheng, C. and Parzen, E. (1997). Unified estimators of smooth quantile and quantile density functions. *Journal of Statistical Planning and Inference*, **59(2)**, 291-307.

Chikkagoudar, M. S. and Shuster, J. S. (1974). Comparison of failure rates using rank tests. *Journal of the American Statistical Association*, **69**, 411-413.

Chesneau, C., Dewan, I. and Doosti, H. (2016). Nonparametric estimation of a quantile density function by wavelet methods. *Computational Statistics and Data Analysis*, **94**, 161-174.

Csorgo, M. and Revesz, P. (1981). Two approaches to constructing simultaneous confidence bounds for quantiles, Carleton Mathematical Series, 176, Carleton University, Ottawa.

Durrett, R. (2019). *Probability: Theory and Examples (Vol. 49)*. Cambridge university press New York (ISBN: 978-0-521-76539-8).

Falk, M. (1986). On the estimation of quantile density function. *Statistics and Probability Letters*, **4(2)**, 69-73.

Fan, C. Ding, G. and Zhang, F. (2020). A kernel nonparametric quantile estimator for right-censored competing risks data. *Journal of Applied Statistics*, **47(1)**, 61-75.

Gilchrist, W. (2000). *Statistical Modelling with Quantile Functions*. Chapman and Hall, New York (ISBN: 978- 0- 4291-1920-0).

Hankin, R. K. S. and Lee, A. (2006). A new family of non-negative distributions. *Australian and New Zealand Journal of Statistics*, **48(1)**, 67-78.

Jones, M. C. (1992). Estimating densities, quantiles, quantile densities and density quantiles. *Annals of the Institute of Statistical Mathematics*, **44(4)**, 721-727.

Karian, A. and Dudewicz, E. J. (2000). *Fitting Statistical Distributions*. CRC, London (ISBN: 978- 1- 5848-8711-9).

Klien, J. P. and Moeschberger, M. L. (1997). *Survival Analysis Techniques for Censored and Truncated data*. Springer, New York (ISBN: 978-0-387-21645-4).

Kochar, S. C. (1979). Distribution-free comparison of two probability distributions with reference to their hazard rates. *Biometrika*, **66(3)**, 437-441.

Kochar, S. C. (1981). A new distribution-free test for the equality of two failure rates. *Biometrika*, **68(2)**, 423-426.

Kumar, V. B. and Nair, N. U. (2011). Ageing concepts: An approach based on quantile function. *Statistics and Probability Letters*, **81(12)**, 2016-2025.

Maladan, D. K. and Sankaran, P. G. (2020). A new family of quantile functions and its applications. *Communications in Statistics-Theory and Methods*. https://doi.org/10.1080/03610926.2020.1713368.

Nair, N. U. and Sankaran, P. G. (2009a). Quantile based reliability analysis. *Communications in Statistics - Theory and Methods* , **38(2)**, 222-232.

Nair, N. U. and Sankaran, P. G. (2009b). Nonparametric estimation of hazard quantile function. *Journal of Nonparametric Statistics*, **21(6)**, 757-767.

Parzen, E. (1979). Nonparametric statistical data modelling. *Journal of the American Statistical Association*, **74**, 105-122.

Peng, L. and Fine, J. P. (2007). Nonparametric quantile inference with competing-risks data. *Biometrika*, **94(3)**, 735-744.

Prakasa Rao, B. L. S. (1983). *Nonparametric Functional Estimation*. Academic Press, New York (ISBN: 978-1-4832-6923-8).

Ramberg, J. S. and Schmeiser, B. W. (1974). An approximate method for generating asymmetric random variables. *Communications of the ACM*, **17(2)**, 78-82.

Soni, P., Dewan, I. and Jain, K. (2012). Nonparametric estimation of quantile density function. *Computational Statistics and Data Analysis*, **56(12)**, 3876-3886.

Van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes With Applications to Statistics*. Springer, New York (ISBN: 978-0-3879-4640-5).

Appendix

A.1: Proofs

Proof:[**Lemma 1**] For arbitrary real numbers λ_1 and λ_2, we consider $T_n(u) = \sqrt{n}h(n)(\lambda_1(\hat{q}_1(u) - q_1(u)) + \lambda_2(\hat{q}_2(u) - q_2(u)))$.

On using central limit theorem, $T_n(u)$ converges in distribution to $N(0, \sigma_n^2(u))$ as $n \to \infty$ where $\sigma_n^2(u) = \lambda_1^2\sigma_{1n}^2(u) + \lambda_2^2\sigma_{2n}^2(u)$.

Using Cramer Wold device, as $n \to \infty$, we get

$S_n(u) = \{\sqrt{n}h(n)(\hat{q}_1(u) - q_1(u), \hat{q}_2(u) - q_2(u))\} \to$ Gaussian process $N(0, \Sigma_n)$ where

$$\Sigma_n = \begin{bmatrix} \sigma_{1n}^2(u) & 0 \\ 0 & \sigma_{2n}^2(u) \end{bmatrix}.$$

For a finite set of numbers $u_1, ..., u_n$ and arbitrary $\lambda_{1i}, \lambda_{2i}$, $\sum_{i=1}^n(\lambda_{1i}\hat{q}_1(u_i) + \lambda_{2i}\hat{q}_2(u_i))$ is sum of independent random variables. Using central limit theorem for the univariate independent random variables and Cramer Wold device, we conclude that the finite dimensional distribution of process $\{S_n(u)\}$ converges weakly to that of a 2-dimensional Gaussian process.

It is well known that the sequences $\sqrt{n}h(n)(\hat{q}_1(u) - q_1(u))$ and $\sqrt{n}h(n)(\hat{q}_2(u) - q_2(u))$ converge weakly in $(D[0,1],.)$ to $B(q_1(u))$ and $B(q_2(u))$ respectively, where $B(q_1(u))$ and $B(q_2(u))$ are Brownian Bridge processes with zero means. Thus, two sequences $\sqrt{n}h(n)(\hat{q}_1(u)-q_1(u))$ and $\sqrt{n}h(n)(\hat{q}_2(u)-q_2(u))$ are asymptotically tight which implies that the process $\{S_n(u)\}$ is also asymptotically tight using (Lemma 1.4.3 and Theorem 1.5.4, van der Vaart and Wellner (1996)).

Distribution of $S_n(u)$ is established using Theorem 1.5.4 of van der Vaart and Wellner (1996). Hence, we conclude that the finite dimensional distribution of the process $\{S_n(u)\}$ converges weakly to that of a 2-dimensional Gaussian process $\{B_1(q_1(u)), B_2(q_2(u))\}$. □

Proof: [**Theorem 1**] From Lemma 1, we have

$$\sqrt{n}h(n)\{(\hat{q}_1(u) - q_1(u)), (\hat{q}_2(u) - q_2(u))\} \xrightarrow{\mathbf{L}} \{B_1(q_1(u)), B_2(q_2(u))\}$$

where B_i are Brownian bridge processes with zero means. Using continuous mapping theorem,

$\sup_{0<u<1} \sqrt{n}h(n)(\hat{q}_1(u) - \hat{q}_2(u))$ converges to $\sup_{0<u<1} (B_1(q_1(u)) - B_2(q_2(u)))$ as $n \to \infty$. □

Proof: [**Theorem 3**] The proof follows using Hadamard differentiability and functional delta method (Theorem 3.9.4, Van der Vaart and Wellner (1996)). Let $BV_1[0,1]$ denote the set of cadlag functions of total variation bounded by M (finite). The map

$\phi(A, B) = \int_0^1 AdB$ from $D[0,1] \times BV_1[0,1]$ to the real line is Hadamard differentiable (using Lemma 3.9.17 of van der Vaart and Wellner (1996)). The Hadamard derivative of

$\phi(A,B)$ is

$$\phi_{(A,B)}(\alpha,\beta)=\int_0^1 Ad\beta+\int_0^1 \alpha dB \tag{21}$$

where $\int Ad\beta$ is defined via integration by parts if β is not of bounded variation.

Let $A=q_1(u)-q_2(u)$, $B=\frac{Q_1(u)+Q_2(u)}{2}$, $\alpha=B_1(q_1(u))-B_2(q_2(u))$ and $\beta=B_3(\frac{Q_1(u)+Q_2(u)}{2})$ where B_3 is a Brownian bridge process with mean zero.

Using Lemma 1 and delta method, we get for large n and under H_0

$$\begin{aligned}\sqrt{n}h(n)T_2 &\to \phi_{(q_1(u)-q_2(u),\frac{\hat{Q}_1(u)+\hat{Q}_2(u)}{2})}(B_1(q_1(u))-B_2(q_2(u)),B_3(\frac{Q_1(u)+Q_2(u)}{2})) \\ &=\int (B_1(q_1(u))-B_2(q_2(u)))d\Big(\frac{(Q_1(u)+Q_2(u)}{2}\Big),\end{aligned} \tag{22}$$

since the first term in (22) is zero under H_0 for large n.

Hence, the limiting random variable is normally distributed with mean zero and variance

$$\sigma^2=Var(\int (B_1(q_1(u))-B_2(q_2(u)))d(\frac{(Q_1(u)+Q_2(u)}{2})). \tag{23}$$

□

A.2: Tables

Table 6: Power comparison for GLD(1,1,2,1) *vs* GLD(1,2,2,1)

		h(n)		
n	Statistics	0.15	0.19	0.25
25	Sup uncensored	0.390(0.636)	0.288(0.614)	0.310(0.824)
	Sup censored	**0.15(0.168)**	**0.18(0.153)**	**0.250(0.266)**
	Int uncensored	0.984(1.000)	0.966(1.000)	0.978(1.000)
	Int censored	**0.262(0.247)**	**0.347(0.365)**	**0.457(0.428)**
	Kochar	0.356	0.356	0.356
	Cheng	0.146	0.146	0.146
50	Sup uncensored	0.422(0.806)	0.712(0.948)	0.836(0.948)
	Sup censored	**0.305(0.585)**	**0.389(0.444)**	**0.491(0.584)**
	Int uncensored	1.000(1.000)	1.000(1.000)	1.000(1.000)
	Int censored	**0.565(0.283)**	**0.767(0.793)**	**0.862(0.923)**
	Kochar	0.524	0.524	0.524
	Cheng	0.146	0.146	0.146
100	Sup uncensored	1.000(1.000)	1.000(1.000)	1.000(1.000)
	Sup censored	**0.496(0.638)**	**0.773(0.82)**	**0.951(0.963)**
	Int uncensored	1.000(1.000)	1.000(1.000)	1.000(1.000)
	Int censored	**0.972(0.981)**	**0.992(0.987)**	**0.997(0.998)**
	Kochar	0.798	0.798	0.798
	Cheng	0.160	0.160	0.160

Table 7: Power comparison for GLD(1,1,2,1) *vs* GLD(1,1.2,2,1)

		h(n)		
n	Statistics	0.15	0.19	0.25
25	Sup uncensored	0.086(0.102)	0.076 (0.080)	0.122 (0.084)
	Sup censored	**0.061(0.078)**	**0.095(0.100)**	**0.086(0.100)**
	Int uncensored	0.196(0.182)	0.260(0.214)	0.304(0.308)
	Int censored	**0.111(0.133)**	**0.127(0.194)**	**0.170(0.193)**
	Kochar	0.102	0.102	0.102
	Cheng	0.109	0.109	0.109
50	Sup uncensored	0.086(0.104)	0.130(0.16)	0.182(0.142)
	Sup censored	**0.122(0.122)**	**0.142(0.165)**	**0.157(0.205)**
	Int uncensored	0.636(0.58)	0.688(0.588)	0.804(0.804)
	Int censored	**0.266(0.343)**	**0.361(0.401)**	**0.539(0.548)**
	Kochar	0.200	0.200	0.200
	Cheng	0.110	0.110	0.110
100	Sup uncensored	0.146(0.118)	0.188(0.222)	0.322(0.358)
	Sup censored	**0.232(0.241)**	**0.283(0.316)**	**0.369(0.405)**
	Int uncensored	0.974(0.968)	0.982(0.986)	0.992(0.996)
	Int censored	**0.712(0.756)**	**0.805(0.819)**	**0.915(0.941)**
	Kochar	0.301	0.301	0.301
	Cheng	0.119	0.119	0.119

Table 8: Power comparison for EXP(1) *vs* EXP(2)

		h(n)		
n	Statistics	0.15	0.19	0.25
25	Sup uncensored	0.270(0.310)	0.400(0.230)	0.350(0.290)
	Sup censored	**0.126(0.106)**	**0.106(0.170)**	**0.186(0.242)**
	Int uncensored	0.570(0.510)	0.620(0.600)	0.600(0.650)
	Int censored	**0.086(0.118)**	**0.122(0.198)**	**0.108(0.144)**
	Kochar	0.694	0.694	0.694
	Cheng	0.740	0.740	0.740
50	Sup uncensored	0.540(0.360)	0.590(0.360)	0.550(0.570)
	Sup censored	**0.240(0.244)**	**0.238(0.192)**	**0.386(0.356)**
	Int uncensored	0.700(0.690)	0.700(0.730)	0.780(0.810)
	Int censored	**0.216(0.154)**	**0.246(0.240)**	**0.254 (0.222)**
	Kochar	0.926	0.926	0.926
	Cheng	0.939	0.939	0.939
100	Sup uncensored	0.380(0.670)	0.490(0.320)	0.620(0.710)
	Sup censored	**0.370(0.250)**	**0.476(0.366)**	**0.538(0.386)**
	Int uncensored	0.810(0.780)	0.900(0.830)	0.830(0.910)
	Int censored	**0.304(0.296)**	**0.172(0.240)**	**0.284(0.338)**
	Kochar	0.998	0.998	0.998
	Cheng	0.999	0.999	0.999

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 95-106

A Cost Effective Approach to the Design and Analysis of Multi–group Experiments

Satya Prakash Singh[1], **Shyamal D. Peddada**[2], **Ori Davidov**[3]
[1]*Department of Mathematics, Indian Institute of Technology Hyderabad, Telangana 502285, India*
[2]*Department of Biostatistics, University of Pittsburgh School of Public Health, 130 DeSoto Street, Pittsburgh, PA 15261, USA*
[3]*Department of Statistics, University of Haifa, Mount Carmel, Haifa 3498838, Israel*

Received: 15 August 2020; Revised: 29 October 2020; Accepted: 01 November 2020

Abstract

In this communication it is shown that employing statistical methods which account for constraints, inherent in some scientific problems, will often lead to a substantial reduction in the sample size required while simultaneously maintaining the power of the study and its scientific validity. In fact a 40%, or even higher, reduction in the required sample size is possible. These savings have the potential to impact individual labs and researchers and will translate to saving of millions of dollars annually for granting authorities and federal agencies such as the NIH.

Key words: Maxi-min designs; Order restricted inference; Power; Sample size.

AMS Subject Classifications: 62F30, 62K05

1. Introduction

Scientific research often requires testing of hypotheses comparing two or more experimental groups. The successful conduct of such investigations requires a study design appropriate for the scientific question at hand, a valid testing procedure for the hypothesis of interest, and an adequate sample size which guarantees suitable power. Sample size determination, or equivalently power calculations, are usually based on two sample and two–sided alternative hypotheses designed to test whether the mean response of the treatment group is different from that of the control group, *cf.*, Ryan (2013). Such calculations are simple and very widely used and numerous software packages, such as SAS and SPSS, have built–in routines for such tasks.

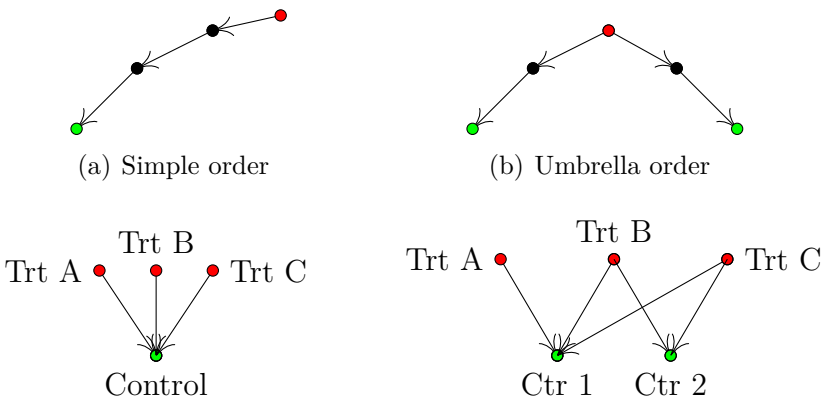
In many applications, such as dose–response studies or multi–drug trials, researchers may have a priori beliefs about the experimental groups. Such prior beliefs are usually based on earlier studies or an understanding of the underlying scientific phenomenon and are often formulated as mathematical inequalities or constraints, known as *order restrictions*.

Correponding Author: Satya Prakash Singh
Email: spsingh@math.iith.ac.in

For example, in a dose–response studies toxicologists may expect that the mean response increases (or decreases) with the dose of a chemical. This constraint is known as the *simple order*. Observational data are also often of this form. For example, in Spiegelhalter *et al.* (1999) the length of the ramus bone of 20 boys was measured at three equally spaced time points from ages 8 to 9. The question of interest was to know whether there was a significant growth spurt during the observed time period. In a time–course gene expression study, the mean expression of a gene may increase up to a certain point, reflecting its biological activity [Peddada *et al.* (2003)] and then decrease. This constraint is known as the *umbrella order*. In clinical trials, a researcher may be interested in demonstrating that the standard treatment is inferior to one of the new treatments, or, that a new treatment is at least as efficacious as the existing ones. This constraint is called the *tree order*. For example, Igari *et al.* (2014) compared the effect of various doses of cytisine on a dysporic–like state in rats. In some cases, the study design may include multiple control and multiple treatment groups. For example, the US National Toxicology Program (NTP) evaluates toxicity and carcinogenicity of chemicals using the concurrent control group as well as historical controls (which are controls collected from similar studies conducted by the NTP). This set up leads naturally to a *bipartite order* restriction [Kanno *et al.* (2003) and Peddada *et al.* (2007)].

The above mentioned order relations are represented graphically in Figure 1 by their corresponding *order graphs*. In each of the Figures, a circle represents a group mean, or more generally any other statistical parameter, and a pointed arrows implies an inequality among the two means or parameters. The *roots* of the order graph are the nodes with the largest means, whereas the *leaves* are the nodes with the smallest means. A variety of other constraints, or order restrictions, arise in applications. There exists over six decades of literature on this subject starting with the pioneering papers of Ayer *et al.* (1955), van Eden (1956) and Bartholomew (1959). Several books summarizing the work done in this field have also been published, e.g., Barlow *et al.* (1972), Robertson *et al.* (1988) and Silvapulle and Sen (2005).

In this article we highlight some important consequences of incorporating order restrictions in both the design and the analysis of experiments. Doing so addresses the scientific questions motivating the study in a principled manner. For if, for example, a standard two–sided test is applied in Figure 1(c), then a significant result tells us that there are differences among the treatments, it does not tell us that one of the treatments is superior to the control. Such inferences, however, are built–in into the procedures of constrained inference. Thus incorporating constraints in the analysis provides more meaningful inferences about the existence of an ordering among the experimental groups. In addition, using the constraints substantially improves efficiency. This means that we can expect considerable improvement in power and therefore the required sample sizes are reduced. In other words, failing to properly incorporate the order restrictions may lead to inflated costs of conducting studies, loss of power and inadequate scientific conclusions.



(a) Simple order (b) Umbrella order

(c) Tree order (d) Bipartite order

Figure 1: Order graphs for some common order restrictions. Circles represent group means and a pointed arrow indicates an inequality among the means. Green circles correspond to leaves of the order graph and red circles to their roots. We refer to the leaves and roots as the extreme groups. The intermediate groups are designated by a black circle.

2. Power, Order and Scientific Discovery

It is well known that tests tailored to accommodate order restrictions, called restricted or constrained tests [Silvapulle and Sen (2005)], are typically more powerful than their unconstrained counterparts. For example, consider the one–way analysis of variance (ANOVA) model

$$Y_{ij} = \mu_i + \epsilon_{ij},$$

where Y_{ij} is the response of j^{th} observation in i^{th} treatment group, $i = 1, \ldots, K$ and $j = 1, \ldots, n_i$ and the errors ϵ_{ij} are independent $\mathcal{N}(0, \sigma^2)$ random variables (RVs). For simplicity, and without any loss of generality, see Remark 2.2 in Singh and Davidov (2020), one may assume that $\sigma^2 = 1$ in which case the unconstrained likelihood ratio test (LRT) is of the form

$$T_n = \sum_{i=1}^{K} n_i (\bar{Y}_i - \hat{\mu}_i)^2$$

where $\bar{Y}_i = n_i^{-1} \sum_{j=1}^{n_i} Y_{ij}$ for $i = 1, \ldots, K$ and $\hat{\mu}_i = \bar{Y} = N^{-1} \sum_{i=1}^{K} n_i \bar{Y}_i$ are the unrestricted estimators. Similarly the constrained LRT is given by

$$T_n = \sum_{i=1}^{K} n_i (\tilde{\mu}_i - \hat{\mu}_i)^2$$

where $\tilde{\mu}_i$ is the i^{th} component of $\tilde{\boldsymbol{\mu}} = \text{argmax}\{\sum_{i=1}^{K} n_i (\bar{Y}_i - \mu_i)^2 : \mathbf{R}\boldsymbol{\mu} \geq \mathbf{0}\}$, the restricted maximum likelihood estimator of $\boldsymbol{\mu}$ which is assumed to satisfy a collection of linear inequalities $\mathbf{R}\boldsymbol{\mu} \geq \mathbf{0}$. It is well known that under the null the unconstrained LRT follows a chi–square distribution whereas the restricted LRT follows, what is known as, a chi–bar–square distribution [Silvapulle and Sen (2005)].

Figure 2 plots the power function of the standard (unconstrained) ANOVA test versus its constrained counterpart as a function of the per–group sample size under a balanced design. Clearly, the constrained test has higher power. Consequently the sample size required to guarantee a prespecified power is smaller when using a constrained test. At the 5% significance level and 80% power the unconstrained test requires 136 observations whereas the constrained test requires only 88 observations. It is evident that the reduction in sample sizes is a substantial 35%.

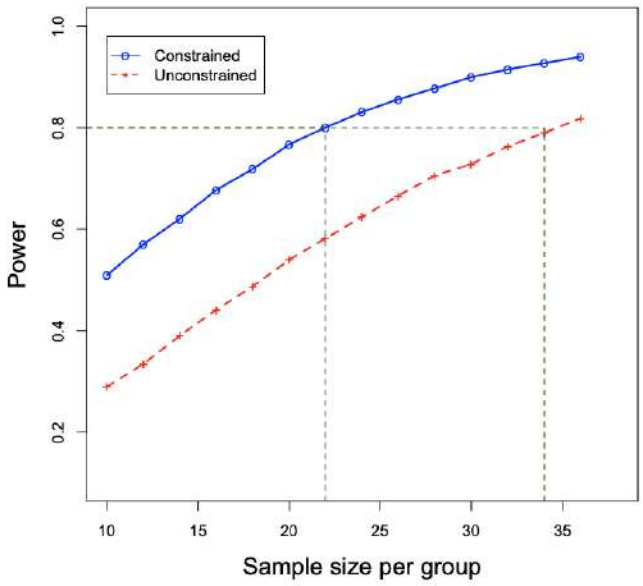


Figure 2: The power of the constrained and unconstrained tests in the ANOVA setting. Data were simulated from normal populations with means 0, 0.25, 0.5 and 0.75, and unit standard deviation

Even more dramatic examples are reported in the literature both in the context of ANOVA [Farnan *et al.* (2014)] as well as a variety of other settings [*e.g.*, Davidov and Herman (2012) and Rosen and Davidov (2017)]. A theoretical proof of the superiority of the restricted LRT is provided by Praestgaard (2012) and Davidov and Iliopoulos (2020). In the following we provide two examples from our own research which demonstrate that using methods which incorporate constraints helps to uncover clinically important features in the data which were missed by standard methods.

Example 1: Uterine fibroids, also known as uterine leiomyomata, are benign smooth muscle hormonally mediated tumors commonly found in pre–menopausal women. Nearly 70% of all women have these tumors. They cause pain, bleeding, urinary incontinence and pregnancy complications. The total annual cost of treating these tumors in US is estimated to be between 4 to 9 billion US dollars. The NIH, [cf. Peddada *et al.* (2008)], conducted a large prospective study of 72 pre–menopausal women (38 black and 34 white). Fibroid volumes were measured by MRI taken at baseline and at 3, 6, and 12 months, with at least two measurements per woman. African American women are known to have greater tumor

burden so a standard ANOVA–based analysis with an interaction between race and was performed. The interaction was found to be barely significant at $p = 0.05$. Since these tumors are known to be estrogen dependent, it is reasonable to hypothesize that tumor growth rates would decrease with age. This hypothesis was investigated in a recent re–analysis of these data [Peddada and Jelsema (2016)] using methods which account for order restrictions. A statistically significant decreasing trend in mean growth rates among whites ($p = 0.015$) but not among blacks ($p = 0.1880$) (Figure 3) was formally discovered. Thus, testing for order restrictions allows us to make a clinically important discovery that was not discovered by the standard ANOVA based methodology.

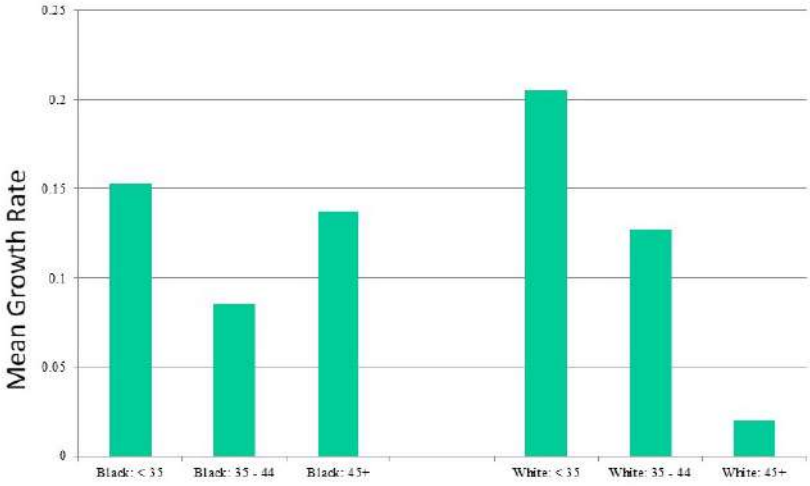


Figure 3: Mammary gland fibroadenoma incidence in female rats

Example 2: The Fish industry uses Malachite Green Chloride as an antifungal agent. The US National Toxicology Program (NTP) conducted a two year cancer bioassay with 48 female rats assigned to each of four dose groups of Malachite Green Chloride, namely, 0, 100, 300 or 600 parts per million. The incidence of mammary gland adenomas and pituitary gland adenoma–carcinomas are reported in Table 1. It is well–known that pituitary gland tumors may be associated with mammary gland tumors via the prolactin pathway [cf. McComb *et al.* (1984) and TR-527 (2005)]. Although these tumors are biologically dependent, the NTP analyzed them separately. The p–values for the corresponding trend tests were not significant, 0.113 for mammary gland adenoma and 0.162 for the pituitary gland adenoma–carcinomas. Davidov and Peddada (2011) developed a nonparametric multivariate ordered test that exploited the underlying dependence among the binary variables to test for trends in multivariate data. Using this constrained trend test Davidov and Peddada (2011) reanalyzed the NTP's Malachite Green Chloride data and discovered a significant increasing trend in both mammary gland adenomas as well as pituitary gland adeno–carcinomas, with a joint p–value of 0.025, suggesting a carcinogenic effect of Malachite Green Chloride on both tumors in a dose–related fashion. This finding reinforce the fact that the methods of constrained inference may discover finding not detected by standard methods.

Another advantage of using the methods of order restricted inference is that it relaxes

Table 1: Tumor incidence rates of control and Malachite Green Chloride treated animals in the NTP study

Tumor type	Estimator	Control	100 ppm	300 ppm	600 ppm
Mammary Gland	Unconstrained	0.050	0.052	0.023	0.130
	Constrained	0.042	0.042	0.042	0.130
Pituitary Gland	Unconstrained	0.607	0.822	0.696	0.756
	Constrained	0.609	0.758	0.758	0.758

parametric assumptions. For example, suppose one is interested in the effect of an allele on a phenotype Y. It is very common to test for "trend" over the alleles aa, Aa, AA by assigning scores $X = 0, 1$ and 2, respectively and performing a linear regression of Y on X. The basic assumption, when using such a modelling framework, is that the change in the mean response from aa to aA is same as from Aa to AA. Such assumptions may not be supported by the data and preclude the possibility of some non–linear but monotonic response such as in Figure 1(a). Such non–parametric curves are easily accommodated by constrained methods. In toxicology, it is also very common to perform linear regression–based tests such as the Cochran–Armitage trend test [Cochran (1954) and Armitage (1955)]. Some investigators use the exact dose as the explanatory variable and others use scores such as 1, 2, 3 and 4. When linearity is not justifiable considerable loss of power is to be expected [Peddada *et al.* (2005a)].

To summarize, incorporating the constraints in the analysis does not only lead to a beautiful and less restrictive statistical theory with improved operating characteristics, it may, much more importantly, help uncover biologically and clinically important results which standard methods fail to detect.

3. Optimal Design: Sample Size and Cost Efficiency

Smucker *et al.* (2018) emphasized that one should *customize the experiment for the setting instead of adjusting the setting to fit a classical design*, a comment that underscores the importance of carefully planned experiments. Recently, Singh and Davidov (2019) developed a rigorous framework for constructing optimal experimental designs which incorporate order restrictions. Their designs, known as Max–Min (MM) designs, maximize power under the worst possible (allowable) configuration in the alternative. They showed that the MM–design is of the form

$$\boldsymbol{\xi}_{\text{MM}} = |\mathcal{V}|^{-1} \sum_{(i,j)\in\mathcal{V}} \boldsymbol{\xi}_{ij}, \tag{1}$$

where $\boldsymbol{\xi}_{ij} = (\boldsymbol{e}_i + \boldsymbol{e}_j)/2$, $\boldsymbol{e}_l$ is the lth standard basis of $\mathbb{R}^K$ and $\mathcal{V}$ is the set of all maximal

pairs. A pair (i, j) where $i \in \mathcal{R}$, the set of roots, and $j \in \mathcal{L}$, the set of leaves, is called a maximal pair if there is a path from i to j. For more details see Singh and Davidov (2019).

The formula (1) is simple and easy to use. MM–designs for some common order restrictions such as the simple, tree, umbrella, and bipartite order (*cf.*, Figure 1) are given in Table 2 along with some other commonly used designs. It turns out that MM–designs allocate observations only to the leaves and roots of the order graph. In fact, if there are N observations, then $N/2$ will be distributed among the leaves and $N/2$ among the roots. When there is more than one root the allocation among the roots is proportional to the degree of the root, *i.e.*, the number of paths to distinct leaves; and similarly for the leaves. Thus, the MM–design for the simple order will allocate $N/2$ observations to the two extreme groups. No observations are allocated to any of the intermediate groups. In the case of the umbrella order, the MM–design assigns $N/2$ observations to the peak of the umbrella and the remaining $N/2$ observations are equally divided among the extreme groups (first and last). Similar logic applies to the tree and bipartite order. We note that MM–designs do not allocate any observations to intermediate treatment groups, and thus do not allow any comparisons among them. This potential practical deficiency can be be addressed and rectified by using Singh and Davidov (2019)'s so called IUT–designs, which, for lack of space, we will not further discuss here.

Table 2: The proportion of the observations allocated by the MM, Balanced, and Dunnetts' design are reported for the order relations depicted in Figure 1. The notation "-" indicates that there is no design to consider

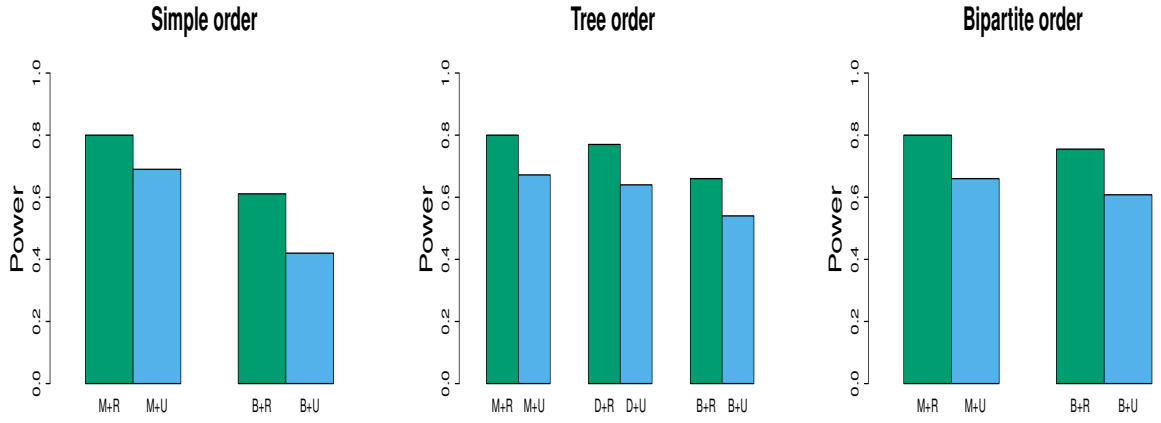
	Order			
Design	Simple	Umbrella	Tree	Bipartite
MM	$(1/2, 0, 0, 1/2)$	$(1/4, 0, 1/2, 0, 1/4)$	$(1/2, 1/8, 1/8, 1/8, 1/8)$	$(3/10, 2/10, 1/10, 2/10, 2/10)$
Balanced	$(1/4, 1/4, 1/4, 1/4)$	$(1/5, 1/5, 1/5, 1/5, 1/5)$	$(1/5, 1/5, 1/5, 1/5, 1/5)$	$(1/5, 1/5, 1/5, 1/5, 1/5)$
Dunnett	-	-	$(1/3, 1/6, 1/6, 1/6, 1/6)$	-

4. Results

The benefits associated with MM–designs were assessed by simulations using data from the published scientific literature. Simulations under the simple order were based on the data of Spiegelhalter *et al.* (1999), whereas the simulations for the tree and bipartite orders were based on data from Igari *et al.* (2014) and Kanno *et al.* (2003), respectively. The substantive scientific problems investigated in these papers were already briefly described. For simplicity, the simulated data is normally distributed with mean values and standard deviations as reported in Table 3. For each ordered alternative, we performed an unconstrained and restricted likelihood ratio test. The results of the simulation study, based on 10^5 simulation runs, are summarized in Figures 4 and 5 which display powers and sample sizes, respectively.

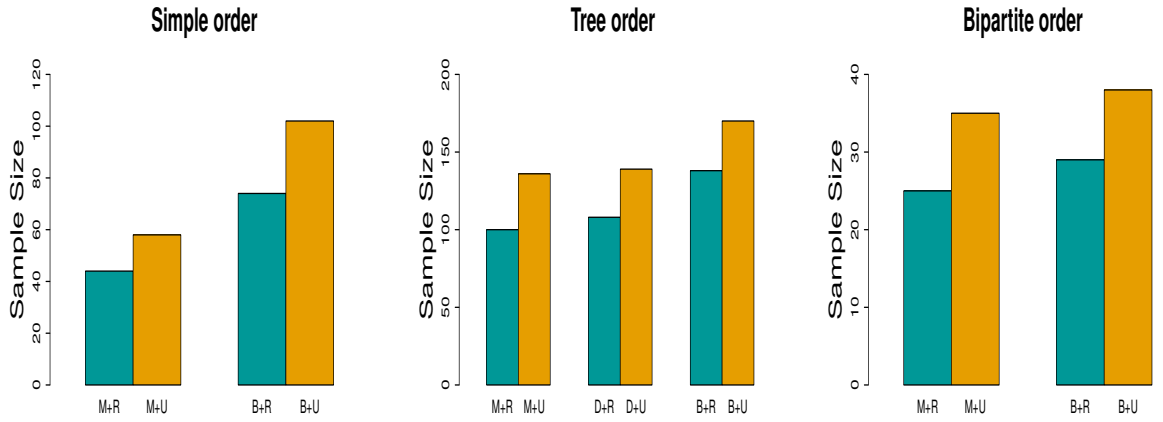
Table 3: A brief summary of the results of Spiegelhalter *et al.* (1999), Igari *et al.* (2014) and Kanno *et al.* (2003). We report on the group size, sample mean and standard deviation as well as the pooled standard deviation (PSD). For the tree order, treatment 1, serves as the control and is compared to the remaining treatments. In the bipartite case, treatments 1 and 2 are the controls. Treatment 1 is compared to treatments 3, 4, and 5, whereas treatment 2 is compared to the 4 and 5

	Treatment Group					
Order	1	2	3	4	5	PSD
Simple	48.66 ± 2.52 20	49.62 ± 2.54 20	50.57 ± 2.63 20			2.56
Tree	97.6 ± 10.39 12	101.6 ± 8.66 12	102.2 ± 4.50 12	103.4 ± 10.04 12	105.9 ± 14.90 12	10.26
Bipartite	29.5 ± 2.95 6	30.0 ± 2.30 6	32.2 ± 3.13 6	34.8 ± 3.48 6	31.8 ± 4.34 6	3.31



(a) (b) (c)

Figure 4: Power comparisons between the Maxi-Min (MM), Balanced (B) and Dunnett's (D) designs when applied with both the unrestricted and restricted test. For example MM+R is the power of the MM design with a restricted test



(a) (b) (c)

Figure 5: Sample sizes required for 80% power under the Maxi-Min (MM), Balanced (B) and Dunnett's (D) designs when applied with both the unrestricted and restricted test. For example MM+R is the sample size required by the MM design with a restricted test

Our simulation study shows, as previously noted, that using the restricted test is always better than using the unrestricted test. It is clear that the MM–design results in improved power relative to the balanced and other designs irrespective of the test being used. For example, Figure 5(b) shows that the MM design analyzed by a restricted test requires a sample of approximately 100 subjects whereas the balanced design with and standard test requires 170 subjects.

5. Summary

This communication shows that accounting for constraints, which occur naturally in a wide variety of scientific investigations, has a huge dividend. In particular it is shown, using examples from the literature, that a substantial reduction in the sample size is achieved when both designing and analyzing data using methods that account for constraints. It is emphasized that the largest benefits are achieved when an experiment is both designed and analyzed using order based methods. The reduction in the required sample sizes, or equivalently the increase in power [Singh and Davidov (2019)], is nothing but phenomenal suggesting that the routine use of order based methods, when appropriate, will result in much more economical and efficient designs. In fact, since in many experimental sciences a substantial portion of the budget is devoted to acquiring a large as possible sample, researchers, pharmaceuticals, granting agencies and others may save millions of dollars on data collection and do much more with a fixed budget. In addition, if the study involves biological samples from animal or human subjects, then these methods would require the participation of fewer animals or human subjects.

It is surprising that although the methodology we describe here traces its roots to the late 1950's it has not had a major impact on data collection and analysis in the sciences.

There are several reasons for that. The first is that the focus of statisticians working in this area had been largely theoretical with little concern for practical issues such as cost reductions. Secondly, appropriate software for analyzing data using these constrained inference based methods were not available until recently. Software such as ORIOGEN [Peddada *et al.* (2005b)] and CLME [Peddada and Jelsema (2016)] have taken the important first steps in this direction and are gaining popularity among users. Finally, the development of experimental designs [Singh and Davidov (2019)] which capitalize on scientific constraints is a recent development with potential far reaching consequences.

Acknowledgements

The research leading to this paper was conducted (mostly) while the first author was a postdoctoral fellow at the University of Haifa, Israel. The work of Satya Prakash Singh was further supported by SEED grant (Indian Institute of Technology, Hyderabad) and is gratefully acknowledged. The work of Ori Davidov is supported by the Israeli Science Foundation Grant No. 456/17 and is gratefully acknowledged.

References

Armitage, P. (1955). Tests for linear trends in proportions and frequencies. *Biometrics*, **11**(**3**), 375–386.

Ayer, M., Brunk, H. D., Ewing, G. M., Reid, W. T. and Silverman, E. (1955). An empirical distribution function for sampling with incomplete information. *Annals of Mathematical Statistics*, **26**(**4**), 641–647.

Barlow, R. E., Bartholomew, D. J., Bremner, J. M. and Brunk, H. D. (1972). *Statistical Inference under Order Restrictions*. Wiley, New York.

Bartholomew, D. J. (1959). A test of homogeneity for ordered alternatives. *Biometrika*, **46**(**1-2**), 36–48.

Cochran, W. G. (1954). Some methods for strengthening the common χ^2 tests. *Biometrics*, **10**(**4**), 417–451.

Davidov, O. and Herman, A. (2012). Ordinal dominance curve based inference for stochastically ordered distributions. *Journal of the Royal Statistical Society, **B***, **74**(**5**), 825–847.

Davidov, O. and Iliopoulos, G. (2020). Semiparametric ordered inference for conditional distributions. *Submitted*.

Davidov, O. and Peddada, S. D. (2011). Order restricted inference for multivariate binary data with application to toxicology. *Journal of the American Statistical Association*, **106**(**496**), 1394–1404.

Farnan, L., Ivanova, A. and Peddada, S. D. (2014). Constrained inference in biological sciences: Linear mixed effects models under constraints. *PLOS ONE*, **9**(**1**), e84778.

Igari, M., Alexander, J. C., Ji, Y., Xiaoli, Q., Papke, R. L. and Bruijnzeel, A. W. (2014). Varenicline and cytisine diminish the dysphoric-like state associated with spontaneous nicotine withdrawal in rats. *Neuropsychopharmacology*, **39**, 445–455.

Kanno, J., Onyon, L., Peddada, S., Ashby, J., Jacob, E. and Owens, W. (2003). The oecd program to validate the rat uterotrophic bioassay. phase 2: Dose-response studies. *Environmental Health Perspectives*, **111**(**12**), 1530–1549.

McComb, D. J., Kovacs, K., Beri, J. and Zak, F. (1984). Pituitary adenomas in old sprague-dawley rats: A histologic, ultrastructural, and immunocytochemical study. *Journal of the National Cancer Institute*, **73**(**5**), 1143–1166.

Peddada, S., Lobenhofer, E., Li, L., Afshari, C. and Umbach, D. (2003). Gene selection and clustering for time-course and dose response microarray experiments using order-restricted inference. *Bioinformatics*, **19**(**7**), 834–841.

Peddada, S. D., Dinse, G. and Haseman, J. (2005a). A survival-adjusted quantal response test for comparing tumor incidence rates. *Journal of the Royal Statistical Society, **C***, **54**(**1**), 51–61.

Peddada, S. D., Dinse, G. and Kissling, G. (2007). Incorporating historical control data when comparing tumor incidence rates. *Journal of the American Statistical Association*, **102**(**480**), 1212–1220.

Peddada, S. D., Harris, S., Zajd, J. and Harvey, E. (2005b). ORIOGEN: order restricted inference for ordered gene expression data. *Bioinformatics*, **21**(**20**), 3933–4.

Peddada, S. D. and Jelsema, C. M. (2016). CLME: An R package for linear mixed effects models under inequality constraints. *Journal of Statistical Software*, **75**(**1**), 1–32.

Peddada, S. D., Laughlin, S. K., Miner, K., Haneke, K., Vahdat, H. L., Semelka, R. C., Kowalik, A., Armao, D., Davis, B. and Baird, D. D. (2008). Growth of uterine leiomyomata among premenopausal black and white women. *Proceedings of the National Academy of Sciences*, **105**(**50**), 19887–19892.

Praestgaard, J. (2012). A note on the power superiority of the restricted likelihood ratio test. *Journal of Multivariate Analysis*, **104**(**1**), 1–15.

Robertson, T., Wright, F. T. and Dykstra, R. L. (1988). *Order Restricted Statistical Inference*. Wiley, New York.

Rosen, S. and Davidov, O. (2017). Ordered regressions. *Scandinavian Journal of Statistics*, **44**(**4**), 817–842.

Ryan, T. P. (2013). *Sample Size Determination and Power*. John Wiley and Sons, Hoboken, NJ.

Silvapulle, M. and Sen, P. (2005). *Constrained Statistical Inference: Order, Inequality, and Shape Constraints*. Hoboken, NJ, New York.

Singh, S. P. and Davidov, O. (2019). On the design of experiments with ordered treatments. *Journal of Royal Statistical Society, **B***, **81**(**5**), 881–900.

Singh, S. P. and Davidov, O. (2020). On Bayes and Nash experimental designs for hypothesis testing problems. *Electronic Journal of Statistics*, **14**(**2**), 3976–4003.

Smucker, B., Krzywinski, M. and Altman, N. (2018). Optimal experimental design. *Nature Methods*, **15**, 557–560.

Spiegelhalter, D., Thomas, A., Best, N. and Gilks, W. (1999). *BUGS 0.5 Examples Volume 2, MRC Biostatistics Unit*. Institute of Public Health. Robinson Way, Cambridge CB2 2SR.

TR-527, N. (2005). NTP technical report on the toxicology and carcinogenesis studies of malachite green chloride and leucomalachite green. Technical report.

van Eden, C. (1956). Maximum likelihood estimation of ordered probabilities. *Proceedings, Koninklijke Nederlandse Akademie van Wetenschappen A*, **59**, 444–455.

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 107–114

A Decision Theory in Non-commutative Domain

Kalyan B. Sinha
JN Centre for Advanced Scientific Research
Indian Institute of Science and
Indian Statistical Institute,
Bangalore, India

Received: 01 October 2020; Revised: 14 October 2020; Accepted: 01 November 2020

Abstract and Prologue

In the later part of my professional life in the Indian Statistical Institute (I.S.I.), when I left Delhi to take up the position of the Director of I.S.I. in Kolkata, Aloke was my pillar of support, my person-to-go-to in any crisis; his was the shoulder to cry on. Those were, in many ways, difficult times for me and often I reflect and wonder how those times would have been without Aloke. In fact, from the mid-nineties, Aloke, in his extremely pleasant, shy and humble way, slowly but surely entered into a very close friendship in my life, which I will cherish forever. My frequent travels to Delhi, well after my retirement from I.S.I., would bring me to I.S.I., Delhi and to me that meant spending time with Aloke, long discussions inevitably ending with a very pleasant lunch with him in the chinese restaurant, opposite the gate of I.S.I. All these will remain only memories now and my next visit to Delhi (delayed by Covid-19) will be empty, "Aloke-heen" (in Bengali), and will make me miss him all the more. The following article, a brief introduction (by a non-expert) to the decision theory in a non-commutative (quantum-) background, is my humble tribute to Aloke and to his friendship for me.

Key words: Decision theory; Quantum theory; Bayes decision rule.

AMS Subject Classifications: 62K05

1. Introduction

The statistical decision-theory or the idea of founding Statistics on a theory of decisions is due to Abraham Wald, enunciated in its originality, in his famous book, "Statistical Decision Functions" (Wald (1950), for a more recent account see the book of Ferguson (1967). There have been attempts, mainly by Holevo (see for example the books of Holevo (2011) and Hayashi (2017)), to recast these ideas in the context of non-commutative probabilistic background. As is well-known (see the first half of the book of Parthasarathy (1992) for an elegant account), the mathematical Quantum Theory represents a model of a non-Kolmogoroffian (or non-commutative) probability theory and hence there should be good reason to explore the possibility of studying an extension of the (classical) decision-theory to this domain. To give a brief account of this is the aim here.

Corresponding Author: Kalyan B. Sinha
Email: kbs@jncasr.ac.in

2. The Mathematical Description of The (Classical) Decision-Theory

As Ferguson (1967) observes in his book, the theory of games, as introduced by von Neumann in the 1940's, has a great deal of similarity with many aspects of decision theory. Both of these two theories start with three basic objects:
(i) a non-empty set of parameter, Θ, parametrizing the possible states of the system;
(ii) a non-empty set, Ω, of decisions (or actions) available to the statistician;
(iii) a function $L : \Theta \times \Omega \to \mathbb{R}$, called the **loss function** (the negative values of L needs to be interpreted as gain).

This triplet (Θ, Ω, L) defines a statistical decision problem or a game with the following interpretation. The nature (or providence!) chooses a point θ in Θ and the statistician, with no knowledge of the choice nature has made, makes a decision (or chooses an action) ω in Ω. As a consequence of these decisions, the statistician loses an amount $L(\theta, \omega)$. While in game-theoretic context, the players are trying simultaneously to minimize their losses, since the nature chooses the state without any such bias (hopefully!), this presents a dilemma for the decision-statistician and she tries to resolve this dilemma by gathering more information on the state by "sampling or by performing many experiments".

Thus for the decision-statistician, there is also a sample space $\mathcal{X}$ (here taken to be a Borel subset of $\mathbb{R}^d$, the d-dimensional Euclidean space) with a family of probability measures $\{\mu_\theta\}_{\theta\in\Theta}$ on $\mathcal{F}(\mathcal{X})$, the Borel σ-algebra of $\mathcal{X}$. The statistical decision problem, given by the triple (Θ, Ω, L) along with the sample space $\mathcal{X}$ of experiments, next chooses a (behavourial) decision map $D : \mathcal{X} \times \mathcal{F}(\Omega) \to \mathbb{R}_+$ such that $D(x, \cdot)$ is a probability measure on the Borel σ-algebra $\mathcal{F}(\Omega)$. Next one writes down the **risk function** $R : \Theta \times \{D\} \to \mathbb{R}$ by

$$R(\theta, D) = \int_\Omega L(\theta, \omega) \int_{\mathcal{X}} \mu_\theta(dx) D(x, d\omega). \tag{1}$$

An instructive way to rewrite (1) is to define the measure $\mu_\theta \circ D : \mathcal{F}(\Omega) \mapsto \mathbb{R}_+$ for every $\theta \in \Theta$ and $\Delta \in \mathcal{F}(\Omega)$ by

$$(\mu_\theta \circ D)(\Delta) = \int_{\mathcal{X}} \mu_\theta(dx) D(x, \Delta) \tag{2}$$

and replacing (1) by

$$R(\theta, D) = \int_\Omega L(\theta, \omega)(\mu_\theta \cdot D)(d\omega), \tag{3}$$

whenever the integral exists. Here we have noted that if $\mathcal{X} \ni x \mapsto D(x, \Delta)$ is measurable, then $\forall\ \theta$, $\mu_\theta \circ D$ is a probability measure on Ω and one can give a meaning to the integral in (3). The risk function R represents the average loss to the statistician when the nature has chosen the state parametrized by θ and the decision made is represented by the decision map D.

At this stage, one is still left with the problem of the "choice of parametrization" $\theta \in \Theta$ of the state and of the several avenues adopted by a statistician, we shall restrict our

discussions here to the use of the "Bayes Principle". This involves putting a structure of a measure space on Θ and assigning a "prior probability measure"π on the σ-algebra $\mathcal{F}(\Theta)$. This leads naturally to the definition of the **Bayes risk of a (behavourial-)decision rule** D **with respect to the prior** π as

$$\begin{aligned}\mathcal{R}(\pi, D) &= \int_{\Theta} \pi(d\theta) R(\theta, D) \\ &= \int_{\Theta} \pi(d\theta) \int_{\Omega} L(\theta, \omega)(\mu_\theta \circ D)(d\omega). \end{aligned} \tag{4}$$

With regard to the definition (4), there are a few technical issues, *e.g.* the sense of measurability of the map $\theta \mapsto \mu_\theta \circ D$ etc., but these can be easily treated; for example in the above mentioned case one can have the assumption that $\theta \mapsto \mu_\theta(\cdot)$ is measurable and refer to [Dunford and Schwarz (1988), pages 156-162]. It is also worth mentioning that often authors (*e.g.* in Wald (1950)) consider the parameter space Θ to be finite or countably infinite. Also note, since all the 3 set-functions are non-negative, one can define a conditional probability measure of the random variable $\widehat{\theta}$ on Θ, given the random variable X on $\mathcal{X}$ (called the **posterior** probability measure of $\widehat{\theta}$ given the observation of X) on the product σ-algebra $\mathcal{F}(\Theta) \times \mathcal{F}(\mathcal{X})$ by

$$(\pi \cdot \mu)(\delta \times \Delta) = \int_{\delta} \pi(d\theta) \mu_\theta(\Delta) \tag{5}$$

for $\delta \in \mathcal{F}(\Theta), \Delta \in \mathcal{F}(\mathcal{X})$. In fact, in Ferguson (1967) the possibility of these two definitions (4) and (5) are pre-conditions for speaking about the "Bayes decision principle". This definition (5) sets up a linear ordering on the set $D(\cdot, \cdot)$ of decision functions and a Bayes decision rule is one that has the smallest Bayes risk, $\mathcal{R}$.

A decision function D_0 is said to be **Bayes with respect to the prior measure** π if

$$\mathcal{R}(\pi, D_0) = \inf_D \mathcal{R}(\pi, D). \tag{6}$$

It may happen that even if the right hand side of (6) exists, that value may **not** be attained for any D_0 and in such a case, one has to be satisfied with a decision D_0; which is "close" to the infimum. Let $\epsilon > 0$. A decision function $D_0 = D_0(\epsilon)$ is said to be ϵ**-Bayes** if

$$\mathcal{R}(\pi, D_0) \leq \inf_D \mathcal{R}(\pi, D) + \epsilon. \tag{7}$$

There are many other questions that arise naturally in the context of the above discussions; however, we shall take a break with the (classical) decision-theory and the rest of this article will be devoted to an attempt to "transport" the theory to the non-commutative (quantum) domain.

The definition (3) sets up a linear ordering (inherited from that of the real line) and the rule that is most preferred by that ordering is called the **minimax decision rule**: a

decision map $D_0 \in \mathcal{D} \equiv$ the set $\{D : \mathcal{X} \times \mathcal{F}(\Omega) \to \mathbb{R}_+ \mid D(x, \cdot)$ is a probability measure with variation norm uniformly bounded w.r.t. $x \in \mathcal{X}\}$ is said to be minimax if

$$\sup_{\theta \in \Theta} R(\theta, D_0) = \inf_{D \in \mathcal{D}} \sup_{\theta \in \Theta} R(\theta, D). \tag{8}$$

If one assumes that (i) Θ and Ω are topological spaces such that Θ is compact, and $L : \Theta \times \Omega \to \mathbb{R}_+$ is continuous, (ii) $\Theta \ni \theta \mapsto \mu_\theta(\cdot)$ is continuous in w^*-topology of probability measures, then it can be seen that $\Theta \times \mathcal{D} \ni (\theta, D) \mapsto R(\theta, D)$ is continuous w.r.t the natural w^*-topology of $\mathcal{D}$, uniformly in θ. Therefore $\sup_\theta R(\theta, D)$ exists and $D \mapsto \sup_\theta R(\theta, D)$ is continuous w.r.t the w^*-topology of $\mathcal{D}$ in which $\mathcal{D}$ is compact. Thus the infimum exists and is attained since $\mathcal{D}$ is compact, *i.e.*, there exists a decision map D_0 with the property that

$$\inf_{D \in \mathcal{D}} \sup_{\theta \in \Theta} R(\theta, D) = \sup_{\theta \in \Theta} R(\theta, D_0).$$

A very similar proof for the partial quantum statistical decision rules can be constructed with $(\mu_\theta \circ D)(\cdot)$ replaced by $Tr_S(\rho_\theta D(\cdot))$ and very similar results can be obtained with same set of assumptions, as explained below.

3. Quantum Theory of Bayes' Decision-rules

If one thinks of the Quantum Theory as one possible model for non-Kolmogoroffian probability (see Partasarthy (1992) for an elaboration of this point of view), then the pair (sample space $\mathcal{X}$, real-valued random variable X) goes over to the relevant pair (Hilbert space h_S, a self-adjoint operator $\widehat{X}$ on it). Furthermore, the probability measure on $\mathcal{F}(X)$, associated with the random variable X is replaced by a density matrix ρ, a positive trace-class operator $(\mathcal{B}_{1+}(h_S))$ of trace 1, on h_S. In the present context of theory of decisions, there are two distinct possibilities:
(i) following Holevo's work (see Partasarthy (1992) and Holevo (1974)), one may have a kind of partial quantum (or non-commutative) statistical decision theory in which the sample space metamorphoses into its corresponding quantum structure, leaving the parameter-set Θ, a classical measure space with a prior probability measure π on it or (ii) a further or fully quantum statistical decision theory, in which the Bayesian part also undergoes a quantum metamorphosis. What turns out to be a remarkable coincidence (at least to the present author) that this second route has all the aspects of "quantum entanglement" (see. *e.g.* Petz (2008) and Parthasarthy (2013)) built in the mathematical structure.

For implementing the route (i), we first note that the sample space $\mathcal{X}$ is replaced by a (separable) Hilbert space h_S, the corresponding real-valued random variable X by a (possibly unbounded) self-adjoint operator $\widehat{X}$ in h_S and the family of probability measures $\{\mu_\theta(\cdot)\}_{\theta \in \Theta}$ by a family of density matrices $\rho \equiv \{\rho_\theta\}_{\theta \in \Theta} \in \mathcal{B}_{1+}(h_S)$ with $Tr_S(\rho_\theta) = 1$ for every θ, where Tr_S stands for the trace taken in the Hilbert space h_S. Furthermore, the triple (Θ, Ω, L) are given as before with Θ and Ω as two measure spaces and $L : \Theta \times \Omega \to \mathbb{R}_+$ measurable loss function. The most important change that takes place here is the replacement of the (behavourial) decision-function $D(x, \Delta)$ (for $x \in \mathcal{X}, \Delta \in \mathcal{F}(\Omega)$) by a map $D : \mathcal{F}(\Omega) \to \mathcal{B}_+(h_S)$, the set of non-negative bounded operators on h_S such that it is countably additive: $\{\Delta_j\}_{j=1}^\infty$ of disjoint subsets in $\mathcal{F}(\Omega)$ such that $\Delta = \bigcup_{j=1} \Delta_j$ implies that

$D(\Delta) = \sum_{j=1}^{\infty} D(\Delta_j)$ (the infinite sum converging in strong operator topology), and $D(\Omega) = I \in \mathcal{B}(h_S)$. This kind of family is called a POVM (positive operator-valued measures) on Ω (see Holevo (2011) and Davies (1976) for some applications of POVM). We can now define the partial quantum risk function (p.q.r.f) as:

$$R(\theta, D) = \int L(\theta, \omega) Tr_S(\rho_\theta D(d\omega)). \tag{9}$$

The right hand side makes sense since the map $\rho \circ D : \Theta \times \mathcal{F}(\Omega) \mapsto \mathbb{R}_+$ given by

$$(\rho \circ D)(\theta, \Delta) = Tr_S(\rho_\theta D(\Delta)) = Tr_S(\rho_\theta^{1/2} D(\Delta) \rho_\theta^{1/2}), \tag{10}$$

is a non-negative countably additive set-function with $(\rho \circ D)(\theta, \Omega) = 1$ and hence defines a probability measure on Ω for every $\theta \in \Theta$. Thus (9) makes sense as a Lebesgue integral and (9) can be rewritten as

$$R(\theta, D) = \int L(\theta, \omega)(\rho \cdot D)(\theta, d\omega). \tag{11}$$

Finally, with the prior probability measure π on $\mathcal{F}(\Theta)$, one has as in (4), the partial quantum Bayes' risk (p.q.B.r) of a (behaviourial) decision rule D:

$$\begin{aligned} \mathcal{R}(\pi, D) &= \int_\Theta \pi(d\theta) R(\theta, D) \\ &= \int_\Theta \pi(d\theta) \int_\Omega L(\theta, \omega)(\rho \cdot D)(\theta, d\omega). \end{aligned} \tag{12}$$

As in the classical case, one can define a (partially quantum) conditional density matrix of the random variable $\widehat{\theta}$ on Θ, given the (quantum) observation of the operator $\widehat{X}$ in h_S (we shall call it as **posterior density matrix** of $\widehat{\theta}$ given $\widehat{X}$ in h_S):

$$(\pi \cdot \rho)(\delta) = \int_\delta \pi(d\theta) \rho_\theta, \ \forall \ \delta \in \mathcal{F}(\Theta), \tag{13}$$

where the integral on the right hand side is the strong Bochner integral in the Banach space $\mathcal{B}_1(h_S)$. It is easy to see that this $\mathcal{B}_{1+}(h_S)$-valued set function on $\mathcal{F}(\Theta)$ is countably additive and $Tr_S(\pi \cdot \rho)(\Theta) = 1$. In fact, the Bayes risk p.q.B.r can be rewritten in terms of the posterior density matrix $(\pi \circ \rho)(\cdot)$ as

$$\mathcal{R}(\pi, D) = \int_{\Theta \times \Omega} L(\theta, \omega) Tr_S((\pi \cdot \rho)(d\theta) D(d\omega)). \tag{14}$$

In analogy, given a prior π, the partially quantum Bayes decision rule is the D which gives the smallest p.q.B.r and a decision D_0 (in $\mathcal{B}_+(h_S)$-valued POVM's on $\mathcal{F}(\Omega)$) is said to be **Bayes with respect to prior** π if

$$\mathcal{R}(\pi, D_0) = \inf_{D \in h_S - povm(\Omega)} \mathcal{R}(\pi, D). \tag{15}$$

In the rest of this article, we consider fully quantum decision theory in which the sample space $\mathcal{X}$ as well as parameter space Θ metamorphoses into two (separable) Hilbert space h_S and h_B, respectively, and $\pi \cdot \mu(\cdot)$ or $\pi \cdot \rho$ are replaced by one density matrix Φ on $\tilde{h} = h_S \otimes h_B$. This structure, in conjunction with the following assumptions constitute the present new proposal.

A1. Ω is a compact Borel space and the loss operator $L : \Omega \to \mathcal{B}_+(h_B)$ is continuous w.r.t the ω^*-topology of $\mathcal{B}(h_B)$;

A2. $D : \mathcal{F}(\Omega) \mapsto \mathcal{B}_+(h_S)$ is a POVM, as mentioned earlier and as in Holevo's theory.

Then we lift these two operator-families to the Hilbert space $\tilde{h}$ by setting

$$\begin{aligned} \tilde{L}(\omega) &= I_S \otimes L(\omega) \text{ for } \omega \in \Omega \text{ and} \\ \tilde{D}(\Delta) &= D(\Delta) \otimes I_B \text{ for } \Delta \in \mathcal{F}(\Omega). \end{aligned} \tag{16}$$

Note that $\tilde{L}(\omega)$ commutes with $\tilde{D}(\Delta)$ in $\tilde{h}$ and we define the **fully quantum risk function**

$$\mathcal{R}(\Phi, D) = \int_\Omega Tr_{\tilde{h}}(\Phi \tilde{L}(\omega)\tilde{D}(d\omega)), \tag{17}$$

which is

$$= \int_\Omega Tr_{\tilde{h}}\{(\tilde{L}(\omega)^{1/2}\Phi\tilde{L}(\omega)^{1/2})\tilde{D}(d\omega)\},$$

showing that $\mathcal{R}(\Phi, D) \geq 0$, if it exists. The issue of the sense in which the integral in (17) can be defined is not a simple one and it is left unresolved in this article, to be dealt with later. However, it should be mentioned that Holevo (see *e.g.* Holevo (1974)) gave a theory to study such integrals. Here we shall restrict ourselves to the simpler case when the density matrix Φ on $\tilde{h}$ is a **finite** linear combination of tensors of density matrices on h_S and h_B:

$$\Phi = \sum_{j=1}^n \rho_j \otimes \pi_j; \ \rho_j \in \mathcal{B}_{1+}(h_S), \pi_j \in \mathcal{B}_{1+}(h_B). \tag{18}$$

In such a case,

$$\Phi\tilde{L}(\omega)\tilde{D}(d\omega) = \sum_{j=1}^n (\rho_j D(d\omega) \otimes (\pi_j L(\omega))$$

and thus we shall be looking at the "integral"

$$\int_\Omega Tr_B(\pi_j L(\omega)) \cdot Tr_S(\rho_j D(d\omega)), \tag{19}$$

which exists as a Lebesgue-type integral since the function $Tr_S(\pi_j L(\cdot))$ is bounded continuous on compact Ω and since the second factor in (19) is clearly a (non-negative) finite measure with total variation $= Tr_S(\rho_j)$. For the rest of the discussion, viz. the one on a kind of

minimax theorem, we shall assume that the integral in (17) exists for all density matrices Φ on $h_S \otimes h_{\mathcal{B}}$.

As we have observed before, by virtue of the assumption A1, the map: density matrices on $\tilde{h} \ni \Phi \mapsto \mathcal{R}(\Phi, D) \in \mathbb{R}_+$ is continuous w.r.t. the w^*-topology on density matrices induced by $\mathcal{B}(\tilde{h})$ after applications of Mazur's theorem. Also note that Alaoglu's theorem implies that in the same topology, the set of density matrices is a (convex) compact set and therefore, there exists a density matrix Φ_0 such that

$$\sup_{\Phi} \mathcal{R}(\Phi, D) = \mathcal{R}(\Phi_0, D). \tag{20}$$

On the other hand, it is easy to see that

$$\begin{aligned} \sup_{\Phi} \inf_{D} \mathcal{R}(\Phi, D) &\leq \inf_{D} \sup_{\Phi} \mathcal{R}(\Phi, D) \\ &= \inf_{D} \mathcal{R}(\Phi_0, D) \\ &\leq \sup_{\Phi} \inf_{D} \mathcal{R}(\Phi, D) \end{aligned}$$

and therefore one has

$$\sup_{\Phi} \inf_{D} \mathcal{R}(\Phi, D) = \inf_{D} \sup_{\Phi} \mathcal{R}(\Phi, D). \tag{21}$$

The left hand side is called the lower value and the right hand side the upper value and equality of these two constitutes the **minimax** decision rule.

The procedure and results, indicated above can be strengthened more, in line with the classical case, if instead we ask the following:

Given $\sigma \in \mathcal{B}_{1+}(h_B)$, let $\mathcal{S}_\sigma = \{\Phi \in \mathcal{B}_{1+}(h_S \otimes h_B) \mid Tr_S \Phi = \sigma\}$.

Then does there exist a POVM D_0 such that $\sup_{\Phi \in \mathcal{S}_\sigma} \inf_{D} \mathcal{R}(\Phi, D) = \sup_{\Phi \in \mathcal{S}_\sigma} \mathcal{R}(\Phi, D_0)$?

Acknowledgement

This endeavour is partly supported by the I.N.S.A. Senior Scientist programme of the Indian National Science Academy, for which the author is grateful.

References

Davies, E. B. (1976). *Quantum Theory of Open Systems.* Academic Press, New York.

Dunford, N. and Schwartz, J. T. (1988). *Linear Operators - Part I.* John Wiley, USA.

Ferguson, T. (1967). *Mathematical Statistics - A Decision Theoretic Approach.* Academic Press, New York.

Hayashi, M. (2017). *Quantum Information Theory- Mathematical Foundations*, Second Edition, Springer-Verlag, Berlin.

Holevo, A. S. (1974). The theory of statistical decisions on an operator algebra. *Dokaldy Akademii Nauk SSSR*, **218(1)**, 54–57.

Holevo, A. S. (2011). *Probabilistic and Statistical Aspects of Quantum Theory*, Second Edition. Edizioni Della Normale, Pisa.

Parthasarathy, K. R. (1992). *Introduction to Quantum Stochastic Calculus*. Birkhauser Verlag, Basel.

Parthasarathy, K. R. (2013). *Coding Theorems of Classical and Quantum Information Theory*. TRIM series, HBA, New Delhi.

Petz, D. (2008). *Quantum Information Theory and Quantum Statistics*. Springer Verlag, Berlin.

Wald, A. (1950). *Statistical Decision Functions*. Wiley, New York.

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 115–124

Nonlinear Error-in-Variables Regression with the Error-in-Variables Distribution Estimated by Censored Data

M. S. Hamada[1] **and K. A. Kaufeld**[1]
[1]*Statistical Sciences*
Los Alamos National Laboratory, New Mexico, USA

Received: 23 October 2020; Revised: 22 November 2020; Accepted: 26 November 2020

Abstract

This article considers the analysis of data using a nonlinear regression model in which the covariate has a distribution, *i.e.*, the error-in-variables case. Moreover, some of the same data consisting of left- and right-censored data are used to estimate the covariate distribution. We show how to simultaneously fit the nonlinear error-in-variables regression model and estimate the covariate distribution using Bayesian inference. The proposed method is illustrated with a simulated data set. We also show the impact of knowing the covariate distribution and the actual covariate values. Furthermore, we show the impact of taking additional data on inference and prediction.

Key words: Bayesian inference; Left-censored; Markov chain Monte Carlo; Prediction; Right-censored.

AMS Subject Classifications: 62F15, 62J02, 62N01, 62P30

1. Introduction

It is our privilege to to contribute this article to the special issue of Statistics and Applications in honor of Professor Dey. The first author met Professor Dey when he visited the University of Waterloo in the late 1980's. At the time, research in the design of experiments for improving quality and productivity in industry had been reinvigorated by the appearance of Taguchi Methods. Professor Dey's 1985 book was timely for its mixed-level orthogonal arrays that were being promoted by the Taguchi Methods. The Wu and Hamada (2009) Experiments book refers to Professor Dey's 1985 book as well as his 1999 book with Professor Mukerjee a number of times for theoretical details and presents tables of his OA$(24, 6^1, 2^{14})$, OA$(54, 2^1, 3^{25})$, and OA$(54, 6^1, 3^{24})$ designs for use by practitioners. The first author fondly remembers Professor Dey as a formal gentleman and seasoned scholar who kindly spent time talking to a young assistant professor about research. In this article, we present a problem that we faced on a project at work. Here we focus on data analysis although there is a design aspect that could be explored.

Suppose that we sample a population each year for I years, $i = 1, \ldots, I$. At year i, we sample a unit and record whether a feature of interest can be observed in the unit. For

Corresponding Author: Michael Hamada
Email: hamada@lanl.gov

example, cracks occur on containers based upon stresses on the container. The initiation crack area or subsequent cracks occur at some time point. If a crack is observed we know that the crack started sometime before that year. Otherwise, we know the crack will start after the recorded time. In terms of chemical reactions, we can think about a reaction occurring at a recorded time. If observed, all that we know is that the mechanism started in the unit before year i; if it is not observed, all that we know is that the mechanism will start in the unit after year i. That is, we assume that the mechanism will start at some time in all units so that there is a start time distribution. The data in which the mechanism has not started are right-censored data. The data in which the mechanism has started are left-censored data. An example of a model that displays similar characteristics is convex degradation where the degradation rate increases with the level of degradation (Meeker and Escobar, 1998). Suppose that the start time distribution is $Lognormal(\mu, \sigma^2)$, say $Lognormal(3, 0.1^2)$ with median 20.1 years and 0.95 probability interval (14.4, 27.0) years. Recall that the log start time distribution is $Normal(\mu, \sigma^2)$. The proportion of the population that the mechanism has started at time t is displayed in Figure 1.

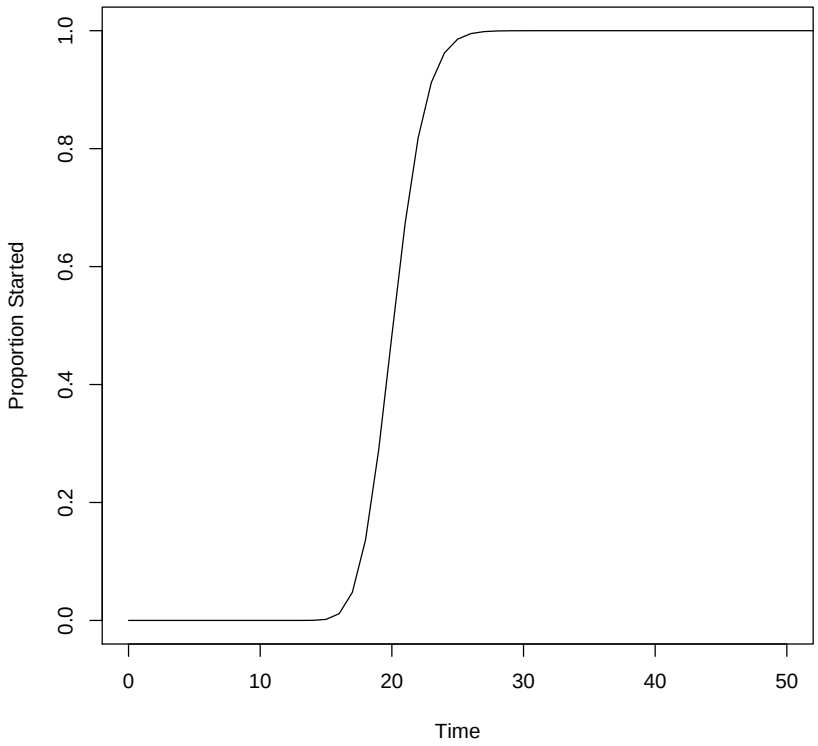


Figure 1: Proportion of population that mechanism has started versus time (years).

Suppose that for those units in which the mechanism has started we observe a quantity Y_t at time t, which is modeled as $Y_t = \beta_0(1-\exp(-\beta_1 e_t)) + \epsilon_t$. This is a nonlinear regression model with mean $\beta_0(1-\exp(-\beta_1 e_t))$, where e_t is the elapsed time between the time when the mechanism started s (*i.e.*, the start time) and time t (*i.e.*, $e_t = t - s$). ϵ_t is the population error assumed to be distributed as $Normal(0, \sigma_\epsilon^2)$ and is assumed independent of the start

time s. Suppose that $\beta_0 = 1000$, $\beta_1 = 0.025$ and $\sigma_\epsilon = 1$. The mean $\beta_0(1 - \exp(-\beta_1 e_t))$ versus elapsed time e_t is displayed in Figure 2.

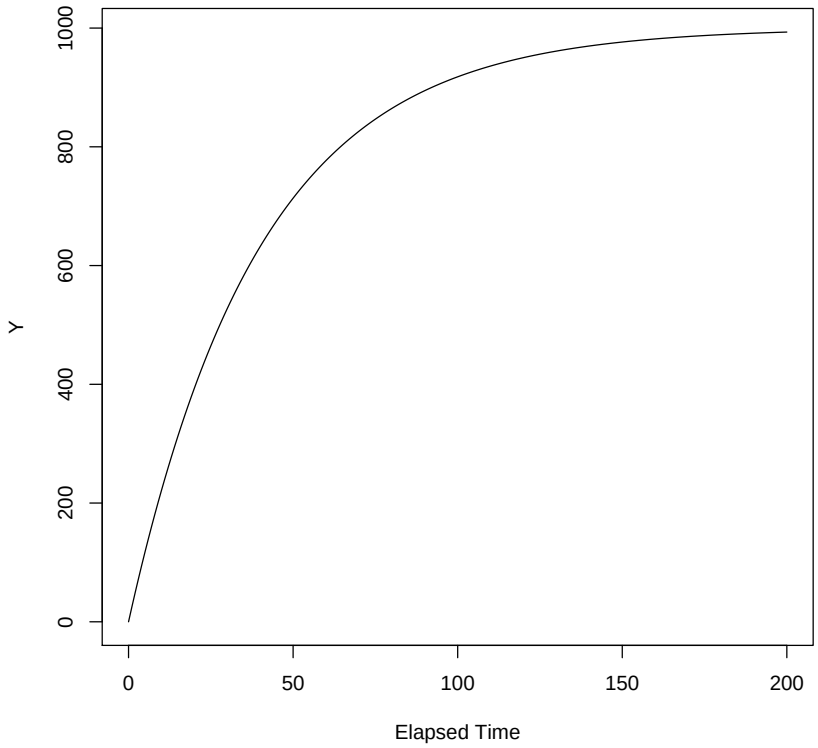


Figure 2: Quantity Y nonlinear regression mean versus elapsed time (years).

2. Data Model and Analysis

In our scenario, we will use example data shown in Table 1, where one unit is sampled per year for 30 years at times 1-30. The StartStar column is 1 if the mechanism is not observed to have started; otherwise, 0 if the mechanism is observed to have started. The Y column is the quantity Y if the mechanism is observed to have started; *e.g.*, if a crack is observed, Y might be the the length of the crack. The elapsed time is not known; we only know that at time t with StartStar equal to 0, the start time s is less than t, *i.e.*, the elapsed time e_t is a random variable, $t - s$, where $s \sim Lognormal(3, 0.1^2)I(0, t)$ and $I(0, t)$ indicates that the lognormal distribution is restricted to the interval $(0, t)$. Because e_t is not known exactly, but has a distribution, the nonlinear regression model of Y is an error-in-variables model where the covariate e_t has a distribution and not an exactly known value.

Further, we use the Time-StartStar data to estimate μ and σ for the $Lognormal(\mu, \sigma^2)$ start time distribution. For a StartStar of 1, say, for Time 2, the likelihood contribution is $1 - \Phi(\frac{\ln(2)-\mu}{\sigma})$, the probability of observing a right-censored datum, where $\Phi()$ is the normal cumulative distribution function. For a StartStar of 0, say, for Time 19, the likelihood contribution is $\Phi(\frac{\ln(19)-\mu}{\sigma})$, the probability of observing a left-censored datum.

Table 1: Example Data (ordered so that right-censored data appear first; Elapsed Time is unknown to the analyst)

Time (year)	StartStar	Elapsed Time (year)	Y
1	1	0	0
2	1	0	0
3	1	0	0
4	1	0	0
5	1	0	0
6	1	0	0
7	1	0	0
8	1	0	0
9	1	0	0
10	1	0	0
11	1	0	0
12	1	0	0
13	1	0	0
14	1	0	0
15	1	0	0
16	1	0	0
17	1	0	0
18	1	0	0
21	1	0	0
19	0	0.47	12.36
20	0	1.15	27.18
22	0	0.51	13.70
23	0	3.66	85.82
24	0	6.62	151.98
25	0	5.41	125.38
26	0	4.68	110.55
27	0	3.66	88.28
28	0	5.92	138.13
29	0	9.97	219.96
30	0	6.99	160.38

We use a Bayesian analysis with the following relatively diffuse prior distributions (Gelman *et al.*, 2013):

- $\beta_0 \sim \mathit{Lognormal}(7, 0.5^2)$ with a 0.95 probability central interval of (411.6, 2921.9)
- $\beta_1 \sim \mathit{Lognormal}(-4, 1^2)$ with a 0.95 probability central interval of (0.003, 0.130)
- $\sigma_t \sim \mathit{HalfNormal}(0, \sqrt{10}^2)$ with a 0.95 probability central interval of (0.099, 7.088)

- $\mu \sim \mathit{HalfNormal}(0, \sqrt{10}^2)$ with a 0.95 probability central interval of (0.099, 7.088)
- $\sigma \sim \mathit{HalfNormal}(0, \sqrt{10}^2)$ with a 0.95 probability central interval of (0.099, 7.088)

These prior distributions are thought to be relatively diffuse, *i.e.*, they are chosen to be quite wide so that the true values of the parameters are thought to fall within these high probability central intervals.

We obtain the following results using a Markov chain Monte Carlo (MCMC) algorithm (Gelman *et al.*, 2013) implemented in JAGS (Plummer, 2003) using the R (R Core Team, 2020) package rjags to call JAGS. The JAGS code for the proposed analysis is given in the Appendix that produces 400,000 draws from the posterior distribution. In the examples, we use 10,000 burnin draws (which are discarded) and 40,000,000 subsequent draws, which we thin by taking every 100th draw. Plots of the posterior draws not shown here display good mixing. Moreover, diagnostics (Gelman and Rubin's convergence diagnostic; Gelman and Rubin, 1992) also suggest convergence, *i.e.*, these draws are from the appropriate posterior distribution.

Table 2 displays the posterior summaries for the model parameters μ and σ for the start time distribution and β_0, β_1 and σ_t for the quantity Y nonlinear regression model. Note that there is substantial uncertainty associated with σ_t.

Table 2: Posterior Summaries of Model Parameters (50, 2.5, 97.5 percentiles) from Table 1 Data

Parameter	True	50%	2.5%	97.5%
μ	3.000	2.975	2.853	3.073
σ	0.100	0.127	0.077	0.269
β_0	1000.000	1106.450	502.993	2583.997
β_1	0.025	0.018	0.007	0.046
σ_t	1.000	2.134	0.100	6.999

2.1. Impact of unknown starting times

There are two impacts of not knowing the starting times. First, the start time distribution parameters are estimated from the left- and right-censored start times. Second, the elapsed times are unknown because of the unknown start times; that is, the covariate in the nonlinear regression model is not known exactly and is referred to as an error-in-variables case. Table 3 shows the impact of using the true error-in-variables (E-I-V) distribution ($\mathit{Lognormal}(3, 0.1^2)$) as well as that of using the actual elapsed times (see Table 1 for the actual elapsed times). We see that using the true EIV distribution provides no improvement, at least for this one data set, but the nonlinear regression model parameters are substantially better estimated (with reduced uncertainty) when the actual elapsed times are used as compared with Table 2.

Table 3: Posterior Summaries of Model Parameters (50, 2.5, 97.5 percentiles) for Some Hypothetical Situations

Parameter	True	50%	2.5%	97.5%
use true E-I-V distribution				
β_0	1000.000	1121.012	509.559	2652.956
β_1	0.025	0.019	0.008	0.049
σ_t	1.000	2.132	0.101	7.023
use actual elapsed times				
β_0	1000.000	1012.378	851.154	1268.457
β_1	0.025	0.025	0.019	0.030
σ_t	1.000	1.016	0.662	1.821

3. Prediction

Suppose that we want to predict a percentile of the Y distribution, say the 90th percentile, at a given time, say 30 years. Suppose that the population size is 1,000. We can draw from the start time distribution to obtain start times and predict Y using the elapsed times (30 minus start times) and the nonlinear regression model, *i.e.* draw 1,000 Y's from the Y distribution. The 90th percentile is the 900th ordered prediction. We do this 10,000 times and take the 95th percentile of the 10,000 90th percentiles to obtain 165.43; for brevity we refer to this as the 90th percentile of the population Y distribution or even shorter as the 90th percentile. For times of 45 and 60 years, the 90th percentiles of the population Y distribution are 426.42 and 605.78, respectively. Based on the proposed analysis, we can obtain a posterior predictive distribution and a 0.95 probability upper bound on the the 90th percentile of the population Y distribution. Table 4 shows the Table 1 data posterior 90th percentile at times 30, 45, and 60 years. The posterior 90th percentiles are somewhat higher that the true 90th percentiles especially at times past the data, *i.e.*, 45 and 60 years.

Table 4: True and Table 1 Data 90th Percentiles at 30, 45, and 60 Years

Time (year)	True Percentile	Table 1 Data Percentile
30	165.43	165.56
45	426.42	440.09
60	605.78	669.36

4. Impact of Taking More Samples

We can also consider the impact of taking more samples per year and taking samples for more than 30 years, *e.g.*, 60 years, using the proposed analysis. Table 5 shows the results when 60 total samples are taken. We use the notation 1@1(1)30 for the Table 1 sampling

scheme, *i.e.*, 1 sample each year from years 1 to 30. Table 5 shows results for 2@1(1)30, 1@1(1)60, and 2@2(2)60; 2@2(2)60 denotes 2 samples in every even year from year 2 to year 60. Note that the first two schemes add to the Table 1 data. The 2@2(2)60 scheme uses data from the even years of the 2@1(1)30 scheme. Table 6 shows the results when 120 total samples are taken; 60 additional samples are added to the data analyzed that produced the Table 5 results. Table 6 shows results for 4@1(1)30, 2@1(1)60, and 4@2(2)60.

Increasing the sample from 1 to 2 to 4 per year (1@1(1)30, 2@1(1)30, 4@1(1)30) helps to estimate σ better; estimation for β_0 seems somewhat worse but recall these results are for one realization of the data. Spreading out the inspections across more years helps much more, *e.g.*, (1@1(1)60, 2@2(2)60) and (2@1(1)60, 4@2(2)60). Inspections on even years and more samples at each inspection helps more than inspecting every year with less samples at each inspection. It is noteworthy that none of sampling schemes had an impact on estimating σ_t so that the posterior distributions are similar to the prior distribution.

Table 5: Posterior Summaries of Model Parameters (50, 2.5, 97.5 percentiles) Using More Samples and More Years (60 total samples)

Parameter	True	50%	2.5%	97.5%
		2@1(1)30		
μ	3.000	3.013	2.938	3.077
σ	0.100	0.113	0.079	0.182
β_0	1000.000	1242.127	574.067	2762.086
β_1	0.025	0.020	0.008	0.049
σ_t	1.000	2.163	0.103	7.061
		1@1(1)60		
μ	3.000	2.998	2.907	3.079
σ	0.100	0.125	0.099	0.165
β_0	1000.000	1058.932	859.737	1523.670
β_1	0.025	0.023	0.013	0.033
σ_t	1.000	2.154	0.098	7.036
		2@2(2)60)		
μ	3.000	2.996	2.915	3.069
σ	0.100	0.108	0.086	0.141
β_0	1000.000	1117.917	916.186	1561.584
β_1	0.025	0.021	0.013	0.029
σ_t	1.000	2.168	0.096	7.199

Like Table 4 for the 1@1(1)30 sampling scheme, Tables 7 and 8 show the posterior 90th percentiles at 30, 45, and 60 years for the various sampling schemes with 60 and 120 total samples, respectively. Overall, the posterior 90th percentiles are quite close to the true 90th percentiles. The results for 2@1(1)30 and 4@1(1)30) are worse caused by the worse estimation for β_0 as noted previously. For some of the schemes, the posterior 90th

Table 6: Posterior Summaries of Model Parameters (50, 2.5, 97.5 percentiles) Using More Samples and More Years (120 total samples)

Parameter	True	50%	2.5%	97.5%
		4@1(1)30		
μ	3.000	3.026	2.983	3.065
σ	0.100	0.091	0.072	0.123
β_0	1000.000	1237.483	580.953	2966.605
β_1	0.025	0.021	0.008	0.053
σ_t	1.000	2.069	0.095	6.726
		2@1(1)60		
μ	3.000	3.004	2.943	3.061
σ	0.100	0.121	0.103	0.146
β_0	1000.000	986.211	866.839	1171.483
β_1	0.025	0.025	0.019	0.032
σ_t	1.000	2.214	0.102	7.207
		4@2(2)60		
μ	3.000	3.019	2.964	3.071
σ	0.100	0.103	0.088	0.124
β_0	1000.000	978.467	881.563	1144.787
β_1	0.025	0.026	0.020	0.032
σ_t	1.000	2.085	0.097	6.875

percentiles are slightly less the true 90th percentiles; again these results are for one realization of the data.

Table 7: 60 Sample Data 90th Percentiles at 30, 45, and 60 Years

Time (year)	True Percentile	1@1(1)30 Percentile	2@1(1)30 Percentile	1@1(1)60 Percentile	2@2(2)60 Percentile
30	165.43	165.56	179.38	163.44	166.44
45	426.42	440.09	502.20	424.13	421.92
60	605.78	669.36	767.29	611.60	611.29

5. Discussion

In this article, we considered a nonlinear regression model with elapsed time as a covariate for a quantity Y. The elapsed time is the difference between the inspection time and the time when a mechanism started. At inspection, we only know that the mechanism has started or not so that the elapsed time is unknown, the error-in-variables case. Our proposed method analyzes the right- and left-censored elapsed time data to estimate the elapsed time distribution. This analysis is achieved simultaneously with analyzing the error-in-variables (E-I-V) nonlinear regression model for the Y data, where the elapsed time distribution is the E-I-V distribution. Besides the original 30 sample scheme, we showed results for various 60

Table 8: 120 Sample Data 90th Percentiles at 30, 45, and 60 Years

Time (year)	True Percentile	4@1(1)30 Percentile	2@1(1)60 Percentile	4@2(2)60 Percentile
30	165.43	187.22	161.56	167.15
45	426.42	520.54	422.68	430.83
60	605.78	803.17	603.25	610.77

sample and 120 schemes. Note that the results are based on one data set for each of these schemes where the smaller schemes data or parts of the smaller schemes data are included in the larger schemes data. Generally, the results improve for more samples per year over more years. A more extensive study using more data sets, say 500 or more, would solidify the results but would require access to a large computer cluster. Future research might consider an optimal sampling scheme that specifies how may samples and what inspection times to takes the samples. It would be natural to use a Bayesian design criterion because of the proposed Bayesian analysis method.

Acknowledgements

We thank Cee Cee Essex for her support and encouragement. We also thank an anonymous referee provided helpful comments on an earlier version of this article.

References

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A. and Rubin, D.B. (2013). *Bayesian Data Analysis Third Edition*. Boca Raton: Chapman & Hall/CRC.

Gelman, A. and Rubin, D.B. (1992). Inference from iterative simulation using multiple sequences. *Statisistical Sciences*, **7**, 457-472.

Meeker, W.Q. and Escobar, L.A. (1998). *Statistical Methods for Reliability Data*. New York: John Wiley & Sons, Inc.

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, March 20-22, Vienna, Austria. R Core Team (2020). *R: a Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna. (http://www.R-project.org)

Wu, C.F.J. and Hamada, M.S. (2009). *Experiments: Planning, Analysis and Optimization, Second Edition*. New York: John Wiley and Sons, Inc.

APPENDIX

This appendix presents JAGS code for the proposed analysis. In the code:

- startStar is 1 if right-censored and 0 if left-censored, *i.e.*, the mechanism has not started or has started, respectively
- right-censored data are ordered first
- N1=19, number of right-censored data for the Table 1 data
- N2=11, number of left-censored data
- inspect is the time of the inspection (sampling)
- start is the unobserved start time
- et is the unobserved elapsed time
- resp is the response or quantity Y
- ra is β_0
- rb is β_1
- sigmaResp is σ_t
- mu is μ
- sigma is σ

```
model
{
for( i in 1 : N1 ) {
startStar[i] ~ dinterval(start[i],inspect[i])
start[i] ~ dlnorm(mu,tau) # second parameter is a precision, \textit{i.e.}, reciprocal varianc
}
for( i in (N1+1) : (N1+N2) ) {
startStar[i] ~ dinterval(start[i],inspect[i])
start[i] ~ dlnorm(mu,tau)
}
for( i in (N1+1) : (N1+N2) ) {
resp[i] ~ dnorm(muResp[i],tauResp) # second parameter is a precision
muResp[i]<- ra*(1-exp(-rb*et[i]))
et[i]<-inspect[i]-start[i]
}

#priors
ra~dlnorm(7,(1/(.5*,5)))
rb~dlnorm(-4,1)
tauResp <- 1/(sigmaResp*sigmaResp)
sigmaResp ~ dnorm(0,1.0E-1)I(0,)
mu ~ dnorm(0.0,1.0E-1)I(0,)
tau <- 1/(sigma*sigma)
sigma ~ dnorm(0,1.0E-1)I(0,)
}
```

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021(New Series), pp 125–140

Unifying Constructions of Group Divisible Designs

Shyam Saurabh[1], **Kishore Sinha**[2] **and Mithilesh Kumar Singh**[3]
[1]*Ranchi University, Ranchi, India*
[2] *Formerly at Birsa Agricultural University, Ranchi, India*
[3] *Formerly at Ranchi University, Ranchi, India*

Received: 30 October 2020; Revised: 28 November 2020; Accepted: 04 December 2020

Abstract

The purpose of this paper is to unify constructions of group divisible designs by making use of certain balanced incomplete block designs, skew-Hadamard matrices, regular Hadamard matrices, balanced generalized Weighing matrices, Conference matrices and generalized Conference matrices. The constructions unify the results of Dey (1977), Dey and Nigam (1985), Parihar and Shrivastava (1988), De and Roy (1990) and generalize some results of Bhagwandas *et al*. (1985), Sinha (1991*b*) and Kadowaki and Kageyama (2009). In the process of investigations, some group divisible designs in the range of r, $k \leq 10$ are found and catalogued. These designs are obtained from the works of other authors but are not reported in Clatworthy (1973) and Sinha (1991*a*).

Keywords: Balanced incomplete block designs; Group divisible designs; Generalized Hadamard matrices; Generalized Conference matrices; Generalized Weighing matrices.

0. Prologue

Dr. Kishore Sinha had the opportunity of working with Professor Aloke Dey, at IASRI, New Delhi as a Post- doctoral research fellow of CSIR, New Delhi during 1977- 1979. It was during this period that he got fascinated with the research work of Professor Aloke Dey especially in the area of Partially Balanced Incomplete Block (PBIB) Designs. His association with Professor Aloke Dey continued growing in strength even after he left IASRI in 1979. His untimely demise has been a personal loss to Kishore in particular and to statistician's fraternity in general.

Various methods of constructions and trial and error solutions of group divisible designs are available and scattered over the literature. To the best of our knowledge, Dey (1977) for the first-time used matrix approach for the constructions of group divisible designs. His works motivated us to take up unification and generalization of constructions of group divisible designs. It is my proud privilege to pay my most respectful homage by dedicating this research paper to his memory.

Corresponding Author: Kishore Sinha
Email: kishore.sinha@gmail.com

1. Introduction

Some relevant definitions in the context of the paper are as follows:

1.1. Group divisible designs

A *Group divisible* (*GD*) *design* is an arrangement of v (= mn; $m, n \geq 2$) treatments into b blocks such that each block contains k ($<v$) distinct treatments, each treatment occurs r times and any pair of distinct treatments which are first associates occur together in λ_1 blocks and in λ_2 blocks if they are second associates. Furthermore, if $r-\lambda_1 = 0$ then the GD design is singular; if $r-\lambda_1 > 0$ and $rk-v\lambda_2 = 0$ then it is semi-regular (SR); and if $r-\lambda_1 > 0$ and $rk-v\lambda_2 > 0$, the design is regular (R). Semi- regular and regular GD designs are denoted by SRGD and RGD respectively. Following Cheng (1995), GD designs with parameters satisfying $b = 4(r-\lambda_2)$ are called family (A) GD designs.

1.2. α- Resolvable design

A block design D (v, b, r, k) whose b blocks can be divided into $t = r/\alpha$ classes, each of size $\beta = v\alpha/k$ and such that in each class of β blocks every treatment of D is replicated α times, is called an α- resolvable design. When α=1 the design is said to be resolvable.

1.3. Hadamard matrices

An $n\times n$ matrix $\mathbf{H} = (H_{ij})$ with entries H_{ij} as ± 1 is called a *Hadamard matrix* if $\mathbf{H}\mathbf{H}'=\mathbf{H}'\mathbf{H}=n\mathbf{I}_n$, where $\mathbf{H}'$ is the transpose of $\mathbf{H}$ and $\mathbf{I}_n$ is the identity matrix of order n. A Hadamard matrix is in normalized form if its first row and first column contain only +1's. A Hadamard matrix $\mathbf{H}$ is said to be of *skew type or skew- Hadamard* if its main diagonal entries are +1 and $\mathbf{H}-\mathbf{I}_n$ is *skew- symmetric*. In other words, a Hadamard matrix is called skew-symmetric if $H_{ij} = -H_{ji}\ \forall i \neq j$ and $H_{ii} = 1\ \forall i$.

Example 1: $\mathbf{H} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ and $\mathbf{H} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$ are Hadamard matrices of order 2 and 4 respectively.

Example 2: $\mathbf{H} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ and $\mathbf{H} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \end{pmatrix}$ are skew- Hadamard matrices of order 2 and 4 respectively.

1.4. Conference matrices

A *Conference matrix* of order n is an $n\times n$ matrix $\mathbf{C}$ with diagonal entries 0 and off-diagonal entries ± 1 such that $\mathbf{C}\mathbf{C}' = (n-1)\mathbf{I}_n$. A *Conference matrix* $\mathbf{C}$ is symmetric if $\mathbf{C} = \mathbf{C}'$ and skew- symmetric if $\mathbf{C} = -\mathbf{C}'$. A Conference matrix of order n is denoted as CM (n).

Example 3: $\mathbf{C} = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & -1 & -1 & 1 \\ 1 & 1 & 0 & 1 & -1 & -1 \\ 1 & -1 & 1 & 0 & 1 & -1 \\ 1 & -1 & -1 & 1 & 0 & 1 \\ 1 & 1 & -1 & -1 & 1 & 0 \end{pmatrix}$ is a symmetric Conference matrix and

$\mathbf{C} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ -1 & 0 & -1 & 1 \\ -1 & 1 & 0 & -1 \\ -1 & -1 & 1 & 0 \end{pmatrix}$ is a skew – symmetric Conference matrix.

1.5. Regular Hadamard matrices

A Hadamard matrix is *regular* if sum of the elements in any row of the matrix is constant. It is known that the order of a regular Hadamard matrix is a perfect square $4t^2$, t a positive integer. The number of entries +1 in any row is a constant, either $2t^2–t$ or $2t^2+t$. In the first case, any two rows will have $t^2 –t$ positions wherein both have entry +1; the second case has $t^2 + t$ positions wherein both have entry +1. For methods of construction, see Crnkovic (2006).

Example 4: $\mathbf{H} = \begin{pmatrix} -1 & -1 & -1 & 1 \\ -1 & -1 & 1 & -1 \\ -1 & 1 & -1 & -1 \\ 1 & -1 & -1 & -1 \end{pmatrix}$ is a regular Hadamard matarix of order 4.

1.6. Generalized weighing matrix

Let $n \geq w \geq 1$. A *Weighing matrix* $\mathbf{W}(n, w)$ of order n and weight w is an $n \times n$ $(0, \pm 1)$ – matrix such that $\mathbf{WW}' = w\mathbf{I}_n$.

A *generalized Weighing matrix* is a $v \times b$ matrix $\mathbf{M} = (m_{ij})$ with entries 0 and elements of a multiplicative group G of order g such that the inner product of any pair of distinct rows contains every element of G same number of times.

A generalized Weighing matrix $v \times b$ with the additional property that every row contains precisely r nonzero entries, each column contains exactly k nonzero entries and the inner product of any pair of distinct rows contains every group element exactly λ/g times, is known as a *generalized Bhaskar Rao design* GBRD $(v, b, r, k, \lambda; G)$. By replacing its nonzero entries by unity, produces the incidence matrix of a BIB design (v, b, r, k, λ).

A *Bhaskar Rao design* BRD (v, b, r, k, λ) is a $v \times b$ $(0, \pm 1)$ – matrix such that the inner product of any pair of distinct rows is zero and replacing –1 by unity, produces the incidence matrix of a BIB design (v, b, r, k, λ).

A GBRD $(v, b, r, k, \lambda; G)$ with $r = k$ and $v = b$ is also known as a *balanced generalized Weighing matrix* BGWM $(v, k, \lambda; G)$.

If the diagonal entries of BGWM $(v, k, \lambda; G)$ are zero and the inner product of any pair of distinct rows contains each element of G exactly λ times, then it is known as *generalized*

Conference matrix, GCM (G; λ). The order of *GCM* (G; λ) is $\lambda g+2$. If $G = \{\pm 1\}$, then *GCM* (G; λ) is a Conference matrix of order $2(\lambda+1)$. For details, see Colbourn and Dinitz (2007) and Tonchev (2009).

Example 5: A GBRD (4, 12, 9, 3, 6; C_6) over a cyclic group $C_6 = \{1, \alpha, \alpha^2, \alpha^3, \alpha^4, \alpha^5\}$ is

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & \alpha^4 & \alpha^2 & \alpha^3 & \alpha & \alpha^5 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & \alpha^2 & \alpha^4 & 0 & 0 & 0 & \alpha^3 & \alpha & \alpha^5 & \alpha^3 & \alpha & \alpha^5 \\ 0 & 0 & 0 & 1 & \alpha^2 & \alpha^4 & \alpha^3 & \alpha^5 & \alpha & 1 & \alpha^2 & \alpha^4 \end{pmatrix}.$$

Example 6: BRD (6, 6, 5, 5, 4) = $\begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & -1 & -1 & 1 \\ 1 & 1 & 0 & 1 & -1 & -1 \\ 1 & -1 & 1 & 0 & 1 & -1 \\ 1 & -1 & -1 & 1 & 0 & 1 \\ 1 & 1 & -1 & -1 & 1 & 0 \end{pmatrix}$.

Example 7: A 5 x 5 BGWM (5, 4, 3; C_3) over a cyclic group $C_3 = \{1, \alpha, \alpha^2\}$ is

$$\begin{pmatrix} 0 & \alpha^2 & \alpha & \alpha^2 & \alpha^2 \\ 1 & 0 & \alpha^2 & \alpha & \alpha^2 \\ 1 & 1 & 0 & \alpha^2 & \alpha \\ \alpha^2 & 1 & 1 & 0 & \alpha^2 \\ 1 & \alpha^2 & 1 & 1 & 0 \end{pmatrix}.$$

Example 8: A GCM (C_3; 2) of order 8 over a cyclic group $C_3 = \{1, \alpha, \alpha^2\}$ is

$$\begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & \alpha^2 & \alpha & \alpha & \alpha^2 & 1 \\ 1 & 1 & 0 & 1 & \alpha^2 & \alpha & \alpha & \alpha^2 \\ 1 & \alpha^2 & 1 & 0 & 1 & \alpha^2 & \alpha & \alpha \\ 1 & \alpha & \alpha^2 & 1 & 0 & 1 & \alpha^2 & \alpha \\ 1 & \alpha & \alpha & \alpha^2 & 1 & 0 & 1 & \alpha^2 \\ 1 & \alpha^2 & \alpha & \alpha & \alpha^2 & 1 & 0 & 1 \\ 1 & 1 & \alpha^2 & \alpha & \alpha & \alpha^2 & 1 & 0 \end{pmatrix}.$$

1.7. Generalized Hadamard matrix and Difference matrix

A *generalized Hadamard matrix* GHM (λ, g) over a group G of order g is a balanced generalized weighing matrix with $v = b = k = \lambda$. For GHM we require that the matrix should be square, but if we relax this condition and allow $v \times b$ ($v \leq b$) matrices, along with the conditions imposed on GHM, we obtain *difference matrices*. For details see Lampio (2015).

Example 9: $\text{GHM}\ (6,3) = \begin{pmatrix} 1 & \alpha & \alpha & 1 & \alpha^2 & \alpha^2 \\ \alpha^2 & \alpha^2 & \alpha & 1 & 1 & \alpha \\ \alpha^2 & \alpha & \alpha^2 & 1 & \alpha & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ \alpha & \alpha^2 & 1 & 1 & \alpha & \alpha^2 \\ \alpha & 1 & \alpha^2 & 1 & \alpha^2 & \alpha \end{pmatrix}$ is a generalized Hadamard

matrix with elements from the cyclic group $C_3 = \{1, \alpha, \alpha^2\}$.

Example 10: A 3×8 difference matrix over a cyclic group $C_4 = \{1, \alpha, \alpha^2, \alpha^4\}$ is

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & \alpha & \alpha & \alpha^2 & \alpha^2 & \alpha^3 & \alpha^3 \\ 1 & 1 & \alpha^2 & \alpha^3 & \alpha & \alpha^3 & \alpha & \alpha^2 \end{pmatrix}.$$

1.8. Kronecker sum of two matrices

Let $\mathbf{A}= (a_{ij})$ and $\mathbf{B}= (b_{ij})$ be two matrices of orders $m\times$n and $p\times q$ respectively over a field. Then the *Kronecker sum* $\mathbf{A} \oplus \mathbf{B}$ is an $mp\times nq$ matrix given by

$$\mathbf{A} \oplus \mathbf{B} = \mathbf{A} \otimes \mathbf{J}_{p,q} + \mathbf{J}_{m,n} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{J}_{p,q} + \mathbf{B} & a_{12}\mathbf{J}_{p,q} + \mathbf{B} & \cdots & a_{1n}\mathbf{J}_{p,q} + \mathbf{B} \\ a_{21}\mathbf{J}_{p,q} + \mathbf{B} & a_{22}\mathbf{J}_{p,q} + \mathbf{B} & \cdots & a_{2n}\mathbf{J}_{p,q} + \mathbf{B} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1}\mathbf{J}_{p,q} + \mathbf{B} & a_{m2}\mathbf{J}_{p,q} + \mathbf{B} & \cdots & a_{mn}\mathbf{J}_{p,q} + \mathbf{B} \end{pmatrix}.$$

where $\mathbf{J}_{v\times b}$ is the $v \times b$ matrix all of whose entries are 1, $\mathbf{A}\otimes\mathbf{B}$ is the Kronecker (or tensor) product of two matrices **A** and **B.**

Here, several methods of constructions of series of GD designs from certain BIB designs, skew Hadamard matrices, regular Hadamard matrices, balanced generalized Weighing matrices, Conference matrices and generalized Conference matrices are described. The constructions unify the results of Dey (1977), Dey and Nigam (1985), Parihar and Shrivastava (1988), De and Roy (1990) and generalize several results of Bhagwandas *et al*. (1985), Sinha (1991*b*) and Kadowaki and Kageyama (2009). A comprehensive coverage of constructions of GD designs may also be found in Arasu *et al.* (1991), Dey and Balasubramanian (1991), Dey (1986, 2010), Raghavarao (1971), Raghavarao and Padgett (2005). In the process of investigations, some group divisible designs in the range of $r, k \leq 10$ are found and catalogued. These designs are obtained from the works of other authors but are not reported in Clatworthy (1973) and Sinha (1991*a*).

The following notations are used: $\mathbf{I}_n$ is the identity matrix of order n, $\mathbf{J}_{v\times b}$ is the $v \times b$ matrix all of whose entries are 1 and $\mathbf{J}_{v\times v} = \mathbf{J}_v$, $\mathbf{A}\otimes\mathbf{B}$ is the Kronecker product of two matrices **A** and **B**, $\mathbf{A}'$ is the transpose of matrix **A** and $\mathbf{O}_n$ is null matrix of order n. *SRX* and *RX* numbers are from Clatworthy (1973). The design numbers *SRXa* and *RXa*, *b*, *c*, *d* are not found in Clatworthy (1973); and these designs are supposed to be located between SRX and SR(X+1) and RX and R(X+1) respectively.

2. The Constructions

2.1. From BIB designs

Theorem 1: There exists a GD design with parameters

$$v^* = vs, b^* = stv, r^* = t(k+s-1), k^* = k+s-1, \lambda_1 = (s-2)t, \lambda_2 = \lambda, m = v,$$

$$n = s; m, s \geq 2; t = r/\alpha \tag{1}$$

where *v, k, λ* are the parameters of an *α*- resolvable BIB design with $\lambda = t[(k+s-1)(k+s-2) - (s-1)(s-2)]/s(v-1)$.

Proof: Let $\mathbf{N}_i$ $(1 \leq i \leq t)$ represent the incidence matrices corresponding to resolution classes of an *α*- resolvable balanced incomplete block (BIB) design with parameters *v*, *b* = *tv*, *r*, *k* and $\lambda = t[(k+s-1)(k+s-2) - (s-1)(s-2)]/s(v-1)$ and also satisfying the condition $\sum_{i=1}^{t}(\mathbf{N_i} + \mathbf{N_i'}) = \lambda(\mathbf{J} - \mathbf{I})_v$.

Then the incidence pattern

$$\mathbf{M} = \mathbf{I_s} \otimes \mathbf{N}_{v \times tv} + (\mathbf{J_s} - \mathbf{I_s}) \otimes (\mathbf{I_v}|\mathbf{I_v}|\cdots|\mathbf{I_v})$$

$$= \begin{bmatrix} (\mathbf{N_1}|\mathbf{N_2}|\cdots|\mathbf{N_t}) & (\mathbf{I_v}|\mathbf{I_v}|\cdots|\mathbf{I_v}) & \cdots & (\mathbf{I_v}|\mathbf{I_v}|\cdots|\mathbf{I_v}) \\ (\mathbf{I_v}|\mathbf{I_v}|\cdots|\mathbf{I_v}) & (\mathbf{N_1}|\mathbf{N_2}|\cdots|\mathbf{N_t}) & \cdots & (\mathbf{I_v}|\mathbf{I_v}|\cdots|\mathbf{I_v}) \\ \vdots & \vdots & \ddots & \vdots \\ (\mathbf{I_v}|\mathbf{I_v}|\cdots|\mathbf{I_v}) & (\mathbf{I_v}|\mathbf{I_v}|\cdots|\mathbf{I_v}) & \cdots & (\mathbf{N_1}|\mathbf{N_2}|\cdots|\mathbf{N_t}) \end{bmatrix}$$

represents a GD design with parameters (1).

For *t* = *s* = 2 in Theorem 1 we obtain:

Corollary 1: There exists a GD design with parameters

$$v^* = 2v, b^* = 4v, r^* = 2(k+1), k^* = k+1, \lambda_1 = 0, \lambda_2 = \lambda = k(k+1)/(v-1),$$

$$m = v, n = 2.$$

For *t* = 2, *s* = 3 in Theorem 1 we obtain:

Corollary 2: There exists a GD design with parameters

$$v^* = 3v, b^* = 6v, r^* = 2(k+2), k^* = k+2, \lambda_1 = 2, \lambda_2 = \lambda = 2k(k+3)/3(v-1),$$

$$m = v, n = 3.$$

Table 1 lists GD designs constructed using Corollaries 1 and 2:

‘

Table 1: RGD from BIB designs

No.	GD: (v, r, k, b, λ_1, λ_2, m, n)	Source: 3- resolvable BIB design (v, r, k, b, λ)
1	$R$106: (10, 8, 4, 20, 0, 3, 5, 2)	(5, 6, 3, 10, 3), Corollary 1
2	$R$128: (26, 8, 4, 52, 0, 1, 13, 2)	(13, 6, 3, 26, 1), Corollary 1
3	$R$150: (15, 10, 5, 30, 2, 3, 5, 3)	(5, 6, 3, 10, 3), Corollary 2
4	$R$160: (39, 10, 5, 78, 2, 1, 13, 3)	(13, 6, 3, 26, 1), Corollary 2

The incidence matrix of a 3- resolvable BIB design with parameters (5, 6, 3, 10, 3) used for constructions of $R$106 and $R$150 in Table 1 can be partitioned as:

$$\mathbf{N_{5\times 10}} = (\mathbf{N_1}|\mathbf{N_2}) = \left(\begin{array}{ccccc|ccccc} 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \end{array}\right)$$

i. e. $\mathbf{N_1}$= *circ* (0 1 0 1 1) and $\mathbf{N_2}$= *circ* (0 1 1 1 0).

Example 11: The blocks of $R$106 using Corollary 1 are given as:
[(2, 3, 5, 6), (1, 3, 4, 7), (2, 4, 5, 8), (1, 3, 5, 9), (1, 2, 4, 10), (3, 4, 5, 6), (1, 4, 5, 7), (1, 2, 5, 8), (1, 2, 3, 9), (2, 3, 4, 10), (1, 7, 8, 10), (2, 6, 8, 9), (3, 7, 9, 10), (4, 6, 8, 10), (5, 6, 7, 9), (1, 8, 9, 10), (2, 6, 9, 10), (3, 6, 7, 10), (4, 6, 7, 8), (5, 7, 8, 9)].

The GD scheme is given as the 5 x 2 array: $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \end{bmatrix}'$.

Example 12: The blocks of $R$150 using Corollary 2 are given as:
[(2, 3, 5, 6, 11), (1, 3, 4, 7, 12), (2, 4, 5, 8, 13), (1, 3, 5, 9, 14), (1, 2, 4, 10, 15), (3, 4, 5, 6, 11), (1, 4, 5, 7, 12), (1, 2, 5, 8, 13), (1, 2, 3, 9, 14), (2, 3, 4, 10, 15), (1, 7, 8, 10, 11), (2, 6, 8, 9, 12), (3, 7, 9, 10, 13), (4, 6, 8, 10, 14), (5, 6, 7, 9, 15), (1, 8, 9, 10, 11), (2, 6, 9, 10, 12), (3, 6, 7, 10, 13), (4, 6, 7, 8, 14), (5, 7, 8, 9, 15), (1, 6, 12, 13, 15), (2, 7, 11, 13, 14), (3, 8, 12, 14, 15), (4, 9, 11, 13, 15), (5, 10, 11, 12, 14), (1, 6, 13, 14, 15), (2, 7, 11, 14, 15), (3, 8, 11, 12, 15), (4, 9, 11, 12, 13), (5, 10, 12, 13, 14)].

The GD scheme is given as the 5 x 3 array: $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \end{bmatrix}'$.

A 3- resolvable solution of BIB design with parameters (13, 6, 3, 26, 1) may be found in Kageyama and Mohan (1983). This solution for the construction of $R$128 and $R$160 in Table 1 can be partitioned as:

$\mathbf{N_{13\times 26}} = (\mathbf{N_1}|\mathbf{N_2})$ where $\mathbf{N_1}$= *circ* (0 0 0 0 1 0 0 0 0 0 1 0 1) and

$\mathbf{N_2}$ = *circ* (0 0 0 0 0 0 0 1 1 0 0 1 0).

Remark 1: The Corollary 1 gives patterned construction for $R106$ whereas an individual solution is given in Dey (1977).

2.2. From skew- Hadamard matrices

Skew- Hadamard matrices are known to exist for the order 2^n, where $n>0$ is an integer; order $4t$, where $4t$-1 is a prime or prime power. For details on existence of skew- Hadamard matrices see Koukouvinos and Stylianou (2008).

Lemma 1: Let **N** be the incidence matrix of a BIB design obtained from the core of a normalized skew- Hadamard matrix of order $4t$. Then

(i) $\mathbf{N}+\mathbf{N}'=(\mathbf{J}-\mathbf{I})_{4t-1}$ (ii) $\mathbf{N}^2+\mathbf{N}=t(\mathbf{J}-\mathbf{I})_{4t-1}$.

Proof: Let **C** be the core of a normalized skew- Hadamard matrix of order $4t$ obtained by deleting first row and first column. Then the diagonal entries of **C** are -1 and

(a) $\mathbf{C}+\mathbf{I}_{4t-1}$ is a skew- symmetric matrix *i.e.* $\mathbf{C}+\mathbf{I}_{4t-1}=-(\mathbf{C}+\mathbf{I}_{4t-1})'$

$$\Rightarrow \mathbf{C}+\mathbf{C}'=-2\mathbf{I}_{4t-1}.$$

(b) $\mathbf{C}\mathbf{C}'=4t\mathbf{I}_{4t-1}-\mathbf{J}_{4t-1}$.

Clearly $\mathbf{N}=(\mathbf{C}+\mathbf{J}_{4t-1})/2$ represents a symmetric $(4t-1, 2t-1, t-1)$ – design and $\mathbf{N}, \mathbf{N}'$ have zeros in diagonals. Then

$$\mathbf{N}+\mathbf{N}'=(\mathbf{C}+\mathbf{C}'+2\mathbf{J}_{4t-1})/2=(\mathbf{J}-\mathbf{I})_{4t-1}$$

$$\begin{aligned}\mathbf{N}^2+\mathbf{N}&=(\mathbf{C}^2+2\mathbf{C}\mathbf{J}_{4t-1}+\mathbf{J}_{4t-1}^2+2\mathbf{C}+2\mathbf{J}_{4t-1})/4\\&=[\mathbf{C}(\mathbf{C}+2\mathbf{I}_{4t-1})-2\mathbf{J}_{4t-1}+(4t-1)\mathbf{J}_{4t-1}+2\mathbf{J}_{4t-1}]/4\\&=(-\mathbf{C}\mathbf{C}'+(4t-1)\mathbf{J}_{4t-1})/4=(-(4t\mathbf{I}_{4t-1}-\mathbf{J}_{4t-1})+(4t-1)\mathbf{J}_{4t-1})/4\\&=t(\mathbf{J}-\mathbf{I})_{4t-1}.\end{aligned}$$

Theorem 2: The existence of a skew- Hadamard matrix of order $4t$ implies the existence of a GD design with parameters

$$v=b=6(4t-1),\ r=k=2(5t-2),\ \lambda_1=5(t-1),\ \lambda_2=2(2t-1),\ m=6,\ n=4t-1. \quad (2)$$

Proof: Let **N** be the incidence matrix of a BIB design obtained from the core of a normalized skew-Hadamard matrix of order $4t$ and **C** be a conference matrix of order 6. Then replacing 0 by $\mathbf{I}_{4t-1}$, 1 by **N** and –1 by $\mathbf{N}'$ in **C** we obtain a (0, 1) - matrix

$$\mathbf{M}=\begin{pmatrix}\mathbf{I}_{4t-1}&\mathbf{N}&\mathbf{N}&\mathbf{N}&\mathbf{N}&\mathbf{N}\\\mathbf{N}'&\mathbf{I}_{4t-1}&\mathbf{N}'&\mathbf{N}'&\mathbf{N}&\mathbf{N}\\\mathbf{N}'&\mathbf{N}'&\mathbf{I}_{4t-1}&\mathbf{N}&\mathbf{N}&\mathbf{N}'\\\mathbf{N}'&\mathbf{N}'&\mathbf{N}&\mathbf{I}_{4t-1}&\mathbf{N}'&\mathbf{N}\\\mathbf{N}'&\mathbf{N}&\mathbf{N}&\mathbf{N}'&\mathbf{I}_{4t-1}&\mathbf{N}'\\\mathbf{N}'&\mathbf{N}&\mathbf{N}'&\mathbf{N}&\mathbf{N}'&\mathbf{I}_{4t-1}\end{pmatrix}.$$

Using the relations $\mathbf{NN'} = \mathbf{N'N} = t\mathbf{I_{4t-1}} + (t-1)\mathbf{J_{4t-1}}$and $\mathbf{N^2} + \mathbf{N} = (\mathbf{N'})^2 + \mathbf{N'} = t(\mathbf{J} - \mathbf{I})_{\mathbf{4t-1}}$ one can see that **M** represents the incidence matrix of a GD design with the parameters (2).

Remark 2: For *t*=3 in Theorem 2 we obtain a BIB design with parameters *v*=*b*=66, *r*=*k*=26, λ=10, reported in Hall (1998) as design number 214.

Remark 3: For $\mathbf{N} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ and *t*=1 in the incidence matrix **M** of the Theorem 2 we obtain *SR*72.

Following Dey and Balasubramanian (1991), series 25 (rephrased), p. no. 288:

If there exists a symmetric BIB design with parameters $v' = 4t-1, k' = 2t-1, \lambda' = t-1 (t \geq 1)$ such that the incidence matrix **N** of the BIB design satisfies $\mathbf{N} + \mathbf{N'} = (\mathbf{J} - \mathbf{I})_{v'}$, then there exists a GD design with parameters: $v = b = pv', r = k = pk' + 1, \lambda_1 = pk', \lambda_2 = p\lambda' + 1, m = v', n = p (\geq 2)$.

It is known that the incidence matrix **N** of a BIB design obtained from the core of a normalized skew- Hadamard matrix of order 4t satisfies $\mathbf{N} + \mathbf{N'} = (\mathbf{J} - \mathbf{I})_{\mathbf{4t-1}}$, see Lemma 4 above.

Theorem 3: The existence of a skew- Hadamard matrix of order 4*t* implies the existence of a 2- parameter GD design with parameters

$$v = b = p(4t-1), r = k = p(2t-1) + 1, \lambda_1 = p(2t-1), \lambda_2 = p(t-1) + 1,$$

$$m = 4t-1, n = p, t \geq 1. \quad (3)$$

Proof: Let **N** be the incidence matrix of a BIB design obtained from the core of a normalized skew-Hadamard matrix of order 4*t*. Then using the relations $\mathbf{NN'} = \mathbf{N'N} = t\mathbf{I_{4t-1}} + (t-1)\mathbf{J_{4t-1}}$and $\mathbf{N} + \mathbf{N'} = (\mathbf{J} - \mathbf{I})_{\mathbf{4t-1}}$ it can be verified that $\mathbf{M} = \mathbf{I_p} \otimes (\mathbf{I_{4t-1}} + \mathbf{N}) + (\mathbf{J} - \mathbf{I})_{\mathbf{p}} \otimes \mathbf{N}$ is the incidence matrix of a GD design with parameters (3).

The following Table lists regular GD designs constructed using Theorem 3:

Table 2: RGD from skew- Hadamard matrices

No.	GD: (v, r, k, b, λ_1, λ_2, m, n)	p, t	Reference
1	*R*177: (14, 7, 7, 14, 6, 3,7, 2)	$p = t = 2$	Clatworthy (1973)
2	R206*a:* (21, 10, 10, 21, 9, 4, 7, 3)	$p = 3, t = 2$	Freeman (1976)

Remark 4: Following Theorem 7 of Bush (1979) and Corollary 4.1.1 of Kageyama and Tanaka (1981) we get:

A GD design with parameters

$$v=b=3(4t-1),\ r=k=2t+1,\ \lambda_1=t-1,\ \lambda_2=1,\ m=3,\ n=4t-1. \quad (4)$$

is obtained from the core of a normalized skew- Hadamard matrix.

2.3. From Conference matrices

Symmetric conference matrices are known to exist for orders 2, 4, 6, 10, … and skew-symmetric conference matrices are known to exist for 2, 4, 8, 12,….

Theorem 4: The existence of a conference matrix of order t (≥4) implies the existence of family (A) regular GD designs with parameters

(i) $v = b = 2t$, $r = k = t-1$, $\lambda_1 = 0$, $\lambda_2 = (t-2)/2$, $m = t$, $n = 2$. (5)

(ii) $v = b = 2t$, $r = k = t+1$, $\lambda_1 = 2$, $\lambda_2 = (t+2)/2$, $m = t$, $n = 2$. (6)

Proof: Let **C** be a conference matric of order t (≥4) and $\mathbf{N_1} = (\mathbf{J_t} - \mathbf{I_t} + \mathbf{C})/2$, $\mathbf{N_2} = (\mathbf{J_t} - \mathbf{I_t} - \mathbf{C})/2$ then we claim that $\mathbf{N} = \begin{pmatrix} \mathbf{N_1} & \mathbf{N_2} \\ \mathbf{N_2} & \mathbf{N_1} \end{pmatrix}$ is the incidence matrix of the GD design with parameters (5). We have

$$\mathbf{N_1N_1'} + \mathbf{N_2N_2'} = [(\mathbf{J_t} - \mathbf{I_t} + \mathbf{C})/2]\,[(\mathbf{J_t} - \mathbf{I_t} + \mathbf{C})/2]' + [(\mathbf{J_t} - \mathbf{I_t} - \mathbf{C})/2]\,[(\mathbf{J_t} - \mathbf{I_t} - \mathbf{C})/2]'$$

$$= (t-1)\mathbf{I_t} + [(t-2)/2](\mathbf{J_t} - \mathbf{I_t}).$$

$$\mathbf{N_1} + \mathbf{N_2} = \mathbf{J_t} - \mathbf{I_t} \Longrightarrow (\mathbf{N_1} + \mathbf{N_2})(\mathbf{N_1} + \mathbf{N_2})' = (\mathbf{J_t} - \mathbf{I_t})^2 = (\mathrm{t}-1)\,\mathbf{I_t} + (\mathrm{t}-2)\,(\mathbf{J_t} - \mathbf{I_t})$$

$$\Longrightarrow \mathbf{N_1N_2'} + \mathbf{N_2N_1'} = [(\mathrm{t}-1)\mathbf{I_t} + (\mathrm{t}-2)(\mathbf{J_t} - \mathbf{I_t})] - [\mathbf{N_1N_1'} + \mathbf{N_2N_2'}]$$

$$\Longrightarrow \mathbf{N_1N_2'} + \mathbf{N_2N_1'} = [(t-2)/2](\mathbf{J_t} - \mathbf{I_t})\ .$$

Thus $\mathbf{N_1}$ and $\mathbf{N_2}$ satisfy the conditions given in Dey (1977). Hence N is the incidence matrix of the GD design with parameters (5). The GD design with parameters (6) is complementary of the design with parameters (5).

The following Table lists GD designs obtained using Theorem 4:

Table 3: RGD from Conference Matrices

No.	GD: (v, r, k, b, λ_1, λ_2, m, n)	Source	Reference
1	*R*54: (8, 3, 3, 8, 0, 1, 4, 2)	CM (4)	Clatworthy (1973)
2	*R*144: (12, 5, 5, 12, 0, 2, 6, 2)	CM (6)	Dey (1977)
3	*R*117*a*: (16, 7, 7, 16, 0, 3, 8, 2)	CM (8)	Dey (1977)
4	*R*197*a*: (20, 9, 9, 20, 0, 4, 10, 2)	CM (10)	Dey (1977)

Theorem 5: The existence of a Conference matrix of order t (≥4) and a BIB design with

$v = 2k,\ b,\ r,\ k,\ \lambda$ implies the existence of a GD design with parameters

$$v^* = tv, b^* = tb, r^* = r(t-1), k^* = k(t-1), \lambda_1^* = (t-1)\lambda, \lambda_2^* = r(t-2)/2,$$

$$m = t, n = v. \tag{7}$$

Proof: Let **N** be the incidence matrix of a BIB design with $v = 2k, b, r, k, \lambda$. Then replacing 0 by $\mathbf{O}_t$, 1 by **N** and –1 by $\bar{\mathbf{N}} = \mathbf{J}_{v\times b} - \mathbf{N}$ in a Conference matrix of order t we obtain a GD design with parameters (7).

For $(t-1)\lambda = r(t-2)/2$ in Theorem 5, we obtain:

Corollary 3: The existence of a conference matrix of order $t = 2(r-\lambda)/(r-2\lambda)$; ($t \geq 4$) and a BIB design with $v=2k, b, r, k, \lambda$ implies the existence of a BIB design with parameters

$$v^* = tv, b^* = tb, r^* = r(t-1), k^* = k(t-1), \lambda^* = (t-1)\lambda.$$

Using BIB design (4, 6, 3, 2, 1) and $t = 4$ in Corollary 3 produces *MR*35; and a BIB design (6, 10, 5, 3, 2) and $t = 6$ produces *MR*427. *MRX* denotes design number *X* in Mathon and Rosa (2007). It is not known if these solutions are isomorphic to theirs.

Remark 5: For $\mathbf{N} = \mathbf{I}_2$ in Theorem 5 we obtain Theorem 4 (i).

2.4. From balanced generalized Weighing matrices and generalized Conference matrices

Let $C_n = \{1, \alpha, \alpha^2, \cdots, \alpha^{n-1}\}$ denote a cyclic group of order n and $\boldsymbol{\beta} = circ$ (0 1 0…0) denote a circulant matrix of order n.

Replacing 1 by $\mathbf{I}_n$ and α^i by $\boldsymbol{\beta}^i$ ($1 \leq i \leq n-1$) in *BGWM* ($v, k, \lambda; C_n$) we obtain:

Theorem 6: The existence of a BGWM ($v, k, \lambda; C_n$) implies the existence of a GD with parameters

$$v^* = b^* = vn, r^* = k^* = k, \lambda_1 = 0, \lambda_2 = \lambda/n, m = v, n. \tag{8}$$

Further replacing 0 by $\mathbf{O}_n$, 1 by $\mathbf{I}_n$ and α^i by $\boldsymbol{\beta}^i$ ($1 \leq i \leq n-1$) in *GCM* ($C_n; \lambda$) of order v we obtain:

Theorem 7: The existence of a GCM ($C_n; \lambda$) of order $v = n\lambda+2$ implies the existence of a GD design with parameters

$$v^* = b^* = vn, r^* = k^* = k, \lambda_1 = 0, \lambda_2 = \lambda, m = v, n \tag{9}$$

where k is the number of nonzero entries in each column of *GCM* ($C_n; \lambda$).

The following Table lists GD designs obtained using Theorems 6 and 7:

Table 4: RGD from balanced generalized Weighing matrices and generalized Conference matrices

No.	GD: (v, r, k, b, λ_1, λ_2, m, n)	Source	Reference
1	$R112$: (14, 4, 4, 14, 0, 1, 7, 2)	BGWM (7, 4, 2; C_2)	Clatworthy (1973)
2	$R114$: (15, 4, 4, 15, 0, 1, 5, 3)	BGWM (5, 4, 3; C_3)	Clatworthy (1973)
3	$R180b$: (24, 7, 7, 24, 0, 2, 8, 3)	GCM (C_3; 2), Order = 8	F (1976)
4	$R182b$: (45, 7, 7, 45, 0, 1, 15, 3)	BGWM (15, 7, 3; C_3)	DR (1990)
5	$R191$: (63, 8, 8, 63, 0, 1, 9, 7)	GCM (C_7; 1), Order = 9	Clatworthy (1973)
6	$R200a$: (38, 9, 9, 38, 0, 2, 19, 2)	BGWM (19, 9, 4; C_2)	DR (1990)
7	$R200c$: (40, 9, 9, 40, 0, 2, 10, 4)	BGWM (10, 9, 8; C_4)	DN (1985)

F (1976), DN (1985) and DR (1990) stand for Freeman (1976), Dey and Nigam (1985) and De and Roy (1990) respectively. The balanced generalized Weighing matrices and generalized Conference matrices used in Table 3 may be found in Colbourn and Dinitz (2007).

2.5. From Kronecker Sum of Hadamard matrices and incidence matrices of BIB designs

Theorem 8 given below gives an algebraic representation of Theorem 1.6 of Parihar and Shrivastava (1988).

Theorem 8: The existence of a Hadamard matrix of order 4t and a BIB design with $v = 2k$, b, r, k, λ implies the existence of a SRGD design with parameters

$$v^* = (4t-1)v, b^* = 4tb, r^* = 4tr, k^* = (4t-1)k, \lambda_1^* = 4t\lambda, \lambda_2^* = 2tr, m = 4t-1,$$

$$n = v. \tag{10}$$

Proof: Let $\mathbf{H}^*$ be a $(4t-1) \times 4t$ matrix obtained by deleting the first row of a normalized Hadamard matrix and $\mathbf{N}$ be the incidence matrix of a BIB design with $v = 2k$, b, r, k, λ. Considering Kronecker sum $\mathbf{M} = \mathbf{H}^* \oplus \mathbf{N}$ of $\mathbf{H}^*$ and $\mathbf{N}$. Then under the transformation: $-1 \rightarrow 1$ in $-\overline{\mathbf{N}} = -(\mathbf{J}_{v \times b} - \mathbf{N})$ and $1 \rightarrow 0$, $2 \rightarrow 1$ in $\mathbf{J}_{v \times b} + \mathbf{N}$, it is easy to see that $\mathbf{M}$ represents incidence matrix of a SRGD with parameters (10).

Removing α ($1 \leq \alpha \leq 4t-3$) rows of blocks of the incidence matrix of the design with parameters (10) we obtain:

Corollary 4: There exists a SRGD design with parameters

$$v^* = (4t-\alpha-1)v, b^* = 4tb, r^* = 4tr, k^* = (4t-\alpha-1)k, \lambda_1^* = 4t\lambda, \lambda_2^* = 2tr,$$

$$m = 4t-\alpha-1, n = v. \tag{11}$$

Remark 6: The Corollary 4 unifies the Theorems 1.2, 1.3, 1.4 and 1.5 of Parihar and Shrivastava (1988).

Theorem 9: The existence of a regular Hadamard matrix of order $4t^2$ and a BIB design with v=$2k, b, r, k, \lambda$ implies the existence of a SRGD design with parameters

$$v^* = 4t^2v, b^* = 4t^2b, r^* = 4t^2r, k^* = 4t^2k, \lambda_1^* = 4t^2\lambda, \lambda_2^* = 2t^2r, m = 4t^2, n = v. \tag{12}$$

Proof: Let **H** be a regular Hadamard matrix of order $4t^2$ and **N** be the incidence matrix of a BIB design with $v = 2k, b, r, k, \lambda$. Considering Kronecker sum $\mathbf{M} = \mathbf{H} \oplus \mathbf{N}$ of **H** and **N**. Then under the transformation: $-1 \rightarrow 1$ in $-\overline{\mathbf{N}} = (\mathbf{J}_{v\times b} - \mathbf{N})$ and $1 \rightarrow 0$, $2 \rightarrow 1$ in $\mathbf{J}_{v\times b} + \mathbf{N}$, it is easy to see that **M** represents incidence matrix of a SRGD with parameters (12).

For $\mathbf{N} = \mathbf{I_2}$ in Theorem 9 we obtain:

Corollary 5: There exists a resolvable SRGD design with parameters

$$v^* = b^* = 8t^2, r^* = k^* = 4t^2, \lambda_1^* = 0, \lambda_2^* = 2t^2, m = 4t^2, n = 2. \tag{13}$$

Removing α ($1 \leq \alpha \leq 4t^2$-2) rows of blocks of the incidence matrix of design with parameters (12) we obtain:

Corollary 6: There exists a SRGD design with parameters

$$v^* = (4t^2 - \alpha)v, b^* = 4t^2b, r^* = 4t^2r, k^* = (4t^2 - \alpha)k, \lambda_1^* = 4t^2\lambda, \lambda_2^* = 2t^2r,$$

$$m = 4t^2 - \alpha, n = v. \tag{14}$$

Remark 7: This theorem is generalization and algebraic representation of the Theorem 2.2 of Bhagwandas *et al.* (1985). For $t = 1$ in Theorem 9 we obtain Theorem 2.2 of Bhagwandas *et al.* (1985).

Theorem 10: The existence of a Hadamard matrix of order $2t$ implies the existence of a resolvable SRGD design with parameters

$$D_i: v^* = b^* = 2^{i+2}t, r = k = 2^{i+1}t, \lambda_1 = 0, \lambda_2 = 2^it, m = 2^{i+1}t, n = 2\ (i \geq 0). \tag{15}$$

Proof: Kadowaki and Kageyama (2009, Theorem 3.3.4) constructed a resolvable SRGD design with parameters

$$D_0: v = b = 4t, r = k = 2t, \lambda_1 = 0, \lambda_2 = t, m = 2t, n = 2. \tag{16}$$

Let $\mathbf{N_0}$ be incidence matrix of a SRGD design D_0 with parameters (16). Considering Kronecker sum $\mathbf{N_i} = \mathbf{H_2} \oplus \mathbf{N_{i-1}} (i \geq 1)$ of $\mathbf{H_2}$ and $\mathbf{N_{i-1}}$, where $\mathbf{H_2}$ is a Hadamard matrix of order 2 and $\mathbf{N_{i-1}}$ represents the incidence matrix of a SRGD design with parameters

$$v' = b' = 2^{i+1}t, r' = k' = 2^it, \lambda_1 = 0, \lambda_2 = 2^{i-1}t, m = 2^it, n = 2\ (i \geq 1).$$

Then under the transformation: $-1 \rightarrow 1$ in $-(\mathbf{J}_{v'\times b'} - \mathbf{N}_{i-1})$ and $1 \rightarrow 0$, $2 \rightarrow 1$ in $\mathbf{J}_{v'\times b'} + \mathbf{N}_{i-1}$, it is easy to see that $\mathbf{N_i}$ represents incidence matrix of a SRGD with parameters (15).

Remark 8: This Theorem generalizes the Theorem 3.3.4 of Kadowaki and Kageyama (2009) and Theorem 2.1 of Sinha (1991*b*). For $i = 0$ we obtain Theorem 3.3.4 of Kadowaki and Kageyama (2009) and for $i = 1$ and 2 we obtain series 2.1 and 2.2 respectively of Sinha (1991*b*).

3. A Catalogue of Group Divisible Designs

In the process of present investigation, some GD designs scattered in literature are found; and those not found in Clatworthy (1973) and Sinha (1991*a*) are catalogued below, to make them available at one place for the convenience of researchers, looking for GD designs in the practical range of $r, k \leq 10$.

Table 5: A Catalogue of GD designs

No.	GD: (v, r, k, b, λ_1, λ_2, m, n)	Reference
1	*SR*109*a*: (50, 10, 10, 50, 0, 2, 10, 5)	GD (1995)
2	*R*208*b*: (49, 10, 10, 49, 1, 2, 7, 7)	S (1991)
3	*R*206*b*: (21, 10, 10, 21, 8, 3, 3, 7)	MD (1995)
4	*R*200*b*: (39, 9, 9, 39, 0, 2, 13, 3)	SS (2021)
5	*R*198*a*: (24, 9, 9, 24, 6, 3, 12, 2)	DK (1993)
6	*R*200*d*: (45, 9, 9, 45, 3, 1, 3, 15)	t = 4 in (4), Bush (1979)
7*	*SR*103*a*: (45, 10, 9, 50, 0, 2, 9, 5)	GD (1995)
8*	*SR*95*a*: (40, 10, 8, 50, 0, 2, 8, 5)	GD (1995)
9*	*SR*86*a*: (35, 10, 7, 50, 0, 2, 7, 5)	GD (1995)

S (1991), DK (1993), MD (1995), GD (1995) and SS (2021) stand for Sastry (1991), Duan and Kageyama (1993), Midha and Dey (1995), Ghosh and Divecha (1995) and Saurabh and Sinha (2020) respectively. The design numbers 1, 7, 8 and 9 were later on also reported by Kadowaki and Kageyama (2009).

*Design No. 7 of Table 5 is obtained by deleting the set of treatments 46, 47, 48, 49, 50 from design No. 1; design No. 8 is obtained by deleting the set of treatments 41, 42, 43, 44, 45 from design No. 7; and design No. 9 is obtained by deleting the set of treatments 36, 37, 38, 39, 40 from design No. 8.

As a special case having t=4, in Remark 4, we get a regular group divisible design with parameters: $v = b = 45, r = k = 9, \lambda_1 = 3, \lambda_2 = 1, m = 3, n = 15$ and the average efficiency E = 0.90. The solution given below is not found elsewhere:

(4 6 7 9 12 14 15 16 31), (1 5 7 9 10 13 15 17 32), (1 2 6 9 10 11 14 18 33),
(2 3 7 10 11 12 15 19 34), (1 3 4 9 11 12 13 20 35), (2 4 5 10 12 13 14 21 36),
(3 5 6 11 13 14 15 22 37), (4 6 7 8 10 11 13 24 39), (1 5 7 8 11 12 14 25 40),
(1 2 6 8 12 13 15 26 41), (1 2 3 4 5 6 7 23 38), (2 3 7 8 9 13 14 27 42),
(1 3 4 8 10 14 15 28 43), (2 4 5 8 9 11 15 29 44), (3 5 6 8 9 10 12 30 45),
(1 19 21 22 24 27 29 30 31), (2 16 20 22 24 25 28 30 32), (3 16 17 21 24 25 26 29 33),
(4 17 18 22 25 26 27 30 34), (5 16 18 19 24 26 27 28 35), (6 17 19 20 25 27 28 29 36),
(7 18 20 21 26 28 29 30 37), (8 16 17 18 19 20 21 22 38), (9 19 21 22 23 25 26 28 39),
(10 16 20 22 23 26 27 29 40), (11 16 17 21 23 27 28 30 41), (12 17 18 22 23 24 28 29 42),
(13 16 18 19 23 25 29 30 43), (14 17 19 20 23 24 26 30 44), (15 18 20 21 23 24 25 27 45),
(1 16 34 36 37 39 42 44 45), (2 17 31 35 37 39 40 43 45), (3 18 31 32 36 39 40 41 44),
(4 19 32 33 37 40 41 42 45), (5 20 31 33 34 39 41 42 43), (6 21 32 34 35 40 42 43 44),

(7 22 33 35 36 41 43 44 45), (8 23 31 32 33 34 35 36 37), (9 24 34 36 37 38 40 41 43), (10 25 31 35 37 38 41 42 44), (11 26 31 32 36 38 42 43 45), (12 27 32 33 37 38 39 43 44), (13 28 31 33 34 38 40 44 45), (14 29 32 34 35 38 39 41 45), (15 30 33 35 36 38 39 40 42).

The GD scheme is defined by the array: 1 2 3 4 … 15
16 17 18 19… 30
31 32 33 34 …45.

Acknowledgement

The authors are grateful to an anonymous reviewer and Dr. G. M. Saha for their kind suggestions leading to considerable improvements in the paper, especially the presentation of the contents.

References

Arasu, K. T., Heamers, W. H., Jungnickel, D. and Pott, A. (1991). Matrix constructions of divisible designs. *Linear Algebra and its Applications*, **153**, 123-133.

Bhagwandas, Banerjee, S. and Kageyama, S. (1985). Patterned constructions of partially balanced incomplete block designs. *Communications in Statistics: Theory and Methods*, **14**(**6**), 1259-1267.

Bush, K. A. (1979). Families of Hadamard group divisible designs. *Journal of Statistical Planning and Inference*, **3**, 387- 394.

Cheng, S. (1995). Matrix constructions of family (A) group divisible designs. *Australasian Jourrnal of Combinatorics*, **11**, 223-231.

Clatworthy, W. H. (1973). Tables of two-associate-class partially balanced designs. *National Bureau of Standards* (*U.S.*), *Applied Mathematics*, Series **63**.

Colbourn, C. J. and Dinitz, J. H. (Eds.) (2007). Balanced generalized weighing matrices and conference matrices. In *Handbook of Combinatorial Designs*, Second Edition, Chapman and Hall/ CRC, New York.

Crnkovic, D. (2006). A series of regular Hadamard matrices. *Designs, Codes and Cryptography,* **39**, 247-251.

De, A. K. and Roy, B. K. (1990). Computer construction of some group divisible designs. *Sankhya*, **B52**, 82-92.

Dey, A. (1977). Construction of regular group divisible designs. *Biometrika*, **64**, 647-649.

Dey, A. (1986). *Theory of Block Designs*. Wiley Eastern, New Delhi

Dey, A. (2010). *Incomplete Block Designs*. Hindustan Book Agency, New Delhi.

Dey, A. and Balasubramanian, K. (1991). Construction of some families of group divisible designs. *Utilitas Mathematica*, **40**, 283-290.

Dey, A. and Nigam, A. K. (1985). Constructions of group divisible designs. *Journal of the Indian Society of Agricultural Statistics*, **37**,163-166.

Duan, X. and Kageyama, S. (1993). Constructions of nested group divisible designs. *Statistics and Probability Letters*, **18**, 41-48.

Freeman, G. H. (1976). A cyclic method of constructing regular group divisible incomplete block designs. *Biometrika*, **63**, 555- 558.

Ghosh, D. K. and Divecha, J. (1995). Some new semi-regular GD designs. *Sankhya*, **B57**, 453-455

Hall, Marshal Jr. (1998). *Combinatorial Theory*. John Wiley, New York.

Kadowaki, S. and Kageyama, S. (2009). Existence of affine α-resolvable PBIB designs with some constructions. *Hiroshima Mathematical Journal*, **39,** 293-326.

Kageyama, S. and Mohan, R. N. (1983). On μ- resolvable BIB designs. *Discrete Mathematics*, **45**, 113-122.

Kageyama, S. and Tanaka, T. (1981). Some families of group divisible designs. *Journal of Combinatorial Theory*, **A15**, 1-10.

Koukouvinos, C. and Stylianou, S. (2008). On skew- Hadamard matrices. *Discrete Mathematics*, **308**, 2723-2731.

Lampio, H. J. P. (2015). Classification of difference matrices and complex Hadamard matrices. *Aalto University Publication Series*, Doctoral Dissertations 177/2015, Finland.

Mathon, R. and Rosa, A. (2007). 2- (v, k, λ) designs of small order. *The CRC Handbook of Combinatorial Designs*, (Eds.): C. J. Colbourn and J. H. Dinitz, CRC Press, 35-58.

Midha, C. K. and Dey, A. (1995). Cyclic group divisible designs. *Calcutta Statistical Association Bulletin*, **45**, 179-180.

Parihar, J. S. and Shrivastava, R. (1988). Methods of construction of group divisible designs. *Journal of Statistical Planning and Inference*, **18**, 399-404.

Raghavarao, D. (1971). *Constructions and Combinatorial Problems in Design of Experiments*. John Wiley, New York.

Raghavarao, D. and Padgett, Lakshmi V. (2005). Block designs. Analysis, combinatorics and applications. *Series on Applied Mathematics*, **17**, World Scientific, Singapore.

Sastry, D. V. S. (1991). A series of group divisible designs. *Communications in Statistics-Theory and Methods*, **20**, 1677-1682.

Saurabh, S. and Sinha, K. (2021). Algebraic constructions of group divisible designs, Under preparation.

Sinha, K. (1991*a*). A list of new group divisible designs. *Journal of Research of the National Institute of Standards and Technology*, **96**, 1-3.

Sinha, K. (1991*b*). Construction of semi-regular group divisible designs. *Sankhya*, **B53**, 229-232.

Tonchev, V. D. (2009). Generalized weighing matrices and self- orthogonal codes. *Discrete Mathematics*, **309**, 4697-4699.

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 141–148

Constant Block-Sum Two-Associate Class Group Divisible Designs

Sudhir Gupta
Department of Statistics
Northern Illinois University, DeKalb, Illinois, USA

Received: 19 September 2020; Revised: 07 December 2020; Accepted: 11 December 2020

Abstract

It is shown that classes of semi-regular and regular group divisible designs do not lead to constant block-sum designs. Construction of constant block-sum designs using singular group divisible designs is discussed in general. For a given singular group divisible design, the construction method is shown to provide a large number of distinct constant block-sum designs. Construction of constant block-sum designs for equispaced treatment levels is also discussed.

Key words: Balanced incomplete block design; Eigenvalue; Eigenvector; Partially balanced.

1. Introduction

Recently Khattree (2018a,b) discussed the concept of constant block-sum designs for quantitative treatment levels. In these designs, the sum of the treatment levels in each block is constant. Non-existence of constant block-sum balanced incomplete designs was established by Khattree (2018a, 2020). Several methods of construction have been presented by Khattree (2019). A general approach to determine whether or not a design can be transformed into a constant block-sum design and its construction if it exists has been developed in Khattree (2020). Bansal and Garg (2020) derived some conditions for existence of partially balanced constant block-sum designs and gave further combinatorial methods of construction. Khattree (2020) discussed some individual examples, including two-associate class group divisible (GD) designs. The purpose of this note is to present results with respect to the property of constant block-sum that apply to the whole class of GD designs. Non-existence of constant block-sum designs is established for classes of semi-regular and regular GD designs. Construction of constant block-sum singular GD designs is discussed in general. Existence of a large number of distinct constant block-sum solutions for a given singular GD design is illustrated with the help of an example. Singular GD constant block-sum designs for equispaced treatment levels are discussed in Section 3.

2. Group Divisible Designs

In two-associate class GD designs, $v = m_1 m_2$ treatments are arranged in m_1 groups of m_2 treatments each. Let the treatments be coded as 1, 2, $\cdots$, $m_1 m_2$. Then it is convenient

Corresponding Author: Sudhir Gupta
Email: sudhirgupta.stat@gmail.com

to form the groups as:

Table 1

1	2	.	.	.	m_2
m_1+1	m_1+2	.	.	.	$2m_2$
			.		
			.		
			.		
$m_2(m_1-1)+1$	$m_2(m_1-1)+2$	.	.	.	m_1m_2

The treatments are first associates if they belong to the same group and second associates otherwise. The parameters of a GD design are $v = m_1m_2$, b, r, k, λ_1, λ_2, m_1, m_2, where the symbols have their standard meaning, see Raghavarao (1971) or Dey (1986) for details. Let

$$\boldsymbol{A} = \boldsymbol{N}\boldsymbol{N}' - \frac{rk}{v}\boldsymbol{J}_v$$

where $\boldsymbol{N}$ is the $v \times b$ incidence matrix and $\boldsymbol{J}_t$ denotes a square matrix of one's of size t. Note that $\mathbf{1}_v$, a vector of ones of size $v \times 1$, is an eigenvector of $\boldsymbol{A}$ corresponding to a zero eigenvalue.

For an equireplicate partially balanced design, Khattree (2020) showed that a necessary condition for existence of a constant block-sum design is that

$$\boldsymbol{A}\boldsymbol{w} = \mathbf{0}$$

where $\boldsymbol{w} \neq \mathbf{1}_v$ is an eigenvector of $\boldsymbol{A}$ corresponding to a zero eigenvalue. Note that this is not a sufficient condition, as it is possible that a vector $\boldsymbol{w}$ satisfying the necessary condition does not have all of its elements different from each other. If the v elements of $\boldsymbol{w}$ are all different from each other, they are taken as v treatment levels to yield a constant block-sum design.

As $\boldsymbol{A}$ and $\boldsymbol{N}\boldsymbol{N}'$ are symmetric matrices, they both admit their spectral decompositions. Also, $\boldsymbol{N}\boldsymbol{N}'\mathbf{1}_v = rk\mathbf{1}_v$, so it can be easily seen that if $\boldsymbol{w} \neq \mathbf{1}_v$ is an eigenvector of $\boldsymbol{A}$ corresponding to a zero eigenvalue then it is also an eigenvector of $\boldsymbol{N}\boldsymbol{N}'$ corresponding to a zero eigenvalue and vice versa. Thus, equivalently, we have the following theorem.

Theorem 1: A necessary condition for the existence of a constant block-sum design is that $\boldsymbol{N}\boldsymbol{N}'$ is singular.

Remark 1: Singularity of $\boldsymbol{N}\boldsymbol{N}'$ in turn implies that the rows of $\boldsymbol{N}$ are not linearly independent.

Remark 2: Statement of Remark 1 is automatically satisfied if $v > b$, since $\text{Rank}(\boldsymbol{N}) \leq \min(v, b) < v$.

The structure of $\boldsymbol{N}\boldsymbol{N}'$ for GD designs as given below and its eigenvectors and eigenvalues given in Lemma 1 are well known, see *e.g.* Nigam, Puri and Gupta (1988).

$$\boldsymbol{NN'} = \begin{bmatrix} (r-\lambda_1)\boldsymbol{I}_{m_2}+\lambda_1\boldsymbol{J}_{m_2} & \lambda_2\boldsymbol{J}_{m_2} & \lambda_2\boldsymbol{J}_{m_2} & \cdots & \lambda_2\boldsymbol{J}_{m_2} \\ \lambda_2\boldsymbol{J}_{m_2} & (r-\lambda_1)\boldsymbol{I}_{m_2}+\lambda_1\boldsymbol{J}_{m_2} & \lambda_2\boldsymbol{J}_{m_2} & \cdots & \lambda_2\boldsymbol{J}_{m_2} \\ \vdots & \vdots & \vdots & \vdots & \\ \vdots & \vdots & \vdots & \vdots & \\ \lambda_2\boldsymbol{J}_{m_2} & \lambda_2\boldsymbol{J}_{m_2} & \lambda_2\boldsymbol{J}_{m_2} & \cdots & (r-\lambda_1)\boldsymbol{I}_{m_2}+\lambda_1\boldsymbol{J}_{m_2} \end{bmatrix}$$

$$= (r-\lambda_1)\boldsymbol{I}_{m_1}\otimes\boldsymbol{I}_{m_2}+(\lambda_1-\lambda_2)\boldsymbol{I}_{m_1}\otimes\boldsymbol{J}_{m_2}+\lambda_2\boldsymbol{J}_{m_1}\otimes\boldsymbol{J}_{m_2}$$

where $\boldsymbol{I}_q$ and $\boldsymbol{J}_q$ denote respectively an identity matrix and a square matrix of one's, both of order q, and $\otimes$ is the (right) kronecker product. Let $\boldsymbol{u}_{1i}$, $i = 1, 2, \cdots, (m_1-1)$ be orthonormal column vectors of size m_1 each, such that $\boldsymbol{u}'_{1i}\mathbf{1}_{m_1} = 0$, $\boldsymbol{u}'_{1i}\boldsymbol{u}_{1i} = 1$, and $\boldsymbol{u}'_{1i}\boldsymbol{u}_{1i_1} = 0$, $i \neq i_1 = 1, 2, \cdots, (m_1-1)$. Similarly, let $\boldsymbol{u}_{2j}$, $j = 1, 2, \cdots, (m_2-1)$ be orthonormal column vectors of size m_2 each, such that $\boldsymbol{u}'_{2j}\mathbf{1}_{m_2} = 0$, $\boldsymbol{u}'_{2j}\boldsymbol{u}_{2j} = 1$, and $\boldsymbol{u}'_{2j}\boldsymbol{u}_{2j_1} = 0$, $j \neq j_1 = 1, 2, \cdots, (m_2-1)$. Without loss of generality, we take normalized orthogonal polynomial contrasts as $\boldsymbol{u}_{1i}$ and $\boldsymbol{u}_{2j}$, $i = 1, 2, \cdots, (m_1-1)$,$j = 1, 2, \cdots, (m_2-1)$.

Lemma 1:

(*a*) $\boldsymbol{w}_{1i} = \boldsymbol{u}_{1i}\otimes\mathbf{1}_{m_2}$, $i = 1, 2, \cdots, (m_1-1)$ constitute a set of (m_1-1) eigenvectors of $\boldsymbol{NN'}$ corresponding to the constant eigenvalue of $\theta_1 = (rk - v\lambda_2)$,
(*b*) $\boldsymbol{w}_{2j} = \mathbf{1}_{m_1}\otimes\boldsymbol{u}_{2j}$, $\boldsymbol{w}_{12ij} = \boldsymbol{u}_{1i}\otimes\boldsymbol{u}_{2j}$, $i = 1, 2, \cdots, (m_1-1)$; $j = 1, 2, \cdots, (m_2-1)$ constitute a set of $m_1(m_2-1)$ eigenvectors of $\boldsymbol{NN'}$ corresponding to the constant eigenvalue of $\theta_2 = (r-\lambda_1)$,
(*c*) $\mathbf{1}_{m_1}\otimes\mathbf{1}_{m_2}$ is an eigenvector of $\boldsymbol{NN'}$ corresponding to the eigenvalue of $\theta_0 = rk$, and
(*d*) the m_1m_2 eigenvectors of $\boldsymbol{NN'}$ in (a), (b), and (c) are mutually orthogonal.

GD designs are called singular if $r = \lambda_1$, semi-regular if $r > \lambda_1$ and $rk = v\lambda_2$, and regular if $r > \lambda_1$ and $rk > v\lambda_2$. Let us first consider the class of semi-regular GD (SRGD) designs. It can be seen that $\theta_1 = 0$ and $\theta_2 > 0$ for SRGD designs. From Lemma 1, the following (m_1-1) eigenvectors of $\boldsymbol{NN'}$ correspond to an eigenvalue of zero as required in Theorem 1.

$$\boldsymbol{w}_{1i} = \boldsymbol{u}_{1i}\otimes\mathbf{1}_{m_2}, \qquad i = 1, 2, \cdots, m_1-1 \ .$$

However, it is easily seen that none of these eigenvectors on its own satisfies the requirement that all of its v elements be different from each other. Note that a linear combination of these m_1-1 eigenvectors is also an eigenvector of $\boldsymbol{NN'}$ corresponding to zero eigenvalue. So, let us consider the following general linear combination $\boldsymbol{t}_{1w}$, where c_i, $i = 1, 2, \cdots, (m_1-1)$ are some constants.

$$\begin{aligned} \boldsymbol{t}'_{1w} &= \sum_{i=1}^{m_1-1} c_i\left(\boldsymbol{u}'_{1i}\otimes\mathbf{1}'_{m_2}\right) \\ &= \left[\left(\sum_{i=1}^{m_1-1} c_iu_{1i1}\right)\mathbf{1}'_{m_2} \quad \left(\sum_{i=1}^{m_1-1} c_iu_{1i2}\right)\mathbf{1}'_{m_2} \quad \cdots \quad \left(\sum_{i=1}^{m_1-1} c_iu_{1im_1}\right)\mathbf{1}'_{m_2}\right] \end{aligned} \tag{1}$$

where $\boldsymbol{u}'_{1i} = (u_{1i1}\ u_{1i2}\ \cdots\ u_{1im_1})\,,\ i = 1, 2, \cdots, (m_1 - 1)$. It is clear from equation (1) that there does not exist a linear combination $\boldsymbol{t}_{1w}$ such that all of its $v = m_1 m_2$ elements are different from each other. Thus we can state the following result.

Theorem 2: There does not exist a constant block-sum semi-regular GD design.

Next, turning attention to the class of regular GD designs, note that both of the eigenvalues θ_1 and θ_2 of $\boldsymbol{N}\boldsymbol{N}'$ for these designs are greater than zero. So, an eigenvector $\boldsymbol{w}$ per the necessary condition of Theorem 1 does not exist for the class of regular GD designs. Thus we have the following.

Theorem 3: There does not exist a regular GD constant block-sum design.

Finally, we now consider singular GD (SGD) designs for which the eigenvalue $\theta_2 = r - \lambda_1 = 0$. From Lemma 1, the following $m_1(m_2 - 1)$ eigenvectors satisfy the necessary condition of Theorem 1 for existence of constant block-sum designs.

$$\begin{aligned} \boldsymbol{w}_{2j} &= \mathbf{1}_{m_1} \otimes \boldsymbol{u}_{2j}, \quad j = 1, 2, \cdots (m_2 - 1)\,, \\ \boldsymbol{w}_{12ij} &= \boldsymbol{u}_{1i} \otimes \boldsymbol{u}_{2j}, \quad i = 1, 2, \cdots, (m_1 - 1)\,;\ j = 1, 2, \cdots, (m_2 - 1) \end{aligned}$$

None of these $m_1(m_2 - 1)$ eigenvectors on its own satisfies the requirement that all of its $m_1 m_2$ elements be different from each other. So, we explore a linear combination $\boldsymbol{t}_{2w}$ of the $m_1(m_2 - 1)$ eigenvectors, that is also an eigenvector of $\boldsymbol{N}\boldsymbol{N}'$ with zero eigenvalue, such that its $m_1 m_2$ elements are different from each other.

$$\boldsymbol{t}_{2w} = \sum_{j=1}^{m_2-1} c_{2j}\boldsymbol{w}_{2j} + \sum_{i=1}^{m_1-1}\sum_{j=1}^{m_2-1} c_{12ij}\boldsymbol{w}_{12ij} \tag{2}$$

where $c_{1j},\ c_{12ij},\ i = 1, 2, \cdots, (m_1 - 1)\,;\ j = 1, 2, \cdots, (m_2 - 1)$ are some constants. For illustration, we consider the following example.

Example 1: Consider the SGD design S21 in Clatworthy (1973) tables with parameters $v = 9,\ b = 3,\ r = 2,\ k = 6,\ \lambda_1 = 2,\ \lambda_2 = 1,\ m_1 = m_2 = 3$:

Block No.	Block contents					
1	1	2	3	4	5	6
2	1	2	3	7	8	9
3	4	5	6	7	8	9

Here, $m_1(m_2 - 1) = 6$ orthonormal eigenvectors of $\boldsymbol{N}\boldsymbol{N}'$ corresponding to zero eigenvalue are as follows.

$$\begin{bmatrix} \boldsymbol{w}'_{21} \\ \boldsymbol{w}'_{22} \\ \boldsymbol{w}'_{1211} \\ \boldsymbol{w}'_{1212} \\ \boldsymbol{w}'_{1221} \\ \boldsymbol{w}'_{1222} \end{bmatrix} = \begin{bmatrix} (-1 & 0 & +1 & -1 & 0 & +1 & -1 & 0 & +1)/\sqrt{6} \\ (+1 & -2 & +1 & +1 & -2 & +1 & +1 & -2 & +1)/3\sqrt{2} \\ (+1 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & +1)/2 \\ (-1 & +2 & -1 & 0 & 0 & 0 & +1 & -2 & +1)/2\sqrt{3} \\ (-1 & 0 & +1 & +2 & 0 & -2 & -1 & 0 & +1)/2\sqrt{3} \\ (+1 & -2 & +1 & -2 & +4 & -2 & +1 & -2 & +1)/6 \end{bmatrix} \tag{3}$$

By taking $c_{21} = -0.03$, $c_{22} = 0.50$, $c_{1211} = -0.42$, $c_{1212} = 0.61$, $c_{1221} = -0.90$, and $c_{1222} = -0.43$, in equation (2) and using (3) we get:

$$\boldsymbol{t}'_{2w} = (-0.0679 \;\; 0.2598 \;\; -0.1920 \;\; -0.2462 \;\; -0.5224 \;\; 0.7686 \;\; 0.7043 \;\; -0.4446 \;\; -0.2598)$$

Adding a same constant value to all the elements of $\boldsymbol{t}_{2w}$ does not break the constant block-sum property. The elements of $\boldsymbol{t}^*_{2w}$ given below, obtained by adding $c_0 = 0.70$ to the elements of $\boldsymbol{t}_{2w}$, can be taken as treatment levels for constant block-sum property.

$$\boldsymbol{t}^{*\prime}_{2w} = (0.6321 \;\; 0.9598 \;\; 0.5080 \;\; 0.4538 \;\; 0.1776 \;\; 1.4686 \;\; 1.4043 \;\; 0.2554 \;\; 0.4402).$$

As a matter of fact, a very large number of solutions for $\boldsymbol{t}^*_{2w}$ can be found by varying the values of the six coefficients c_{21}, c_{22}, c_{1211}, c_{1212}, c_{1221}, c_{1222} of the linear combination $\boldsymbol{t}_{2w}$. Any set of six values of these coefficients that results in all the elements of $\boldsymbol{t}_{2w}$ to be different from each other would satisfy the constant block-sum property. Table 2 lists 5 other solutions for the treatment levels vector $\boldsymbol{t}^*_{2w}$ obtained by trial and error. The corresponding values of the six coefficients are listed in Table 3, where c_0 is the constant value added to the elements of $\boldsymbol{t}_{2w}$ to obtain $\boldsymbol{t}^*_{2w}$. Many more solutions can be found simply by taking other values for the coefficients such that all the elements of $\boldsymbol{t}_{2w}$ are different from each other.

Table 2: Further solutions for Example 1

$\boldsymbol{t}^{*\prime}_{2w}$ No.	$\boldsymbol{t}^{*\prime}_{2w}$								
1	0.7980	0.3685	1.8336	0.8232	1.1953	0.9815	0.8612	0.0221	2.1168
2	0.7480	0.4685	1.7836	0.9232	0.9953	1.0815	0.8112	0.1221	2.0668
3	0.6980	0.5685	1.7336	1.0232	0.7953	1.1815	0.7612	0.2221	2.0168
4	0.8771	0.6185	1.5044	1.2773	0.6953	1.0274	0.9403	0.2721	1.7876
5	1.5412	1.1447	0.9140	1.5833	1.9138	0.1029	0.7722	0.6829	2.1450

Table 3: Coefficient values for $\boldsymbol{t}^{*\prime}_{2w}$ listed in Table 2

$\boldsymbol{t}^{*\prime}_{2w}$ No.	c_{21}	c_{22}	c_{1211}	c_{1212}	c_{1221}	c_{1222}	c_0
1	1.00	1.00	0.11	0.30	0.57	1.00	1.00
2	1.00	1.00	0.11	0.30	0.57	0.70	1.00
3	1.00	1.00	0.11	0.30	0.57	0.40	1.00
4	0.50	1.00	0.11	0.30	0.57	0.25	1.00
5	-0.30	-0.10	1.00	0.40	1.07	1.00	1.20

Remark 3: For comparing treatments with respect to their effects, it is natural that treatment levels will be determined by subject matter specialists based on the objectives of their study. Example 1 illustrates the conundrum the experimenter is confronted with. What if none of the solutions illustrated in the example is a good choice of treatment levels for the study objectives? Note that for a $\boldsymbol{t}^*_{2w}$ of Table 2, $f_1\boldsymbol{t}^*_{2w} + f_2\mathbf{1}_9$ also satisfies the property of constant block-sum, where $f_1 > 0$ is a constant and f_2 is another constant such that all the treatment levels are greater than zero. Of course, we can also include more solutions

in Table 2 and hope that one of the solutions meets the study objectives. However, a systematic, perhaps algebraic, method of deriving possible solutions for unequally spaced and equispaced treatment levels in general deserves further research. Khattree (2019) has provided a detailed discussion on optimizing constant block-sum and nearly constant block-sum designs.

Sometimes a choice of $\boldsymbol{u}_{1i}$'s and $\boldsymbol{u}_{2i}$'s other than the orthogonal polynomial contrasts may yield a analytical solution directly without the need of forming linear combinations of eigenvectors. For instance, suppose in Example 1 we take $\boldsymbol{u}'_{11} = (1,\ 2,\ -3)/\sqrt{14}$, $\boldsymbol{u}'_{12} = (1,\ -1.25,\ -0.5)/\sqrt{2.8125}$, $\boldsymbol{u}'_{21} = (-5,\ 4,\ 1)/\sqrt{42}$, $\boldsymbol{u}'_{22} = (1,\ 2,\ -3)/\sqrt{14}$. Then, using Lemma 1,

$$\boldsymbol{w}'_{1211} = \boldsymbol{u}_{11} \otimes \boldsymbol{u}_{21} = (-5\ \ 4\ \ 1\ \ -10\ \ 8\ \ 2\ \ 15\ \ -12\ \ -3)/\sqrt{588} \tag{4}$$

is an eigenvector of $\boldsymbol{N}\boldsymbol{N}'$ with zero eigenvalue having all of its elements different from each other. Thus,

$$\boldsymbol{t}^{*\prime}_{2w} = f_1\,(-5\ \ 4\ \ 1\ \ -10\ \ 8\ \ 2\ \ 15\ \ -12\ \ -3) + c_0\mathbf{1}'_9\,,$$

where $f_1 > 0$ and $c_0 > 12$ are some constants, satisfies the property of constant block-sum. The constants f_1 and c_0 can be chosen appropriately to suit experimenter's requirements with respect to the magnitude of treatment levels.

3. Equispaced Treatment Levels

The general approach illustrated in the previous section shows many possibilities for constant block-sum designs with unequally spaced treatment levels. However, if equispaced treatment levels are desired, SGD designs based on BIB designs in particular afford a solution directly without making use of the eigenvectors of $\boldsymbol{N}\boldsymbol{N}'$. Consider a BIB design D with parameters $v_0 = m_1$, b_0, r_0, k_0, λ_0, with treatments coded as 1, 2, $\cdots$, m_1. Let D_{SGD} denote the design obtained by replacing treatment i in the BIB design by m_2 treatments $(i-1)m_2+1$, $(i-1)m_2+1$, $\cdots$, im_2, $i = 1, 2, \cdots, m_1$. Then D_{SGD} is a SGD design (Bose and Connor, (1952)) with parameters $v = m_1m_2$, $b = b_0$, $r = r_0$, $k = m_2k_0$, $\lambda_1 = r$, $\lambda_2 = \lambda_0$, m_1, m_2, with m_1 groups of treatments as given in Table 1. Let $\boldsymbol{T}$ be the vector of treatments given by,

$$\boldsymbol{T} = (1,\ 2,\ \cdots,\ m_2,\ m_2+1,\ m_2+2,\ \cdots,\ 2m_2, \cdots,\ m_1m_2)'\ . \tag{5}$$

Now suppose it is desired to transform SGD design D_{SGD} into a constant block-sum design for m_1m_2 equispaced treatment levels ℓ_i, $i = 1, 2, \cdots, m_1m_2$, where $\ell_i = \ell_1 + (i-1)d$, $d = \ell_i - \ell_{i-1}$, $i = 2, 3, \cdots, m_1m_2.$, ℓ_1 being the lowest dose or treatment level. Let the vector of equispaced treatment levels can be written as,

$$\boldsymbol{T}_\ell = \ell_1\mathbf{1}_v + d\,\{0,\ 1, 2, \cdots, (m_1m_2-1)\}'\ . \tag{6}$$

In fact we only need to work with $\boldsymbol{T}_{\ell 0}$ as defined below, since $\boldsymbol{T}_\ell = \ell_1\mathbf{1}_v + d\boldsymbol{T}_{\ell 0}$,

$$\boldsymbol{T}_{\ell 0} = \{0,\ 1,\ 2,\ \cdots,\ (m_1m_2-1)\}'\ . \tag{7}$$

The sum of the m_1m_2 elements of $\boldsymbol{T}_{\ell 0}$, say ℓ_{SUM}, is then given by

$$\ell_{SUM} \;=\; \boldsymbol{T}'_{\ell 0}\mathbf{1}_v \;=\; \{m_1m_2\,(m_1m_2-1)\}\,/2\,.$$

Further suppose that it is possible to partition the $v = m_1m_2$ elements of $\boldsymbol{T}_{\ell 0}$ into m_1 groups of size m_2 each such that the sum of the m_2 elements within all the m_1 groups is equal to each other. Clearly, then the sum of m_2 elements in each group is equal to ℓ_{SUM}/m_1. Let the ith group, say G_i be denoted by,

$$\boldsymbol{G}_i \;=\; \left\{\ell^*_{\{(i-1)m_2+1\}},\; \ell^*_{\{(i-1)m_2+2\}},\; \cdots\;,\; \ell^*_{im_2}\right\}\,,$$

$$\sum_{j=1}^{m_2}\ell^*_{\{(i-1)m_2+j\}} \;=\; \frac{\ell_{SUM}}{m_1} \;=\; \frac{m_2\,(m_1m_2-1)}{2}, \qquad\qquad i \;=\; 1,\,2,\,\cdots,\,m_1\,,$$

$$\left\{\ell^*_{\{(i-1)m_2+1\}},\; \ell^*_{\{(i-1)m_2+2\}},\; \cdots\;,\; \ell^*_{im_2}\right\} \in \{0,\;1,\;2,\;\cdots\;,\;(m_1m_2-1)\}\;,$$

$$\boldsymbol{G}_1 \cup \boldsymbol{G}_2\;\cdots\;\cup \boldsymbol{G}_{m_1} \;\equiv\; \boldsymbol{T}_{\ell 0} \;=\; \{0,\;1,\;2,\;\cdots\;,\;(m_1m_2-1)\}\;.$$

Then a constant block-sum design equispaced treatment levels vector $\boldsymbol{t}^*_{2w}$ is given by,

$$\boldsymbol{t}^*_{2w} \;=\; \ell_1\mathbf{1}_v + d\left(\ell^*_1,\,\ell^*_2,\,\cdots,\,\ell^*_{m_2},\,\ell^*_{m_2+1},\,\ell^*_{m_2+2},\,\cdots,\ell^*_{m_1m_2}\right)'\;. \tag{8}$$

An equispaced constant block-sum design D^*_{SGD} is obtained by replacing the ith element of $\boldsymbol{T}$ of (5) in design D_{SGD} by the $i$the element of $\boldsymbol{t}^*_{2w}$ of (8). The block size being m_2k_0, the treatment levels (8) imply that the constant block-sum equals $k_0\ell_{SUM}/m_1$. Alternatively, D^*_{SGD} can be obtained by replacing treatment i in the BIB design D by the m_2 elements of $\ell_1\mathbf{1}_{m_2} + d\boldsymbol{G}_i$, $i = 1,\,2,\,\cdots,\,m_1$. For illustration let us consider Example 1 again.

Example 1 continued: Let D be the BIB design with parameters $v_0 = b_0 = 3$, $r_0 = k_0 = 2$, $\lambda_0 = 1$, with blocks given by [1 2], [1 3], [2 3]. Then the SGD design S21 of Clatworthy (1973) is obtain by replacing treatment i in D by $m_2 = 3$ treatments as described above. Thus, replace treatments 1, 2, 3 in D by the treatment groups (1, 2, 3), (4, 5, 6) and (7, 8, 9) respectively to obtain the SGD design S21 or D_{SGD}. From (3.3) we have

$$\boldsymbol{T}_{\ell 0} \;=\; (0,\;1,\;2,\;3,\;4,\;5,\;6,\;7,\;8)'\;,$$

with $\ell_{SUM} = 36$. Taking $\boldsymbol{G}_1 = (0,\,4,\,8)$, $\boldsymbol{G}_2 = (1,\,5,\,6)$ and $\boldsymbol{G}_3 = (2,\,3,\,7)$, gives the sum of elements in each group to be $\ell_{SUM}/m_1 = 12$. Suppose $\ell = 1.5$ and $d = 0.3$. Then the requisite equispaced constant block-sum design D_{SGD} is obtained by replacing treatment i in the BIB design D by $m_2 = 3$ elements of $1.5\mathbf{1}_3 + d\boldsymbol{G}_i$, $i = 1,\,2,\,3$. The designs S22, S23, S24, and S25 in Clatworthy (1973) are obtained by taking replications of design S21. Corresponding constant block-sum designs can then be obtained by taking replications of D_{SGD}

Most of the SGD designs listed in Clatworthy (1973) are constructed using irreducible BIB designs. Let $D^{m_1}_{k_0}$ denote the irreducible BIB design for $v_0 = m_1$ treatments in blocks of size k_0. Then, the groups $\boldsymbol{G}_i$ for $m_2 = 2$ are as below, where the subscript 2 indicates the value of m_2,

$$\boldsymbol{G}_{2i} \;=\; \{(i-1)\,,\;\;(2m_1-i)\}\,, \qquad\qquad i = 1,\,2,\,\cdots,\,m_1\,.$$

The D_{SGD} designs corresponding to S1 to S20 can thus be obtained using $\boldsymbol{G}_{2i}, i = 1, 2, \cdots, m_1$. Constant block-sum designs for some other values of m_2 can also be similarly developed. The reader is also referred to Khattree (2019) for constructions of some equispaced SGD constant block-sum designs.

Acknowledgements

I thank Dr. Ravi Khattree for several constructive suggestions, including the analytical solution of equation (2.4) and the equispaced constant block-sum design for Example 1, which led me to develop the general formulation presented in Section 3. Sincere thanks are also due to a referee for helpful comments.

References

Bansal, N. and Garg, D. K. (2020). Construction and existence of constant sum PBIB designs. *Communications in Statistics - Theory and Methods,* https://doi.org/10.1080/03610926.2020.1772308

Clatworthy, W. H. (1973). *Tables of Two-Associate Class Partially Balanced Designs. National Bureau of Standards Applied Mathematics Series 63.*

Dey, A. (1986). *Theory of Block Designs. Wiley Eastern Limited: New Delhi*

Khattree, R. (2018a). A Note on the Nonexistence of the Constant Block-Sum Balanced Incomplete Block Designs. *Communications in Statistics - Theory and Methods,* https://doi.org/10.1080/03610926.2018.1508715.

Khattree, R. (2018b). The Parshvanath Yantram Yields a Constant-Sum Partially Balanced Incomplete Block Design. *Communications in Statistics - Theory and Methods,* https://doi.org/10.1080/03610926.2017.1417439.

Khattree, R. (2019). On Construction of Constant Block-Sum Partially Balanced Incomplete Block Designs. *Communications in Statistics - Theory and Methods,* https://doi.org/10.1080/03610926.2019.1576895.

Khattree, R. (2020). On construction of equireplicated constant block-sum designs. *Communications in Statistics - Theory and Methods,* https://doi.org/10.1080/03610926.2020.181481.

Nigam, A. K., Puri, P. D. and Gupta, V. K. (1988). *Characterizations and Analysis of Block Designs. Wiley Eastern Limited: New Delhi*

Raghavarao, D. (1971). *Constructions and Combinatorial Problems in Design of Experiments. Wiley: New York*

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), 149-172

Small Area Estimation – Some Applications in NSSO Surveys

A. K. Srivastava and Hukum Chandra
ICAR-Indian Agricultural Statistics Research Institute, New Delhi.

Received: 05 December 2020; Revised: 06 January 2021; Accepted: 09 January 2021

Abstract

The purpose of this article is to use small area estimation (SAE) method to produce district level estimates for some of the important indicators such as living condition, poverty incidence and working population ratio. For this purpose, data from 68th round (2011-12) of National Sample Survey Office (NSSO) pertaining to Household Consumer Expenditure Survey (HCES) and Employment and Unemployment Survey (EUS) for Uttar Pradesh has been used along with the 2011 Population Census data. The empirical results, evaluated through set of internal and external diagnostics measures, show that the district-level estimates generated through SAE approach are precise than the direct estimates. Spatial maps showing district level inequality in distribution of living condition, poverty incidence and working population ratio in Uttar Pradesh are also produced. These maps and districts level estimates are important for target oriented effective policy planning, monitoring and decision-making. In this article we deliberately consider two types of estimates *viz*. averages and proportions and use two different survey data of NSSO for producing district level estimates. We then illustrate how the existing survey data can be linked with Census data to produce reliable, timely and cost-effective district-level estimates of averages and proportions. The SAE methodology, illustration and guidelines set out in this paper can be adopted in other existing surveys for generating the disaggregate level estimates.

Key words: NSSO survey*;* Small area estimation; Precision; Living condition; Working population ratio.

0. Prologue

This paper is a tribute in honour and loving memory of Dr. Aloke Dey who had been a close friend to me all along for more than five and a half decades. Right from our student days to the entire professional career, he had been a source of strength and inspiration to all of us. His intense concern for maintaining high standards and values in research and teaching had a deep influence on his friends, colleagues, and students. Improvement in statistical system of the country was also remarkably close to his heart. This paper is an effort towards bringing in Small Area Estimation Techniques closer to application into some of the NSSO surveys – a hearty tribute from our side. - A K Srivastava.

Corresponding Author: A. K. Srivastava
E-mail: arunsrivast@gmail.com

1. Introduction

The NSSO surveys are generally conducted to generate a huge range of invaluable and crucial data, separately for the rural and urban sectors of the country, for States and Union Territories, and for different socio-economic groups. However, there is a rapidly growing demand for disaggregate level estimates (*e.g*. district or further disaggregate level) in India as the country is moving towards more decentralized system of governance. The disaggregate level estimates are also inevitable for several sustainable development goals (SDGs) related indicators. Just to mention some early attempts in India, an expert committee on small area statistics (SAS) was set up by (then) Ministry of Planning and Programme Implementation, Government of India (Government of India, 1997) under the chairmanship of Professor J. Roy. The committee deliberated upon the implications of 73rd and 74th amendments in the Constitution in view of data needs and its availability and highlighted the need for methodological studies for generating small area statistics appropriate to Indian conditions. This paper particularly concentrates on providing district level estimates for NSSO surveys where estimates are mostly generated at state-level. The SAE techniques provide a viable approach for producing estimates at smaller levels (Rao and Molina, 2015). The models used in SAE are commonly grouped as area level or unit level model. Area-level modelling is typically used when unit-level data are unavailable, or, as is often the case, where model covariates or auxiliary variables are only available in aggregate form. In this article, we motivate the SAE method based on area level small area modeling because in India the auxiliary variables are often accessible and available at aggregate (*e.g.* district) level. In this context, Fay–Herriot model (Fay and Herriot, 1979) is a widely used area level model in SAE. But this model is suited for continuous data. If the variable of interest is binary and the aim is to estimate small area proportions, then the area level generalized linear mixed model with logit link function, also referred to as the logistic linear mixed model (LLMM) is generally used (Johnson *et al*., 2010; Chandra *et al*., 2011 and Chandra *et al*., 2019). Srivastava (2007) used Fay–Herriot method of SAE to generate district level estimates for monthly per capita consumer expenditure (MPCE) using the 2004-05 Household Consumer Expenditure Survey (HCES) data of NSSO for the state of Uttar Pradesh. Srivastava (2009) further used the same data for estimating several poverty indicators at district level. Singh. *et. al.* (2005) used NSSO data for application of spatio-temporal models in SAE. More recently, Anjoy *et al.* (2020) used All India Debt and Investment Survey 2012-13 of NSSO for estimating the district-wise proportions of indebted households in rural areas of Karnataka. Chandra (2020) applied SAE method to estimate the incidence of food insecurity in different districts of rural areas of the state of Uttar Pradesh using the 2011-12 HCES of NSSO.

In this article, we consider SAE methods to produce district level estimates of the average household MPCE, the proportion of poor households and the employment rate for both rural and urban sectors for the state of Uttar Pradesh. Throughout this article, the proportion of poor households (*i.e.* proportion of households below poverty line) is also referred by poverty incidence and poverty rate (PR). The employment rate (UR) is referred as the proportion of persons employed to total persons. Alternatively, the worker-population ratio (WPR), also referred as work-force participation rate (WFPR) is defined as the number of persons employed per 1000 persons (*i.e.* WPR=1000×UR). The work force in the usual status includes the persons who worked for a relatively long part of the 365 days preceding the date of survey and the persons from among the remaining population who had worked at least for 30 days during the reference period of 365 days preceding the date of survey. The estimates of average household MPCE and poverty rate from the HCES of NSSO and the estimates of employment rate from the employment and unemployment survey (EUS) are common statistics generated by all the

states and used by different departments and ministries. This article deliberates these parameters and illustrates how the existing HCES and EUS data can be used to generate precise district level estimates. We elaborate two types of estimates *viz*. average and proportion (rate) and use two different survey data (HCES and EUS) of NSSO linking with Census data for producing district level estimates. This example can also be used as guidelines for generating the district level estimates of other commonly required parameters from the other existing surveys.

The paper is organized as follows. Besides introductory part in section 1, we describe data sources, different indicator variables of interest, and choice of auxiliary variables for SAE modelling in Section 2. Section 3 briefly delineates methodology used in the applications considered in this paper. Some aspects of the methodological framework have been discussed in Srivastava (2007, 2009) and Chandra (2020). In fact, Chandra (2020) applied the approach to estimate district-wise proportion of food insecure households in rural areas of Uttar Pradesh. However, for the sake of clarity and completeness, approach is described briefly in Section 3. The empirical results including essential diagnostic measures and discussions are deliberated in Section 4. Finally, Section 5 concludes the paper with some final remarks and recommendations.

2. Data Sources and Model Selection

The small area applications reported in this paper are based on the HCES and the EUS data from 68th round (2011-12) of NSSO for both rural and urban sectors of Uttar Pradesh and the 2011 Population Census. The 2011-12 HCES data is used to estimate the average household MPCE and the proportion of poor households (*i.e.* poverty ratio or PR) at district level for both rural and urban sectors in Uttar Pradesh. On the other hand, the estimation of UR (or WPR) is based on the 2011-12 EUS data. The household MPCE and the binary variable indicating whether a household is poor or not are the target variables of interest in 2011-12 HCES data. In this application a household having MPCE below the state poverty line is defined as poor. The poverty line used in this study (Rs. 768 for rural and Rs. 941 for urban) is the same as that set by the then Planning Commission, Govt. of India, for 2011-12. The parameters of interest are the average household MPCE and the PR within each district. In 2011-12 EUS data, the parameter of interest is the UR or WPR. In 2011-12 HCES, a total of 5916 rural and 3102 urban households from the 71 districts of Uttar Pradesh were surveyed. The district sample sizes for rural areas ranged from 32 to 128 with average of 83. Similarly, the district sample sizes for urban areas varied from 30 to 128 with average of 44. On the other hand, the 2011-12 EUS enumerated 49513 persons (33738 in rural areas and 15775 in urban areas). The district level sample sizes are relatively small for generating precise district level estimates.

The 2011 Population Census has a range of auxiliary variables (covariates) which can be explored for SAE modelling. However, we identified few relevant auxiliary variables for this study. We also used Principal Component Analysis (PCA) to derive composite scores for selected groups of auxiliary variables, separately for both rural and urban areas. Using district aggregates of rural data, we did PCA for two groups of auxiliary variables, denoted as R1and R2. The first group (R1) consisted of the proportions of main worker by gender, proportions of main cultivator by gender and proportions of main agricultural labourer by gender. The second group (R2) consisted of proportions of marginal cultivator by gender and proportions of marginal agriculture labourers by gender. The first principal component (R11) for the first group explained 44% of the variability in the R1, while adding the second component (R12) increased explained variability to 69%. The first principal component (R21) for the second

group explained 52% of the variability in the R2 group, while addition of second component (R22) increased explained variability to 90%. For urban areas, we further applied PCA, separately for two groups of variables, as defined in rural data, but using district aggregates of urban data. These are denoted as U1 and U2. Here, the first principal component (U11) explained 53% of the variability and addition of the second component (U12) explained 83% variability to in the U1. The first principal component (U21) for the U2 explained 63% of the variability, while adding the second component (U22) enhanced explained variability to 87%.

For both rural and urban data, we separately fitted a linear model using district-wise direct estimates of MPCE as the response variable and the PCA scores and other auxiliary variables as covariates. The final model with selected covariates was used to produce district-wise estimates of average household MPCE. The model was fitted using the lm() function in R using the district specific sample sizes as the weight. We also fitted a generalized linear model (GLM) using direct estimates of proportions of poor households versus the PCA scores and other auxiliary variables for each group of data. The model was fitted using the glm() function in R and specifying the family as "binomial" and the district wise sample sizes as the weight. We also fitted a GLM using direct estimates of employment rates versus a set of auxiliary variables. In each case, model fitting was used for selection of final model for SAE analysis. Table 1 provides list of selected covariates which were used in SAE of average household MPCE, poverty incidence and employment rate.

Table 1: Selected auxiliary variables for SAE of the average MPCE, the proportion of poor households (poverty incidence) and the working population ratio

Parameter	Rural	Urban
MPCE	SC (Proportion of scheduled caste to total population), Literacy rate, R11, R21and R22	Literacy rate and TWPR (Proportion of worker to total population)
Poverty incidence	SC, Literacy rate, R11, R21and R22	Literacy rate and TWPR
Employment rate	SC, TWPR, Number of households and Total population	SC, Literacy rate and TWPR

3. Methodological Framework

This Section briefly introduces the SAE methods applied for producing the district level estimates of average household MPCE, poverty incidence and employment rate (or WPR) and their measure of precision for rural and urban areas of the state of Uttar Pradesh. Let y_{ij} denote the value of the variable of interest for unit $j(j=1,\ldots,N_i)$ in district $i(i=1,\ldots,D)$, where N_i and D denote the population size of district i and total number of districts in the population respectively. The quantity of interest in district i is the population mean (or proportion, in case of binary variable) m_i defined as $m_i = N_i^{-1}\sum_{j=1}^{N_i} y_{ij}$. Let n_i denotes the sample size in district i, then the direct estimator of m_i is $\hat{m}_i^{Direct} = \left(\sum_{j=1}^{n_i} w_{ij}\right)^{-1}\sum_{j=1}^{n_i} w_{ij}y_{ij}$, where w_{ij} is inverse of the inclusion probability for unit j in district i. The estimate of variance of direct estimator $\hat{m}_i^{Direct}$ is $v(\hat{m}_i^{Direct}) \approx \left(\sum_{j=1}^{n_i} w_{ij}\right)^{-2}\sum_{j=1}^{n_i} w_{ij}(w_{ij}-1)(y_{ij}-\hat{m}_i^{Direct})^2$. Let $\hat{m}_i^{Direct}$ be the observed

direct estimate of average MPCE for district i. Let $\mathbf{x}_i$ be the k-vector of known district level auxiliary variables, related to the population parameter m_i. Then district specific Fay and Herriot (1979) model is described as $\hat{m}_i^{Direct} = m_i + e_i$ and $m_i = \mathbf{x}_i^T\boldsymbol{\beta} + u_i$. Alternatively, this model can be expressed as

$$\hat{m}_i^{Direct} = \mathbf{x}_i^T\boldsymbol{\beta} + u_i + e_i;\ i = 1,...,D. \tag{1}$$

Here $\boldsymbol{\beta}$ is a k-vector of unknown fixed effect parameters, u_i's are independently and identically distributed normal random errors with $E(u_i) = 0$ and $Var(u_i) = \sigma_u^2$, and $e_{\alpha i}$'s are independent sampling errors normally distributed with $E(e_i \mid m_i) = 0$, $Var(e_i \mid m_i) = \psi_i$. The two errors are independent of each other within and across districts. Let $\hat{\sigma}_u^2$ denote the estimator of σ_u^2 and $\hat{\boldsymbol{\beta}}$ the empirical best linear unbiased estimator of $\boldsymbol{\beta}$. The empirical best linear unbiased predictor (EBLUP) estimate of m_i is then

$$\hat{m}_i^{EBLUP} = \mathbf{x}_i^T\hat{\boldsymbol{\beta}} + \hat{u}_i;\ i = 1,...,D. \tag{2}$$

Here, $\hat{u}_i = \hat{\gamma}_i(\hat{m}_i^{Direct} - \mathbf{x}_i^T\hat{\boldsymbol{\beta}})$, where $\hat{\gamma}_i = \hat{\sigma}_u^2\left(\hat{\sigma}_u^2 + \psi_{\alpha i}\right)^{-1}$ defines the shrinkage effect for district *i*. The mean squared error (MSE) estimation of EBLUP (2) follows from Rao and Molina (2015). Readers can also refer to Chandra (2013) for the expression of MSE estimate of EBLUP (2).

It is worth noting that the direct estimate of proportions (*e.g.* PR and ER) can also be modelled by Fay-Herriot model (1) and the EBLUP estimate of district level proportions can be obtained. However, the estimate of district level proportions derived from the EBLUP (2) might be inconsistent in the sense that they might not be within the [0,1] interval. We describe approach to model district-specific proportions under a LLMM to produce precise district level estimates of PR and ER. For example, for estimating PR, the binary variable y_{ij} takes value 1 when household *j* in district *i* is poor and 0 otherwise. Similarly, in case of ER, it assumes value 1 when person *j* in district *i* is employed and 0 otherwise. In this case, population parameter of interest in district *i* is the district level proportion. Let $y_{si} = \sum_{j \in s_i} y_{ij}$ denotes the sample count in district *i*, which follows a Binomial distribution with parameters n_i and π_i, *i.e.* $y_{si} \mid v_i \sim \text{Bin}(n_i, \pi_i)$, where π_i is a success probability. The model linking π_i with the covariates $\mathbf{x}_i$ is the LLMM of form

$$logit(\pi_i) = \ln\left\{\pi_i(1-\pi_i)^{-1}\right\} = \eta_i = \mathbf{x}_i^T\boldsymbol{\alpha} + v_i, \tag{3}$$

with $\pi_i = \exp(\mathbf{x}_i^T\boldsymbol{\alpha} + v_i)\left\{1 + \exp(\mathbf{x}_i^T\boldsymbol{\alpha} + v_i)\right\}^{-1}$, where $\boldsymbol{\alpha}$ is the k-vector of regression coefficients and v_i is the district-specific random effect with v_i : $N(0, \sigma_v^2)$. Here, the sampling information has been incorporated by replacing the "actual sample size" and the "actual sample count" with the "effective sample size" and the "effective sample count" respectively, see for

example, Chandra *et al*. (2019). Assuming $N_i >>> n_i$, a plug-in empirical predictor (EPP) of proportion (*e.g*., PR or ER) m_i in district *i* is

$$\hat{m}_i^{EPP} = \exp(\mathbf{x}_i^T\hat{\boldsymbol{\beta}} + \hat{v}_i)\left(1+\exp(\mathbf{x}_i^T\hat{\boldsymbol{\beta}} + \hat{v}_i)\right)^{-1}. \quad (4)$$

The expression for the estimate of MSE of EPP (4) is given in Chandra *et al.* (2019).

4. Results and Discussions

This Section illustrates and discusses the district-wise estimates of average household MPCE, incidence of poverty and UR generated by direct and SAE methods for both rural and urban areas of Uttar Pradesh. The EBLUP (2) under FH model is used to produce the district-wise estimates of average household MPCE and the EPP (4) is applied for generating the district-wise estimates of incidence of poverty (or PR) and employment rate for rural and urban areas. The corresponding estimates of MSE are also computed to assess the reliability of estimates and also to construct the confidence interval for the estimates. The district-specific estimates of average household MPCE, PR and WPR along with their SEs and CVs generated by the Direct and SAE methods for Uttar Pradesh are provided in the Appendices (Tables A1-A6).

A set of diagnostics measures are implemented before making inferences about small area estimates. Such diagnostics measures are (i) the model diagnostics, and (ii) the small area estimates diagnostics. The model diagnostics are tested to verify the assumptions of the underlying model. For example, the small area models (1) and (3) assume that the random district specific effects have a normal distribution with mean zero and fixed variance. The district specific residuals are expected to be randomly distributed around zero if the model assumptions are satisfied. Histogram and q-q plot are also checked to inspect the normality assumption. For this study, the district level residuals are randomly distributed around zero and the histograms as well as the q-q plots also provide evidence in support of the normality assumption. In addition, we use the Shapiro-Wilk (SW) test to examine the normality of the district random effects. The other diagnostics are demonstrated to examine the level of validity and accuracy of the small area estimates. Three commonly used diagnostics measures for evaluating the validity and the reliability of the small area estimates: the bias diagnostic, the percent coefficient of variation (CV) diagnostic and the 95% confidence interval (CI) diagnostic. The first diagnostics assesses the validity and last two review the improved precision of the small area estimates level (Chandra *et al*., 2011). For bias diagnostic we plot direct estimates (*Y*-axis) vs. small area estimates (*X*-axis) and we looked for divergence of the fitted least squares regression line from the line of equality. Although results not reported here, the bias diagnostic plots revealed that the district level estimates of MPCE, poverty incidence and WPR for both rural and urban are less extreme when compared to the corresponding direct estimates. We also use a Goodness of Fit (GoF) diagnostic, which is equivalent to a Wald test, for whether the differences $D_i = \hat{m}_i^{Direct} - \hat{m}_i$ between direct estimates $\hat{m}_i^{Direct}$ and small area estimates $\hat{m}_i$ of a population parameter (m_i) are statistically different. The null hypothesis is that the direct and small area estimates are statistically equivalent. The alternative is that the direct and small area estimates are statistically different. This Wald test statistic is computed as $\mathrm{W} = \sum_i \left\{ D_i^2 \left[var(\hat{m}_i^{Direct}) + mse(\hat{m}_i) \right]^{-1} \right\}$. Assuming $\hat{m}_i^{Direct}$ and $\hat{m}_i$ are independently distributed, which is not unreasonable for large sample sizes, the value of W can be compared with an appropriate critical value from a chi square distribution with degrees of freedom *D*

equal to the number of districts. For our analysis, $D = 71$, with a critical value of 91.67 at a 5% level of significance calculated using qchisq function in R. A small value (<91.67 here) of W indicates no statistically significant difference between small area and direct estimates. The results from GoF diagnostic are given in Table 2. The values of W are smaller than the 91.67, which indicates that small area estimates are consistent with the direct estimates. In general, the bias diagnostics reflect that the small area estimates are consistent with the direct survey estimates.

Table 2: Goodness of fit diagnostic

Estimate	Rural	Urban
MPCE	11.87	7.70
Poverty	28.39	13.04
WPR	26.20	44.17

We computed the CV to compare the extent to which the small area estimates of MPCE, poverty incidence and WPR improve in precision compared to the corresponding direct estimates. There is no standard, universally accepted definition of what constitutes large or small CV values. However, different organizations have different cut-offs: for instance, the UK Office for National Statistics has a cut-off CV value of 20% for acceptable estimates, while in the US the National Center for Health Statistics has a cut-off of 30% for county-level health statistics (Baffour *et al*., 2019). Figure 1 displays the district-wise values of CV for small area estimates and direct estimates in increasing order of sample sizes. The distribution of CV in Figure 1 shows that in most of the districts, the CVs of the small area estimates are significantly smaller than those of the direct survey estimates, implying that the small area estimates are less variable, and hence relatively more precise than the direct survey estimates. The improvement CV is higher for the districts with smaller sample sizes as compared to the larger sample sizes. A set of summary statistics for the direct and small area estimates along with associated standard errors and CV of the MPCE, poverty incidence and WPR over 71 districts for rural and urban area are reported in Table 3. As expected, the average values of MPCE, poverty incidence and WPR estimates generated by SAE are almost identical to those of the direct estimates but with lower variability (*i.e*. smaller values of standard deviation). For example, the standard deviations of MPCE estimates for rural area generated by the direct and the SAE methods are 50 and 24, respectively. From Table 3, it is obvious that the small area estimates of MPCE, poverty incidence and WPR are more precise and representative than the direct estimates for both rural and urban areas. We now examine the 95 % confidence interval for the direct estimates compared to the small area estimates. For more precise estimates, we expect the width of the confidence interval to be narrower. The district-wise plots of the 95% confidence intervals (CIs) for the average household MPCE, poverty incidence and WPR generated by direct and SAE methods (EBLUP for average household MPCE and EPP for both poverty incidence and WPR) are displayed in Figure 2. These plots show that the 95% CIs for the direct estimates are wider than the 95% CIs for the small area estimates for the average household MPCE, poverty incidence and WPR. We further note that in many districts the 95% CI for direct estimates are invalid (for example, negative values for poverty incidence) due to large standard errors. Finally, we examine the aggregation property of the small area district-level estimates generated by SAE methods at higher (*e.g*. State) level. Let $\hat{m}_i$ and N_i denote the estimate of an average or proportion y_i and population size for district *i*. The divisional and state-level estimates an average or proportion is then calculated as $\hat{m} = \sum_{i=1}^{D} N_i \hat{m}_i \Big/ \sum_{i=1}^{D} N_i$.

Table 4 reports the state level estimates of the average household MPCE, poverty indicator and WPR generated by direct and SAE methods. Comparing these estimates, we see that the small area estimates are close to the direct survey estimates at state level.

Table 3: Summary distribution of direct and model-based small area estimates along with their standard error (SE) and percent CV of MPCE, poverty incidence and WPR

Rural							
Parameter	Statistics	Direct estimate			Small area estimate		
		Estimate	SE	CV	Estimate	SE	CV
Average household MPCE	Minimum	774	38	3.73	791	37	3.67
	Maximum	1958	309	25.35	1558	139	14.43
	Average	1083	83	7.44	1059	68	6.40
	Std. deviation	224	50	3.61	165	24	1.96
Poverty incidence	Minimum	0.002	0.011	12.91	0.060	0.024	12.75
	Maximum	0.578	0.169	99.38	0.506	0.083	42.32
	Average	0.249	0.071	35.08	0.251	0.055	24.37
	Std. deviation	0.137	0.025	17.44	0.107	0.014	6.72
WPR	Minimum	224	22	6.80	280	20	5.86
	Maximum	507	79	21.68	430	36	10.16
	Average	338	37	11.03	337	26	7.65
	Std. deviation	57	10	3.25	33	3	0.99
Urban							
Average household MPCE	Minimum	791	61	6.08	796	60	6.02
	Maximum	6453	609	22.42	4762	482	19.39
	Average	1623	185	10.97	1569	172	10.61
	Std. deviation	835	123	3.83	637	99	3.34
Poverty incidence	Minimum	0.003	0.003	11.98	0.023	0.010	11.54
	Maximum	0.736	0.136	105.41	0.667	0.122	56.94
	Average	0.281	0.077	36.73	0.278	0.061	26.57
	Std. deviation	0.183	0.032	18.47	0.161	0.024	10.89
WPR	Minimum	222	28	7.51	260	18	5.75
	Maximum	487	57	21.51	393	27	7.72
	Average	313	41	13.39	312	20	6.51
	Std. deviation	51	7	2.73	25	2	0.49

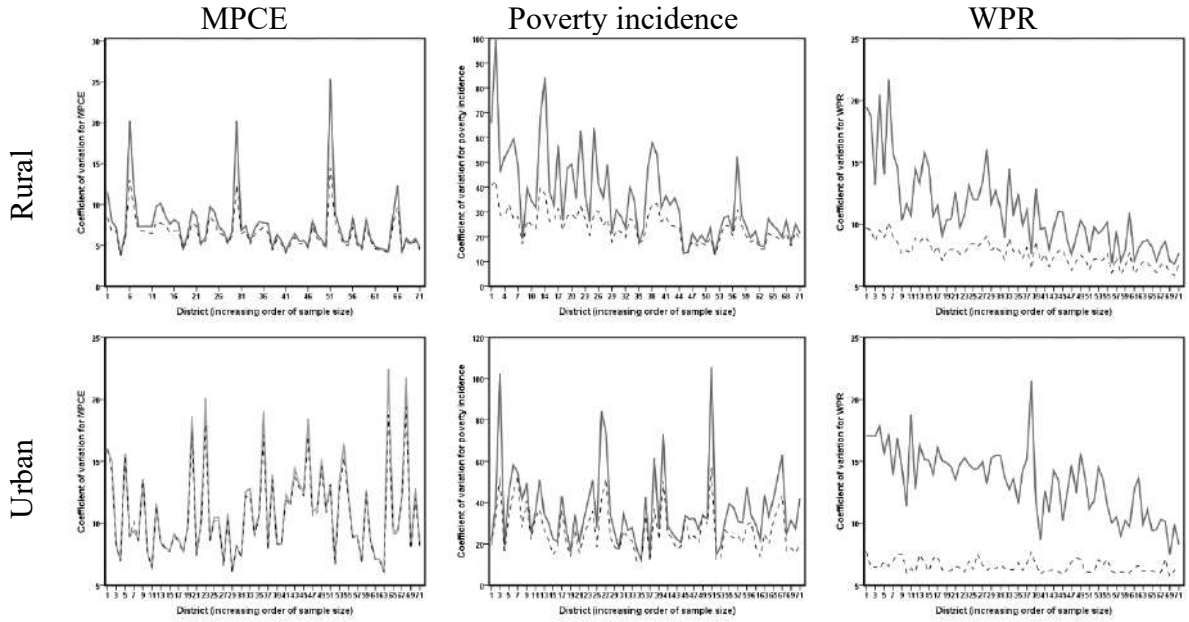


Figure 1: District-wise coefficient of variation (%) for the small area estimates (solid line) and the direct estimates (dash line). Districts are arranged in increasing order of sample sizes.

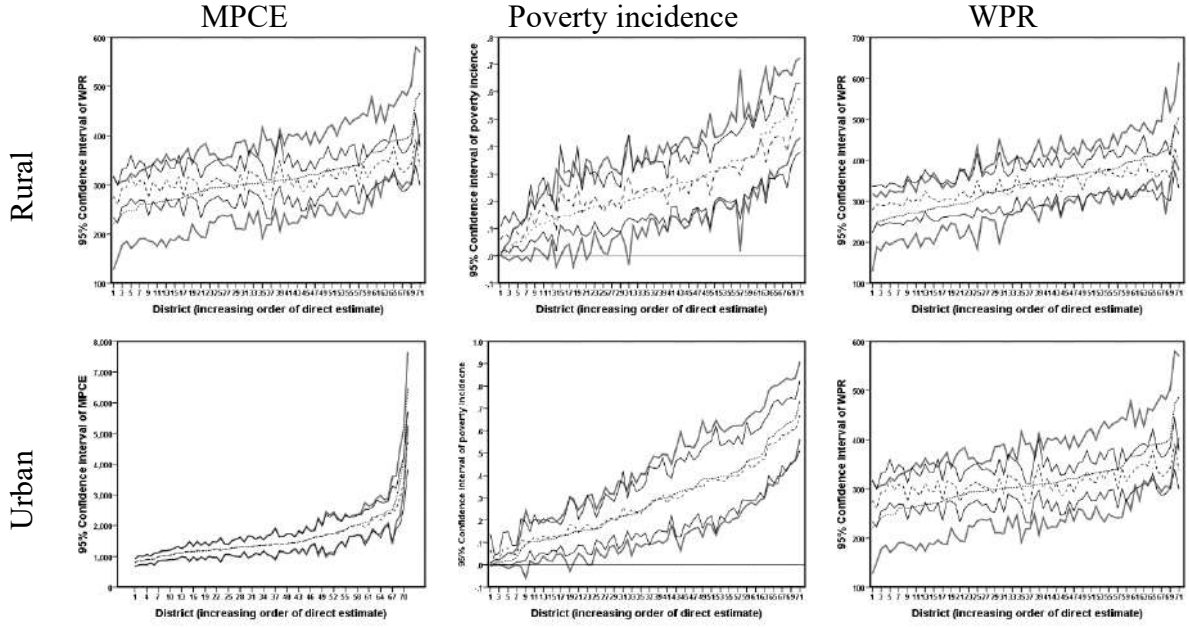


Figure 2: District-wise 95 percentage nominal confidence interval (95% CI) for the direct (solid line) and small area (thin line) estimates. Direct (dotted point) and model-based estimates (dash point) are shown in the 95% CI.

The district-specific estimates of average household MPCE, poverty incidence and WPR along with their CVs and and 95% CIs generated by the direct and SAE methods are provided in the Appendices (Table A1-A6). The diagnostics measures clearly demonstrate that the small area estimates are more efficient, precise, and representative than the direct estimates. Consequently, statistical inferences and conclusions based on the small area estimates of MPCE, poverty incidence and WPR are expected to offer better and effective policy decisions. Therefore, hereafter in discussion we focus on the estimates of MPCE, poverty incidence and WPR generated by SAE methods. Figures 3-5 provides maps showing spatial distribution of MPCE, poverty incidence and WPR estimates respectively at district level for rural and urban areas of Uttar Pradesh produced from the SAE methods. Darker areas of the maps correspond to the areas with high values of estimates. These maps supplement the district-wise estimates along with CVs and 95% CIs set out in Appendices (Table A1-A6).

Table 3: Aggregated level estimates generated by direct and SAE methods

Parameters	Rural		Urban	
	Direct	SAE	Direct	SAE
MPCE (Rs)	1073	1050	1942	1934
Poverty rate (%)	25.8	25.7	19.2	19.4
WPR	338	336	317	310

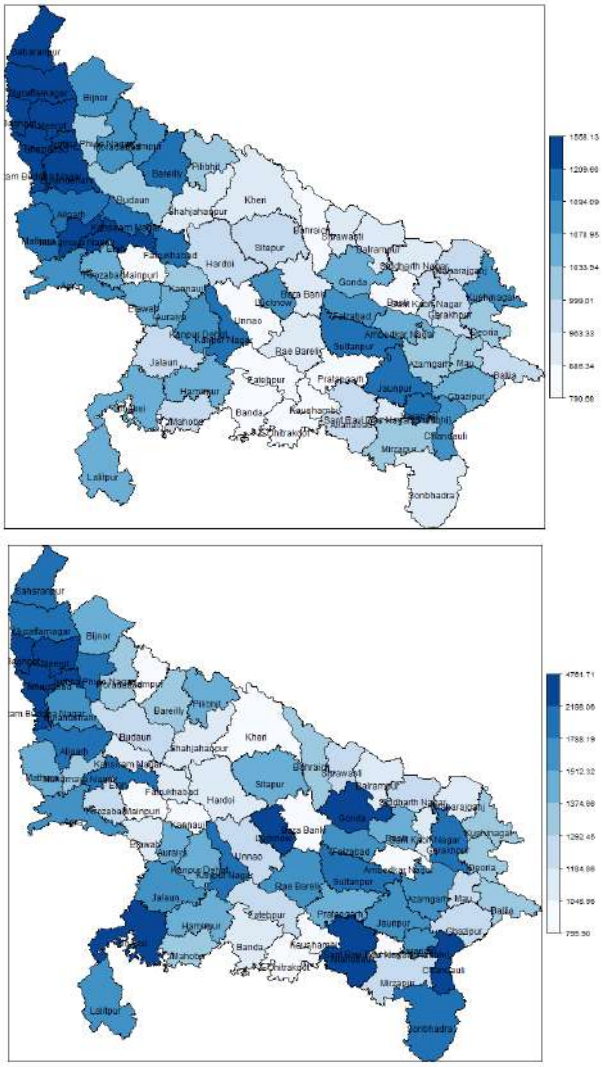


Figure 3: District-wise mapping of MPCE for rural (left) and urban (right) areas in the state of Uttar Pradesh generated by small area estimation method, 2011-12

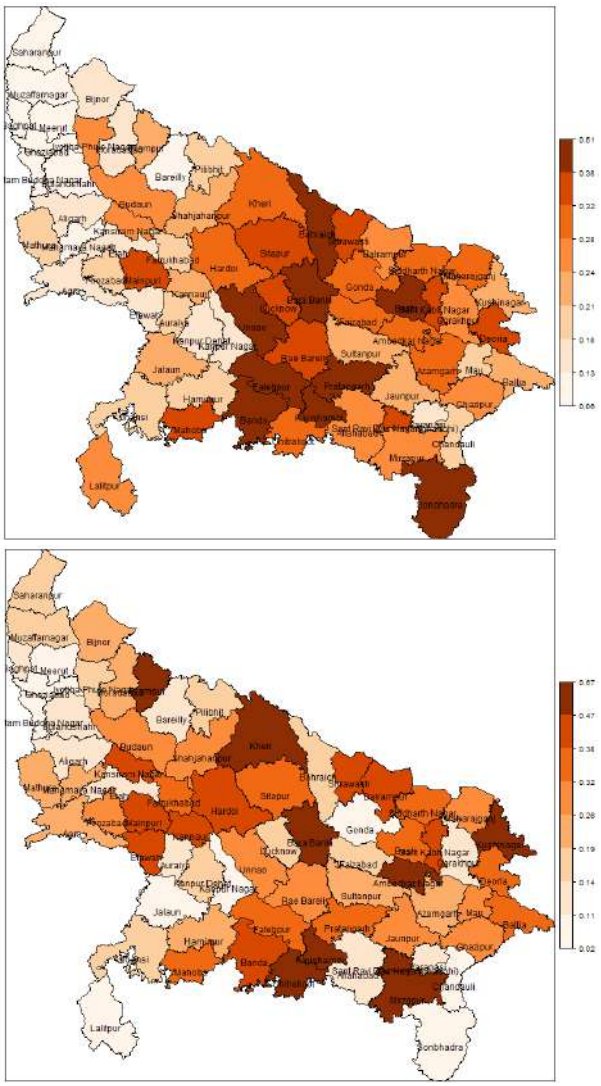


Figure 4: District-wise mapping of poverty incidence for rural (left) and urban (right) areas in the state of Uttar Pradesh generated by small area estimation method, 2011-12

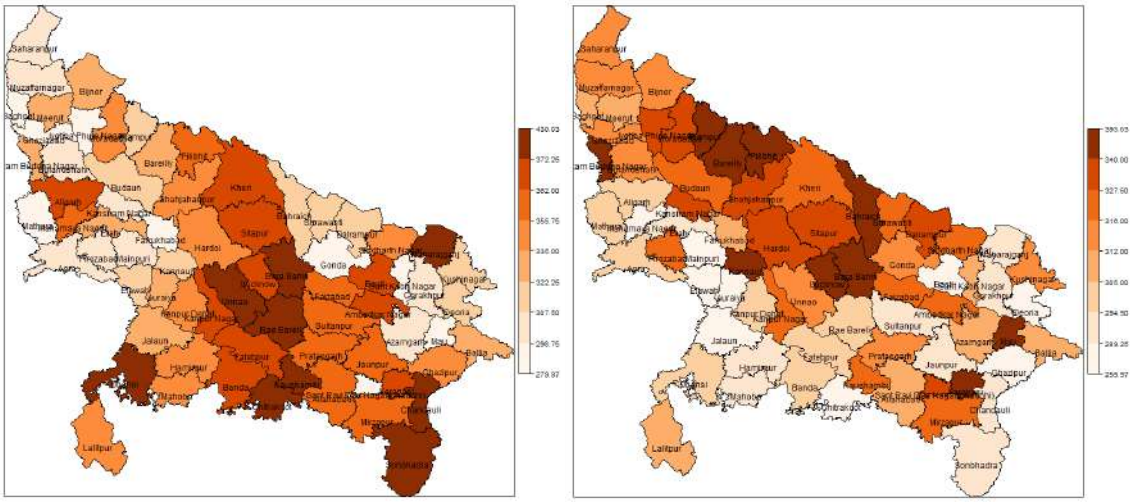


Figure 5: District-wise mapping of worker population ratio for rural (left) and urban (right) areas in the state of Uttar Pradesh generated by small area estimation method, 2011-12

5. Concluding Remarks and Recommendations

In India, Censuses are usually limited as they tend to focus mainly on the basic socio-demographic and economic data and are not available for every time - period. On the other hand, country is fortunate to have regular NSSO surveys for generating number of socio-economic indicators. The NSSO surveys are aimed to generate estimates at national and state level. They do not provide sub-state level statistics. There is no regular flow of estimates at districts and further disaggregate levels. It is known that state and national estimates usually mask variations (heterogeneity) at the sub-state or district level and render little information for micro level planning and allocation of resources. Recently, there has been a pressing demand for disaggregate level sustainable development goals (SDGs) related indicators in various departments in central and state governments and United Nations agencies in the country. Therefore, need for SAE has again achieved momentum. Despite the importance and urgent requirements, there are several virtual reasons for this topic not being implemented in the system. To the best of our knowledge and understanding, one such reason is technicality involved in SAE method. For example, SAE is combination of statistical modelling and survey estimation and there is no unique solution for all type of problems encountered. In order to develop a team of personnel with technical knowledge and experience in the field, adequate stability of the staff needs to be ensured.

This article demonstrated application of SAE approach to generate district level reliable and representative of the average household MPCE, poverty incidence and working population ratio for rural and urban areas of Uttar Pradesh by linking the latest round of 2011-12 HCES and 2011-12 EUS data of NSSO with the 2011 Population Census. The diagnostic measures clearly confirm that the estimates generated by SAE have reasonably good precision. The SAE method has also generated reliable estimates for the districts with smaller sample sizes. The district level estimates, and spatial mapping can provide useful information for the purpose of better strategic decision and policy planning. For example, many programmes are launched by Government of India with an objective to uplift the socio-economic condition of masses. NITI Aayog requires values of some socio-economic parameters for the backward districts, which they have identified, to see the impact of policy interventions and for future planning in these backward districts. NITI Aayog has identified 114 backward districts in rural India and 112 backward districts in urban India. They are monitoring some indicators related to socio-economic parameter on a continuous basis and thus providing district level estimates is very much apt for these districts. Further, the district level estimates are likely to be advantageous for allocating budget to target welfare interventions through recognizing the districts or regions with low average MPCE (or high poverty rate) and working population ratio. The indicators chosen here are based on HCES and EUS Surveys. In fact, NSSO conducts several other important Household Surveys such as Health Surveys, Education Surveys, Situation Analysis Surveys, AIDIS Surveys besides Establishment Surveys. There are several well identified indicators of interest for each of these surveys. The methodology and application presented in this paper can be used as guideline for producing reliable, timely and cost-effective estimates using survey data from different sectors.

Acknowledgments

The authors would like to acknowledge the valuable comments and suggestions of the Chair Editor, and the reviewer. These led to a considerable improvement in the paper.

References

Anjoy, P., Chandra, H. and Parsad, R. (2020). Estimation and spatial mapping of incidence of indebtedness in the state of Karnataka in India by combining survey and census data. *Statistics and Applications*, **18**(**1**) (New Series), 21-33.

Baffour, B., Chandra, H. and Martinez, A. (2019). Localised estimates of dynamics of multidimensional disadvantage: an application of the small area estimation technique using Australian survey and census data. *International Statistical Review*, **87**(**1**), 1-23.

Chandra, H. (2020). District-level estimates of extent of food insecurity for the state of Uttar Pradesh in India by combining survey and census data. In*: Special Proceeding of the 22nd Annual Conference of SSCA held at Savitribai Phule Pune University, Pune, during January* 02-04, 2020*,* 25-38.

Chandra, H. (2013). Exploring spatial dependence in area level random effect model for disaggregate level crop yield estimation. *Journal of Applied Statistics*, **40**, 823-842.

Chandra, H., Chambers, R. and Salvati, N. (2019). Small area estimation of survey weighted counts under aggregated level spatial model. *Survey Methodology*, **45**(**1**), 31-59.

Chandra, H., Salvati N. and Sud U. C. (2011). Disaggregate-level estimates of indebtedness in the state of Uttar Pradesh in India-an application of small area estimation technique. *Journal of Applied Statistics*, **38**(**11**), 2413-2432.

Fay, R. E. and Herriot, R. A. (1979). Estimation of income from small places: an application of James-Stein procedures to census data. *Journal of the American Statistical Association,* **74**, 269-277.

Government of India (1997). Report of the expert committee on small area statistics. *Department of Statistics, Ministry of Planning and Programme Implementation, Government of India, New Delhi*, April 1997.

Johnson F. A., Chandra H., Brown J. and Padmadas S. (2010). Estimating district-level births attended by skilled attendants in Ghana using demographic health survey and census data: an application of small area estimation technique. *Journal of Official Statistics,* **26** (**2**), 341–359.

Rao, J. N. K. and Molina, I. (2015). *Small Area Estimation*. 2nd Edition. John Wiley & Sons.

Singh, B. B., Shukla, G. K. and Kundu, D. (2005). Spatio-temporal models in small area estimation. *Survey Methodology*, **31**, 183195.

Srivastava A. K. (2009). Some aspects of estimating poverty at small area level. *Journal of the Indian Society of Agricultural Statistics*, **63** (**1**), 1-23.

Srivastava A. K. (2007). Small area estimation – a perspective and some applications. *Journal of the Indian Society of Agricultural Statistics*, **61** (**3**), 295-309.

Appendix

Table A1: District-wise sample size, direct estimates (Direct) and small area estimates (SAE) along with their standard error (SE) and percentage coefficient of variations (CV) of MPCE for rural areas of Uttar Pradesh in 2011-12

Region	District	Sample size	Direct			SAE		
			MPCE	SE	CV	MPCE	SE	CV
Western	Saharanpur	96	1419	89	6.30	1361	78	5.76
	Muzaffarnagar	128	1366	73	5.31	1345	66	4.91
	Bijnor	96	1068	59	5.55	1087	56	5.12
	Moradabad	128	1081	51	4.71	1083	49	4.49
	Rampur	64	1092	80	7.31	1081	71	6.61
	Jyotiba Phule Nr	64	1012	74	7.28	1032	67	6.47
	Meerut	64	1958	191	9.75	1558	123	7.87
	Baghpat	32	1885	218	11.56	1542	130	8.44
	Ghaziabad	64	1454	147	10.14	1430	110	7.71
	Gautam B Nr	32	1547	123	7.93	1465	98	6.68
	Bulandshahar	96	1247	53	4.29	1247	51	4.07
	Aligarh	95	1135	87	7.66	1151	76	6.61
	Hathras	64	1546	133	8.58	1360	102	7.50
	Mathura	64	1109	84	7.59	1116	74	6.63
	Agra	96	1063	58	5.50	1080	55	5.07
	Firozabad	64	1014	83	8.14	1075	73	6.82
	Etah	64	1436	111	7.71	1338	91	6.79
	Mainpuri	64	836	39	4.68	853	38	4.46
	Budaun	96	1016	65	6.42	1018	61	5.98
	Bareilly	95	1168	55	4.67	1168	52	4.43
	Pilibhit	64	1021	65	6.33	1024	60	5.84
	Shahjahanpur	96	921	51	5.52	939	49	5.18
	Farrukhabad	64	1149	107	9.28	1146	89	7.74
	Kannauj	64	973	84	8.63	1023	75	7.33
	Etawah	64	1045	55	5.28	1045	52	5.00
	Auraiya	64	1087	65	5.97	1076	60	5.59
	Kashiramnagar	32	1230	88	7.19	1161	79	6.79
Central	Kheri	128	936	76	8.11	947	69	7.25
	Sitapur	128	1002	59	5.86	993	55	5.58
	Hardoi	128	967	46	4.80	965	45	4.63
	Unnao	96	861	48	5.59	867	46	5.32
	Lucknow	64	1130	110	9.69	1083	92	8.51
	Rae Bareli	128	930	43	4.60	930	41	4.44
	Kanpur Dehat	64	1104	101	9.15	1090	85	7.80
	Kanpur Nagar	64	1139	83	7.27	1126	74	6.53
	Fatehpur	96	777	38	4.91	791	37	4.69

Southern	Jalaun	64	993	67	6.74	987	62	6.25
	Jhansi	64	1070	58	5.39	1056	54	5.14
	Lalitpur	32	1061	40	3.73	1052	39	3.67
	Hamirpur	32	1079	67	6.25	1069	62	5.81
	Banda	64	774	52	6.71	793	49	6.22
	Chitrakoot	32	839	170	20.23	879	114	12.99
	Mahoba	32	975	114	11.68	976	92	9.40
Eastern	Mahrajganj	96	1012	81	7.99	984	73	7.38
	Pratapgarh	128	870	40	4.54	880	38	4.35
	Kaushambi	63	809	59	7.29	819	56	6.83
	Allahabad	128	991	43	4.32	999	41	4.14
	Barabanki	96	900	56	6.26	906	53	5.86
	Faizabad	64	1378	278	20.18	1080	132	12.22
	Ambedkar Nagar	96	1047	59	5.67	1041	55	5.32
	Sultanpur	128	1313	115	8.78	1197	92	7.71
	Bahraich	96	828	40	4.85	833	39	4.69
	Shrawasti	64	888	61	6.88	887	57	6.47
	Balrampur	63	892	65	7.29	895	61	6.77
	Gonda	128	1063	131	12.31	1034	100	9.68
	Siddharthnagar	96	1220	309	25.35	962	139	14.43
	Basti	96	861	78	9.02	885	70	7.86
	Sant K Nagar	64	1006	74	7.36	991	67	6.77
	Gorakhpur	128	993	42	4.25	996	41	4.10
	Kushinagar	128	1108	65	5.89	1087	61	5.59
	Deoria	96	988	70	7.11	999	64	6.45
	Azamgarh	128	1020	54	5.28	1020	51	5.00
	Mau	64	1000	55	5.51	1007	52	5.16
	Ballia	96	955	52	5.44	976	49	5.07
	Jaunpur	128	1115	65	5.86	1098	61	5.53
	Ghazipur	128	1051	50	4.71	1050	47	4.50
	Chandauli	64	1092	76	6.98	1087	69	6.31
	Varanasi	96	1136	61	5.41	1152	57	4.99
	Sant Ravidas Nr	64	873	69	7.91	929	63	6.83
	Mirzapur	96	1023	84	8.18	1024	74	7.21
	Sonbhadra	64	946	73	7.74	928	67	7.21

Table A2: District-wise sample size, direct estimates (Direct) and small area estimates (SAE) along with their standard error (SE) and percentage coefficient of variations (CV) of poverty incidence for rural areas of Uttar Pradesh in 2011-12

Region	District	Sample Size	Direct			SAE		
			Poverty Incidence	SE	CV	Poverty Incidence	SE	CV
Western	Saharanpur	96	0.068	0.0362	53.23	0.089	0.0297	33.33
	Muzaffarnagar	128	0.052	0.0272	52.31	0.083	0.0257	30.95
	Bijnor	96	0.165	0.0524	31.78	0.160	0.0417	26.07
	Moradabad	128	0.131	0.0371	28.36	0.149	0.0365	24.48
	Rampur	64	0.240	0.0825	34.36	0.232	0.0574	24.76
	Jyotiba Phule Nr	64	0.268	0.0850	31.71	0.253	0.0589	23.28
	Meerut	64	0.002	0.00	0.00	0.060	0.0237	39.44
	Baghpat	32	0.138	0.0909	65.86	0.092	0.0374	40.67
	Ghaziabad	64	0.054	0.0367	68.04	0.063	0.0251	39.84
	Gautam B. Nr	32	0.018	0.0179	99.38	0.066	0.0279	42.32
	Bulandshahar	96	0.103	0.0377	36.59	0.110	0.0308	28.02
	Aligarh	95	0.181	0.0861	47.56	0.160	0.0452	28.23
	Hathras	64	0.013	0.0110	84.27	0.084	0.0311	37.08
	Mathura	64	0.179	0.0684	38.22	0.192	0.0504	26.25
	Agra	96	0.192	0.0628	32.73	0.179	0.0443	24.73
	Firozabad	64	0.252	0.0824	32.70	0.187	0.0496	26.52
	Etah	64	0.126	0.0716	56.79	0.139	0.0437	31.44
	Mainpuri	64	0.451	0.1198	26.57	0.343	0.0772	22.51
	Budaun	96	0.233	0.0829	35.57	0.245	0.0591	24.11
	Bareilly	95	0.047	0.0272	57.88	0.099	0.0329	33.20
	Pilibhit	64	0.177	0.0844	47.71	0.192	0.0546	28.43
	Shahjahanpur	96	0.270	0.0820	30.38	0.229	0.0549	23.96
	Farrukhabad	64	0.184	0.0907	49.27	0.190	0.0553	29.11
	Kannauj	64	0.308	0.1089	35.37	0.221	0.0601	27.19
	Etawah	64	0.093	0.0582	62.61	0.160	0.0516	32.23
	Auraiya	64	0.148	0.0540	36.51	0.205	0.0555	27.07
	Kashiramnagar	32	0.161	0.0744	46.19	0.257	0.0740	28.80
Central	Kheri	128	0.295	0.0717	24.30	0.288	0.0588	20.42
	Sitapur	128	0.324	0.0634	19.57	0.321	0.0570	17.76
	Hardoi	128	0.260	0.0579	22.26	0.287	0.0550	19.18
	Unnao	96	0.566	0.0756	13.36	0.499	0.0669	13.41
	Lucknow	64	0.347	0.0903	26.02	0.326	0.0670	20.55
	Rae Bareli	128	0.367	0.0604	16.46	0.360	0.0548	15.21
	Kanpur Dehat	64	0.152	0.0971	63.89	0.183	0.0543	29.68
	Kanpur Nagar	64	0.115	0.0473	41.16	0.161	0.0488	30.30
	Fatehpur	96	0.520	0.0722	13.88	0.453	0.0618	13.64
Southern	Jalaun	64	0.213	0.0760	35.69	0.236	0.0578	24.49
	Jhansi	64	0.117	0.0573	48.95	0.187	0.0506	27.06
	Lalitpur	32	0.144	0.0751	52.15	0.260	0.0752	28.91
	Hamirpur	32	0.169	0.0939	55.54	0.201	0.0659	32.78
	Banda	64	0.486	0.1038	21.35	0.434	0.0773	17.82
	Chitrakoot	32	0.204	0.1213	59.47	0.290	0.0780	26.89
	Mahoba	32	0.349	0.1693	48.51	0.295	0.0828	28.08

Eastern	Mahrajganj	96	0.354	0.0757	21.38	0.354	0.0652	18.42
	Pratapgarh	128	0.451	0.0711	15.77	0.403	0.0604	14.99
	Kaushambi	63	0.450	0.0863	19.17	0.430	0.0734	17.07
	Allahabad	128	0.244	0.0664	27.22	0.242	0.0517	21.35
	Barabanki	96	0.501	0.0895	17.86	0.437	0.0711	16.28
	Faizabad	64	0.287	0.0884	30.81	0.295	0.0677	22.97
	Ambedkar Nagar	96	0.310	0.0639	20.60	0.303	0.0535	17.65
	Sultanpur	128	0.210	0.0512	24.37	0.221	0.0449	20.34
	Bahraich	96	0.488	0.0873	17.90	0.437	0.0716	16.39
	Shrawasti	64	0.359	0.1008	28.09	0.364	0.0797	21.89
	Balrampur	63	0.196	0.0773	39.42	0.257	0.0678	26.39
	Gonda	128	0.274	0.0625	22.82	0.277	0.0544	19.64
	Siddharthnagar	96	0.263	0.0616	23.44	0.295	0.0583	19.77
	Basti	96	0.578	0.0746	12.91	0.506	0.0645	12.75
	Sant Kabir Nagar	64	0.325	0.0765	23.53	0.320	0.0634	19.81
	Gorakhpur	128	0.283	0.0563	19.89	0.275	0.0504	18.33
	Kushinagar	128	0.214	0.0567	26.52	0.238	0.0511	21.47
	Deoria	96	0.347	0.0746	21.49	0.322	0.0597	18.56
	Azamgarh	128	0.322	0.0585	18.16	0.315	0.0517	16.40
	Mau	64	0.146	0.0582	39.88	0.198	0.0539	27.24
	Ballia	96	0.267	0.0738	27.62	0.232	0.0564	24.31
	Jaunpur	128	0.177	0.0443	25.01	0.216	0.0473	21.91
	Ghazipur	128	0.236	0.0509	21.56	0.248	0.0477	19.25
	Chandauli	64	0.190	0.0663	34.87	0.205	0.0530	25.86
	Varanasi	96	0.192	0.0546	28.43	0.170	0.0415	24.40
	Sant Ravidas Nr	64	0.506	0.0880	17.40	0.380	0.0673	17.71
	Mirzapur	96	0.237	0.0522	22.05	0.247	0.0509	20.60
	Sonbhadra	64	0.375	0.0854	22.77	0.382	0.0698	18.27

Table A3: District-wise sample size, direct estimates (Direct) and small area estimates (SAE) along with their standard error (SE) and percentage coefficient of variations (CV) of MPCE for urban areas of Uttar Pradesh in 2011-12

Region	District	Sample size	Direct			SAE		
			MPCE	SE	CV	MPCE	SE	CV
Western	Saharanpur	64	2118	262	12.39	2047	247	12.06
	Muzaffarnagar	64	2057	184	8.95	2012	179	8.87
	Bijnor	64	1405	127	9.02	1397	125	8.94
	Moradabad	64	1363	94	6.89	1360	93	6.85
	Rampur	32	988	69	6.99	988	69	6.96
	Jyotiba Phule Nr	32	2108	328	15.58	1945	300	15.44
	Meerut	96	2401	225	9.35	2334	215	9.2
	Baghpat	32	2290	205	8.97	2219	198	8.91
	Ghaziabad	96	4180	504	12.06	3439	416	12.1
	Gautam Buddha Nr	32	6453	609	9.44	4762	482	10.12
	Bulandshahar	64	1803	229	12.68	1756	218	12.43
	Aligarh	64	2009	173	8.61	1968	168	8.55
	Hathras	32	1335	116	8.67	1336	114	8.57
	Mathura	64	1445	103	7.14	1446	102	7.06
	Agra	96	1714	373	21.77	1715	333	19.39
	Firozabad	64	1229	87	7.07	1236	86	6.99

	Etah	32	2354	320	13.6	2191	294	13.4
	Mainpuri	32	1026	78	7.61	1030	78	7.54
	Budaun	32	1234	80	6.49	1229	80	6.48
	Bareilly	64	1311	80	6.12	1313	80	6.08
	Pilibhit	32	1419	164	11.58	1410	161	11.39
	Shahjahanpur	32	1175	100	8.49	1175	99	8.42
	Farrukhabad	32	1150	92	8.04	1157	92	7.92
	Kannauj	32	1027	80	7.82	1035	80	7.72
	Etawah	32	1118	102	9.13	1130	101	8.96
	Auraiya	32	1401	122	8.7	1412	120	8.53
	Kashiramnagar	32	1158	90	7.77	1158	89	7.72
Central	Kheri	32	894	87	9.72	902	86	9.56
	Sitapur	32	1400	261	18.64	1410	246	17.43
	Hardoi	32	1046	78	7.45	1051	77	7.37
	Unnao	32	1273	126	9.88	1285	124	9.65
	Lucknow	128	2318	296	12.79	2296	276	12.02
	Rae Bareli	32	1742	350	20.11	1756	316	18
	Kanpur Dehat	32	1499	129	8.62	1509	127	8.43
	Kanpur Nagar	128	1956	162	8.29	1966	159	8.07
	Fatehpur	32	1214	127	10.45	1229	125	10.17
Southern	Jalaun	32	1659	174	10.47	1659	169	10.18
	Jhansi	64	2507	562	22.42	2407	451	18.74
	Lalitpur	32	1620	108	6.66	1629	107	6.55
	Hamirpur	32	1437	155	10.78	1457	152	10.41
	Banda	32	1120	68	6.08	1127	68	6.02
	Chitrakoot	32	791	65	8.18	796	64	8.1
	Mahoba	32	1179	87	7.39	1184	87	7.31
Eastern	Mahrajganj	32	1328	167	12.58	1335	163	12.21
	Pratapgarh	32	1458	186	12.78	1477	181	12.23
	Kaushambi	32	867	79	9.11	878	79	8.95
	Allahabad	63	3436	564	16.41	2940	450	15.3
	Barabanki	32	911	99	10.83	923	98	10.59
	Faizabad	32	1632	310	19.01	1668	286	17.13
	Ambedkar Nagar	32	868	70	8.03	875	69	7.92
	Sultanpur	31	1847	277	15	1832	260	14.17
	Bahraich	32	1313	183	13.93	1313	178	13.52
	Shrawasti	30	1224	196	16	1190	190	15.97
	Balrampur	32	1076	90	8.36	1077	89	8.29
	Gonda	32	2488	207	8.32	2414	199	8.25
	Siddharthnagar	32	1178	145	12.33	1186	143	12.02
	Basti	32	1371	159	11.62	1375	156	11.34
	Sant Kabir Nagar	32	1153	165	14.34	1173	161	13.74
	Gorakhpur	64	1820	172	9.45	1820	168	9.21
	Kushinagar	32	1376	180	13.09	1368	175	12.79
	Deoria	32	1306	163	12.5	1306	160	12.23
	Azamgarh	32	1734	320	18.44	1719	293	17.03
	Mau	32	1210	132	10.93	1235	130	10.54
	Ballia	32	1348	151	11.23	1361	148	10.89
	Jaunpur	32	1522	231	15.15	1513	220	14.56
	Ghazipur	32	1280	143	11.13	1288	140	10.87
	Chandauli	32	2875	377	13.13	2552	336	13.15
	Varanasi	96	1572	127	8.11	1585	126	7.93
	Sant Ravidas Nr.	32	902	61	6.72	905	60	6.67
	Mirzapur	32	1169	157	13.41	1201	153	12.77
	Sonbhadra	31	2039	169	8.3	2021	165	8.17

Table A4: District-wise sample size, direct estimates (Direct) and small area estimates (SAE) along with their standard error (SE) and percentage coefficient of variations (CV) of poverty incidence for urban areas of Uttar Pradesh in 2011-12

Region	District	Sample size	Direct			SAE		
			Poverty Incidence	SE	CV	Poverty Incidence	SE	CV
Western	Saharanpur	64	0.153	0.0603	39.43	0.165	0.0407	24.69
	Muzaffarnagar	64	0.166	0.0621	37.43	0.179	0.0407	22.76
	Bijnor	64	0.199	0.0614	30.85	0.206	0.0490	23.78
	Moradabad	64	0.231	0.0698	30.21	0.235	0.0473	20.14
	Rampur	32	0.576	0.1144	19.86	0.547	0.0918	16.78
	Jyotiba Phule Nr	32	0.148	0.0678	45.83	0.175	0.0595	34.00
	Meerut	96	0.035	0.0184	52.68	0.050	0.0187	37.42
	Baghpat	32	0.068	0.0395	58.08	0.110	0.0492	44.72
	Ghaziabad	96	0.010	0.0063	63.25	0.023	0.0100	43.48
	Gautam B Nr	32	0.010	0.0055	54.77	0.026	0.0138	53.02
	Bulandshahar	64	0.104	0.0493	47.40	0.116	0.0339	29.23
	Aligarh	64	0.113	0.0381	33.70	0.129	0.0392	30.42
	Hathras	32	0.226	0.0938	41.51	0.244	0.0683	28.01
	Mathura	64	0.201	0.0587	29.18	0.208	0.0382	18.37
	Agra	96	0.222	0.0555	25.00	0.224	0.0373	16.64
	Firozabad	64	0.314	0.0696	22.18	0.305	0.0424	13.91
	Etah	32	0.103	0.0507	49.22	0.124	0.0508	40.96
	Mainpuri	32	0.417	0.1053	25.24	0.401	0.0896	22.35
	Budaun	32	0.219	0.0764	34.89	0.256	0.0828	32.35
	Bareilly	64	0.114	0.0498	43.68	0.127	0.0316	24.90
	Pilibhit	32	0.156	0.0796	51.00	0.180	0.0663	36.81
	Shahjahanpur	32	0.296	0.1018	34.39	0.314	0.0841	26.78
	Farrukhabad	32	0.333	0.1003	30.13	0.322	0.0707	21.96
	Kannauj	32	0.433	0.0970	22.40	0.415	0.0625	15.07
	Etawah	32	0.474	0.1001	21.12	0.433	0.0756	17.47
	Auraiya	32	0.132	0.0570	43.19	0.134	0.0487	36.33
	Kashiramnagar	32	0.399	0.1081	27.10	0.399	0.0815	20.44
Central	Kheri	32	0.630	0.1039	16.50	0.579	0.0822	14.19
	Sitapur	32	0.385	0.1321	34.30	0.344	0.0923	26.83
	Hardoi	32	0.486	0.0999	20.57	0.466	0.0727	15.61
	Unnao	32	0.293	0.0948	32.36	0.294	0.0713	24.24
	Lucknow	128	0.160	0.0437	27.31	0.161	0.0253	15.71
	Rae Bareli	32	0.329	0.1356	41.22	0.302	0.0904	29.95
	Kanpur Dehat	32	0.158	0.0802	50.75	0.160	0.0575	35.96
	Kanpur Nagar	128	0.102	0.0430	42.17	0.106	0.0212	20.01
	Fatehpur	32	0.365	0.1012	27.74	0.359	0.0663	18.48
Southern	Jalaun	32	0.092	0.0775	84.20	0.105	0.0438	41.73
	Jhansi	64	0.149	0.0497	33.36	0.146	0.0295	20.20
	Lalitpur	32	0.021	0.0155	73.77	0.043	0.0221	51.48
	Hamirpur	32	0.243	0.0834	34.33	0.231	0.0541	23.43
	Banda	32	0.414	0.1063	25.67	0.396	0.0736	18.59
	Chitrakoot	32	0.600	0.1041	17.34	0.551	0.0957	17.36
	Mahoba	32	0.291	0.1002	34.43	0.284	0.0752	26.47

Eastern	Mahrajganj	32	0.386	0.1033	26.77	0.373	0.0852	22.84
	Pratapgarh	32	0.395	0.1098	27.80	0.377	0.0804	21.34
	Kaushambi	32	0.609	0.1071	17.59	0.579	0.0692	11.95
	Allahabad	63	0.121	0.0370	30.59	0.127	0.0341	26.82
	Barabanki	32	0.736	0.0882	11.98	0.667	0.0802	12.03
	Faizabad	32	0.188	0.0801	42.62	0.192	0.0643	33.51
	Ambedkar Nr	32	0.654	0.0932	14.25	0.604	0.0697	11.54
	Sultanpur	31	0.212	0.0834	39.32	0.201	0.0667	33.19
	Bahraich	32	0.137	0.0845	61.68	0.177	0.0656	37.05
	Shrawasti	30	0.460	0.0991	21.54	0.463	0.0887	19.16
	Balrampur	32	0.388	0.1064	27.43	0.387	0.0859	22.20
	Gonda	32	0.015	0.0110	73.03	0.073	0.0367	50.33
	Siddharthnagar	32	0.340	0.0982	28.88	0.337	0.0856	25.39
	Basti	32	0.395	0.1015	25.69	0.372	0.0856	23.00
	Sant Kabir Nr	32	0.477	0.1067	22.36	0.445	0.0785	17.64
	Gorakhpur	64	0.113	0.0475	42.07	0.122	0.0387	31.75
	Kushinagar	32	0.504	0.1044	20.71	0.473	0.0884	18.70
	Deoria	32	0.390	0.1317	33.76	0.382	0.1217	31.87
	Azamgarh	32	0.249	0.0796	31.95	0.252	0.0559	22.17
	Mau	32	0.301	0.0975	32.40	0.276	0.0660	23.92
	Ballia	32	0.341	0.0922	27.04	0.335	0.0685	20.44
	Jaunpur	32	0.271	0.0917	33.84	0.264	0.0901	34.11
	Ghazipur	32	0.318	0.1020	32.07	0.295	0.0868	29.42
	Chandauli	32	0.003	0.0032	105.4	0.065	0.0370	56.94
	Varanasi	96	0.131	0.0407	31.10	0.132	0.0239	18.09
	Sant Ravidas Nr	32	0.640	0.0964	15.06	0.602	0.0745	12.38
	Mirzapur	32	0.571	0.1061	18.58	0.537	0.0699	13.01
	Sonbhadra	31	0.018	0.0184	102.4	0.073	0.0382	52.34

Table A5: District-wise sample size, direct estimates (Direct) and small area estimates (SAE) along with their standard error (SE) and percentage coefficient of variations (CV) of worker population ratio for rural areas of Uttar Pradesh in 2011-12

Region	District	Sample size	Direct			SAE		
			WPR	SE	CV	WPR	SE	CV
Western	Saharanpur	557	287	29.65	10.33	307	23.02	7.50
	Muzaffarnagar	785	291	25.08	8.62	303	20.49	6.76
	Bijnor	557	321	30.99	9.66	327	23.66	7.24
	Moradabad	768	343	24.19	7.05	335	20.49	6.12
	Rampur	407	321	34.27	10.68	317	24.70	7.79
	Jyotiba Phule Nr.	365	260	34.02	13.09	296	24.49	8.28
	Meerut	354	349	37.80	10.83	333	25.69	7.71
	Baghpat	184	224	48.55	21.68	280	28.46	10.16
	Ghaziabad	392	272	34.56	12.71	298	24.49	8.22
	Gautam Buddha Nr.	171	360	50.61	14.06	323	28.98	8.97
	Bulandshahar	500	272	26.15	9.61	303	21.21	7.00
	Aligarh	536	399	38.81	9.73	363	26.65	7.34
	Hathras	348	413	42.74	10.35	357	28.28	7.92
	Mathura	412	249	31.07	12.48	291	23.24	7.99
	Agra	582	297	29.04	9.78	307	22.14	7.21
	Firozabad	379	286	34.30	11.99	299	24.49	8.19
	Etah	346	366	42.11	11.51	328	27.02	8.24
	Mainpuri	387	274	43.95	16.04	301	27.20	9.04
	Budaun	539	323	35.71	11.06	322	25.30	7.86

	Bareilly	515	341	33.11	9.71	328	24.90	7.59
	Pilibhit	349	432	44.95	10.41	358	28.64	8.00
	Shahjahanpur	541	355	39.04	11.00	335	26.65	7.95
	Farrukhabad	400	252	36.51	14.49	290	25.69	8.86
	Kannauj	379	328	43.88	13.38	320	27.57	8.62
	Etawah	334	294	43.13	14.67	320	27.57	8.62
	Auraiya	315	294	39.09	13.30	323	27.57	8.54
	Kashiramnagar	205	282	44.25	15.69	306	27.39	8.95
Central	Kheri	686	363	34.33	9.46	364	26.83	7.37
	Sitapur	701	359	28.38	7.90	366	24.08	6.58
	Hardoi	727	343	27.67	8.07	355	23.24	6.55
	Unnao	470	475	36.25	7.63	430	27.57	6.41
	Lucknow	350	373	36.53	9.79	369	28.11	7.62
	Rae Bareli	700	377	26.45	7.01	377	22.58	5.99
	Kanpur Dehat	283	350	51.15	14.61	347	29.66	8.55
	Kanpur Nr	299	396	46.09	11.64	369	29.15	7.90
	Fatehpur	445	340	33.62	9.89	362	26.27	7.26
Southern	Jalaun	349	277	34.91	12.60	325	26.08	8.02
	Jhansi	286	409	42.21	10.32	402	30.66	7.63
	Lalitpur	134	324	63.01	19.45	353	34.21	9.69
	Hamirpur	160	308	63.02	20.46	350	33.47	9.56
	Banda	390	378	44.12	11.67	367	28.98	7.90
	Chitrakoot	145	507	66.83	13.18	398	34.35	8.63
	Mahoba	145	423	79.24	18.73	378	36.06	9.54
Eastern	Mahrajganj	549	360	32.45	9.01	355	24.90	7.01
	Pratapgarh	765	369	30.08	8.15	356	23.02	6.47
	Kaushambi	308	373	40.06	10.74	380	29.33	7.72
	Allahabad	745	351	30.60	8.72	359	24.70	6.88
	Barabanki	595	395	37.95	9.61	383	27.20	7.10
	Faizabad	345	386	41.20	10.67	357	27.39	7.67
	Ambedkar Nr.	546	381	28.85	7.57	362	22.80	6.30
	Sultanpur	708	371	26.19	7.06	361	21.68	6.01
	Bahraich	488	281	36.15	12.86	310	26.65	8.60
	Shrawasti	317	270	42.64	15.79	308	27.75	9.01
	Balrampur	310	300	43.37	14.46	314	27.93	8.89
	Gonda	705	255	27.93	10.95	289	22.36	7.74
	Siddharthnagar	541	350	30.69	8.77	337	23.87	7.08
	Basti	534	410	32.60	7.95	371	24.29	6.55
	Sant Kabir Nr.	394	265	30.10	11.36	291	22.80	7.84
	Gorakhpur	777	286	23.12	8.08	301	20.25	6.73
	Kushinagar	727	314	27.07	8.62	317	22.14	6.98
	Deoria	584	328	30.45	9.28	316	22.80	7.22
	Azamgarh	896	290	22.18	7.65	304	20.49	6.74
	Mau	374	261	31.35	12.01	294	24.90	8.47
	Ballia	630	333	33.79	10.15	324	24.70	7.62
	Jaunpur	796	376	26.43	7.03	361	22.14	6.13
	Ghazipur	816	364	24.76	6.80	354	20.74	5.86
	Chandauli	395	420	37.56	8.94	373	26.65	7.14
	Varanasi	636	410	28.27	6.89	372	22.58	6.07
	Sant Ravidas Nr.	467	265	29.50	11.13	288	23.45	8.14
	Mirzapur	566	360	27.94	7.76	358	22.80	6.37
	Sonbhadra	347	400	36.46	9.11	383	27.02	7.05

Table A6: District-wise sample size, direct estimates (Direct) and small area estimates (SAE) along with their standard error (SE) and percentage coefficient of variations (CV) of worker population ratio for urban areas of Uttar Pradesh in 2011-12

Region	District	Sample size	Direct			SAE		
			WPR	SE	CV	WPR	SE	CV
Western	Saharanpur	298	332	34.90	10.51	313	18.71	5.98
	Muzaffarnagar	322	294	30.04	10.22	315	18.97	6.02
	Bijnor	354	281	31.08	11.06	313	19.49	6.23
	Moradabad	349	302	29.97	9.92	330	20.25	6.14
	Rampur	171	473	54.68	11.56	393	26.83	6.83
	Jyotiba Phule Nr.	163	336	44.27	13.18	337	20.98	6.22
	Meerut	470	296	30.56	10.32	305	18.44	6.05
	Baghpat	126	268	47.82	17.84	311	19.75	6.35
	Ghaziabad	402	321	30.59	9.53	312	18.97	6.08
	Gautam Buddha Nr.	117	309	52.71	17.06	350	27.02	7.72
	Bulandshahar	289	322	32.38	10.06	314	19.24	6.13
	Aligarh	339	247	31.45	12.73	296	19.24	6.50
	Hathras	194	315	39.27	12.47	287	20.74	7.23
	Mathura	336	294	28.56	9.71	302	18.17	6.02
	Agra	509	298	30.20	10.13	295	20.98	7.11
	Firozabad	390	303	28.53	9.42	317	19.49	6.15
	Etah	159	321	43.45	13.54	294	19.24	6.54
	Mainpuri	151	306	44.39	14.51	270	20.25	7.50
	Budaun	194	247	38.66	15.65	316	22.36	7.08
	Bareilly	315	390	35.20	9.03	377	23.02	6.11
	Pilibhit	183	368	46.58	12.66	346	21.45	6.20
	Shahjahanpur	170	334	45.43	13.60	334	20.98	6.28
	Farrukhabad	199	261	35.81	13.72	305	18.71	6.13
	Kannauj	191	386	39.50	10.23	346	20.49	5.92
	Etawah	161	261	38.46	14.74	274	18.97	6.92
	Auraiya	148	226	38.16	16.88	260	19.49	7.50
	Kashiramnagar	186	391	42.69	10.92	338	20.98	6.21
Central	Kheri	190	302	40.70	13.48	323	19.75	6.11
	Sitapur	160	367	56.16	15.30	329	20.00	6.08
	Hardoi	151	394	44.86	11.39	329	19.75	6.00
	Unnao	158	313	46.56	14.88	312	19.24	6.17
	Lucknow	656	367	30.46	8.30	340	20.49	6.03
	Rae Bareli	152	304	56.98	18.74	303	19.24	6.35
	Kanpur Dehat	159	339	50.02	14.75	299	19.24	6.43
	Kanpur Nr.	548	308	30.54	9.92	320	20.00	6.25
	Fatehpur	174	247	35.13	14.22	297	18.44	6.21
Southern	Jalaun	157	275	44.29	16.10	276	20.25	7.34
	Jhansi	272	311	35.65	11.46	302	21.45	7.10
	Lalitpur	163	262	39.92	15.24	305	19.75	6.47
	Hamirpur	134	270	42.39	15.70	294	20.25	6.89
	Banda	125	288	49.17	17.07	297	19.24	6.48
	Chitrakoot	153	286	46.58	16.29	274	20.49	7.48
	Mahoba	143	264	45.44	17.21	294	19.24	6.54
Eastern	Mahrajganj	162	298	44.59	14.96	290	19.24	6.63
	Pratapgarh	187	330	46.89	14.21	315	19.75	6.27
	Kaushambi	154	311	47.11	15.15	322	21.91	6.80
	Allahabad	267	259	35.20	13.59	309	20.49	6.63
	Barabanki	154	344	51.86	15.08	342	20.98	6.13
	Faizabad	166	335	51.88	15.49	319	20.25	6.35

	Ambedkar Nr.	201	328	36.72	11.19	318	19.24	6.05
	Sultanpur	176	269	40.93	15.22	294	20.00	6.80
	Bahraich	161	338	48.61	14.38	340	21.45	6.31
	Shrawasti	155	270	37.74	13.98	316	22.58	7.15
	Balrampur	168	328	44.89	13.69	335	21.45	6.40
	Gonda	125	280	47.65	17.02	315	20.49	6.51
	Siddharthnagar	152	392	50.02	12.76	326	19.75	6.06
	Basti	162	304	43.75	14.39	272	20.49	7.53
	Sant Kabir Nr.	158	285	41.50	14.56	305	18.97	6.22
	Gorakhpur	344	242	33.07	13.66	293	19.24	6.56
	Kushinagar	179	359	41.52	11.57	315	20.98	6.66
	Deoria	177	222	47.75	21.51	276	21.21	7.69
	Azamgarh	202	293	34.59	11.80	305	18.71	6.13
	Mau	168	402	51.15	12.72	342	21.45	6.27
	Ballia	192	293	36.41	12.43	310	19.49	6.29
	Jaunpur	205	311	45.03	14.48	294	20.74	7.05
	Ghazipur	192	266	39.16	14.72	287	19.75	6.88
	Chandauli	166	300	46.43	15.48	290	19.24	6.63
	Varanasi	541	372	27.93	7.51	352	20.25	5.75
	Sant Ravidas Nr.	179	487	42.23	8.67	338	20.00	5.92
	Mirzapur	157	299	44.96	15.04	320	20.00	6.25
	Sonbhadra	144	319	44.26	13.88	290	20.25	6.98

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 173–186

Some Existence on Ordered Multi-designs

Kazuki Matsubara[1] **and Sanpei Kageyama**[2]
[1]*Faculty of Commerce, Chuo Gakuin University, Abiko, Japan*
[2]*Hiroshima University, Higashi-Hiroshima, Japan*

Received: 21 December 2020; Revised: 20 January 2021; Accepted: 24 January 2021

Abstract

Two variants of an orthogonal array, orthogonal arrays of type I and of type II, were introduced by Rao in 1961. Furthermore, as generalizations of an orthogonal array and an orthogonal array of type II, an orthogonal multi-array and a perpendicular multi-array have been introduced by Brickell in 1984 and by Li *et al.* in 2018, respectively. In this paper, as a generalization of the orthogonal array of type I, an ordered multi-design is newly introduced from a combinatorial viewpoint. Necessary conditions for the existence of an ordered multi-design are discussed and several constructions of the ordered multi-design are provided by use of group divisible designs and self-orthogonal latin squares, through a difference technique. As main results, the existence of a family of ordered multi-designs is provided and also the sufficiency of necessary conditions for existence is shown for a class of ordered multi-designs with one possible exception.

Key words: Ordered multi-design; Perpendicular multi-array; Self-orthogonal Latin square; Group divisible design.

AMS Subject Classifications: 05B15, 05B05

1. Introduction

An *ordered multi-design* of size $N \times k$, denoted by $\mathrm{OMD}_\lambda(k \times c, v)$, is an $N \times k$ multi-array, $\mathcal{A} = (A_{ij})$, on a set V of v points, which satisfies the following conditions:

(C1) each entry A_{ij} ($|A_{ij}| = c$) is a c-subset of V and kc distinct points occur in k entries of each row of $\mathcal{A}$, and
(C2) for any ordered pair (j_1, j_2) of integers with $1 \le j_1 < j_2 \le k$ and for any ordered pair (x_1, x_2) of distinct points in V, there are exactly λ rows of $\mathcal{A}$ such that the points x_1 and x_2 appear in the j_1th and the j_2th entries, *i.e.*, in the j_1th and the j_2th columns, of each of the λ rows, respectively.

Note that the conditions (C1) and (C2) lead to $N = \lambda v(v-1)/(c^2)$. Moreover, $k \ge 2$ is assumed at least to validate the condition (C2).

Let us illustrate the definition of the $\mathrm{OMD}_\lambda(k \times c, v)$ by the following example.

Corresponding Author: Kazuki Matsubara
Email: additivestructure@gmail.com

Example 1: An $\mathrm{OMD}_2(3 \times 2, 6)$ on $V = \mathbb{Z}_5 \cup \{\infty\}$ is given by

$$\left(\begin{array}{c|c|c} \infty, 0 & 1, 4 & 2, 3 \\ \infty, 1 & 2, 0 & 3, 4 \\ \infty, 2 & 3, 1 & 4, 0 \\ \infty, 3 & 4, 2 & 0, 1 \\ \infty, 4 & 0, 3 & 1, 2 \\ 2, 3 & \infty, 0 & 1, 4 \\ 3, 4 & \infty, 1 & 2, 0 \\ 4, 0 & \infty, 2 & 3, 1 \\ 0, 1 & \infty, 3 & 4, 2 \\ 1, 2 & \infty, 4 & 0, 3 \\ 1, 4 & 2, 3 & \infty, 0 \\ 2, 0 & 3, 4 & \infty, 1 \\ 3, 1 & 4, 0 & \infty, 2 \\ 4, 2 & 0, 1 & \infty, 3 \\ 0, 3 & 1, 2 & \infty, 4 \end{array}\right)$$

with $k = 3$ (three columns), $c = 2$, $N = 15$ (fifteen rows), entries of the first row $A_{11} = \{\infty, 0\}$, $A_{12} = \{1, 4\}$, $A_{13} = \{2, 3\}$, entries of the second row $A_{21} = \{\infty, 1\}$, $A_{22} = \{2, 0\}$, $A_{23} = \{3, 4\}, \ldots$, entries of the sixth row $A_{61} = \{2, 3\}$, $A_{62} = \{\infty, 0\}$, $A_{63} = \{1, 4\}$, etc. The condition (C2) with $\lambda = 2$ can be checked, *e.g.*, 0 and 1 occur in the first and the second columns, respectively, of the first and the last rows.

From now on, each row of an $\mathrm{OMD}_\lambda(k \times c, v)$ is separately displayed in the form of

$$(a_{11}, a_{12}, \ldots, a_{1c} \mid a_{21}, a_{22}, \ldots, a_{2c} \mid \ldots \mid a_{k1}, a_{k2}, \ldots, a_{kc})$$

by use of kc points on V or $(A_{i1} \mid A_{i2} \mid \ldots \mid A_{ik})$ by use of k entries A_{ij} $(1 \leq i \leq N)$.

It is clear that the $\mathrm{OMD}_\lambda(k \times 1, v)$ coincides with the ordered design, denoted by $\mathrm{OD}_\lambda(k, v)$, defined in Rao (1961), who call the ordered design by the other name "an orthogonal array of Type I". An orthogonal array and a perpendicular array (called by the other name "an orthogonal array of Type II" in Rao, 1961) have been generalized to an orthogonal multi-array (OMA) in Brickell (1984) and a perpendicular multi-array (PMA) in Li *et al.* (2018), respectively. Furthermore, applications of the OMA and the PMA to design of experiments and coding theory are discussed in Brickell (1984), Li *et al.* (2015), Li *et al.* (2018), Mukerjee (1998) and Sitter (1993). On the other hand, as far as the authors know, the ordered multi-design has never been discussed in literature.

In this paper, the existence on an $\mathrm{OMD}_\lambda(k \times 2, v)$, *i.e.*, $c = 2$, is mainly discussed from a viewpoint of combinatorics. In Section 2, a construction and a fundamental property of the OMD, and combinatorial structures used in later sections are presented. In Section 3, necessary conditions for the existence of an $\mathrm{OMD}_\lambda(k \times c, v)$ are discussed. In Section 4, constructions of a cyclic $\mathrm{OMD}_\lambda(3 \times 2, v)$ are provided by use of difference techniques. In Sections 5 and 6, methods of constructing an OMD are presented by use of a group divisible design (GDD) and self-orthogonal latin squares (SOLS), respectively. In Section 7, the existence of an $\mathrm{OMD}_\lambda(k \times 2, q)$ for any prime power q is provided. Furthermore, it is shown that the necessary conditions discussed in Section 3 are also sufficient for the existence of an $\mathrm{OMD}_\lambda(3 \times 2, v)$ with one possible exception, as in the following main results.

Theorem 1: There exists an $\mathrm{OMD}_\lambda(k \times 2, q)$ for any prime power q, any $\lambda \equiv 0 \pmod 2$ and any k with $2 \le k \le \lceil (q-1)/2 \rceil$.

Theorem 2: Let v be a positive integer with $v \ge 6$. Then there exists an $\mathrm{OMD}_\lambda(3 \times 2, v)$ if and only if $v \equiv 1 \pmod 4$ or $\lambda \equiv 0 \pmod 2$ with a possible exception of $(v, \lambda) = (9, 1)$.

As the appendix, some individual examples, which cannot be obtained by methods in this paper, will be presented to be utilized in the proof of Theorem 2 in Section 7.

2. Preliminaries

At first, the perpendicular multi-array discussed in Li *et al.* (2018) and Matsubara and Kageyama (2021) is reviewed. The perpendicular multi-array $\mathcal{A} = (A_{ij})$, denoted by $\mathrm{PMA}_\lambda(k \times c, v)$, is defined by the condition (C1) and the following condition (C3):

(C3) for any two columns of $\mathcal{A}$ and for any unordered pair $\{x_1, x_2\}$ of distinct points in V, there are exactly λ rows of $\mathcal{A}$ such that the points x_1 and x_2 separately appears in the two entries of each of the λ rows.

Since the condition (C2) involves the condition (C3), it follows that any $\mathrm{OMD}_\lambda(k \times c, v)$ can be regarded as a $\mathrm{PMA}_{2\lambda}(k \times c, v)$.

On the other hand, it is known (see Bierbrauer, 2007) that the existence of an $\mathrm{OD}_1(k, v)$, *i.e.*, $\mathrm{OMD}_1(k \times 1, v)$, is equivalent to the existence of $k-2$ idempotent mutually orthogonal latin squares. The review of results on the existence and applications of the $\mathrm{OD}_\lambda(k, v)$ can be found in Bierbrauer (2007), Bierbrauer and Edel (1994), Kunert and Martin (2000) and Majumdar and Martin (2004). Especially, the following result will be useful for the construction of an $\mathrm{OMD}_\lambda(k \times c, v)$ described in Section 5.

Lemma 1 (Bierbrauer, 2007)**:** There exists an $\mathrm{OD}_1(k, k)$ for any prime power k.

A direct construction of an $\mathrm{OMD}_2(k \times 2, v)$ can be obtained as follows.

Lemma 2: Let q be an odd prime power. Then there exists an $\mathrm{OMD}_2(k \times 2, q)$ with $k = (q-1)/2$.

Proof: Let $V = GF(q)$. Then a direct sum decomposition of $GF(q)$ can be given by

$$GF(q) = \{0\} \cup B_1 \cup B_2 \cup \ldots \cup B_{\frac{q-1}{2}},$$

where $B_j = \{a_j, -a_j\}$ with $a_j \in GF(q)$ $(1 \le j \le (q-1)/2)$. Now consider $(q-1)/2$ rows:

$$(\alpha^\ell B_1 \mid \alpha^\ell B_2 \mid \ldots \mid \alpha^\ell B_{\frac{q-1}{2}}),\ 1 \le \ell \le \frac{q-1}{2},$$

where $\alpha^\ell B_j = \{\alpha^\ell a_j, -\alpha^\ell a_j\}$ and α is a primitive element of $GF(q)$. Hence any two entries $\{\alpha^\ell a_{j_1}, -\alpha^\ell a_{j_1}\}, \{\alpha^\ell a_{j_2}, -\alpha^\ell a_{j_2}\}$ in the same row yield four pairs as

$$(\alpha^\ell a_{j_1}, \alpha^\ell a_{j_2}), (-\alpha^\ell a_{j_1}, -\alpha^\ell a_{j_2}), (\alpha^\ell a_{j_1}, -\alpha^\ell a_{j_2}), (-\alpha^\ell a_{j_1}, \alpha^\ell a_{j_2})$$

for the condition (C2). Furthermore, for any pair (x, y) it holds that

$$\{(x+t, y+t) \mid t \in GF(q)\} = \{(x', y') \mid x', y' \in GF(q), x' - y' = x - y\}.$$

Since $\alpha^{(q-1)/2} = -1$ and $\{\alpha^\ell, -\alpha^\ell \mid 1 \leq \ell \leq (q-1)/2\} = GF(q)\backslash\{0\}$, both of

$$\{(\alpha^\ell a_{j_1} + t, \alpha^\ell a_{j_2} + t), (-\alpha^\ell a_{j_1} + t, -\alpha^\ell a_{j_2} + t) \mid 1 \leq \ell \leq \frac{q-1}{2}, t \in GF(q)\}$$

and

$$\{(\alpha^\ell a_{j_1} + t, -\alpha^\ell a_{j_2} + t), (-\alpha^\ell a_{j_1} + t, \alpha^\ell a_{j_2} + t) \mid 1 \leq \ell \leq \frac{q-1}{2}, t \in GF(q)\}$$

are equal to $\{(x, y) \mid x, y \in GF(q), x \neq y\}$. Therefore the required $\mathrm{OMD}_2(k \times 2, q)$ with $k = (q-1)/2$ can be obtained from the following $(q-1)q/2$ rows:

$$(\alpha^\ell B_1 + t \mid \alpha^\ell B_2 + t \mid \ldots \mid \alpha^\ell B_{\frac{q-1}{2}} + t),\ 1 \leq \ell \leq \frac{q-1}{2},\ t \in GF(q),$$

where $\alpha^\ell B_j + t = \{\alpha^\ell a_j + t, -\alpha^\ell a_j + t\}$. □

Next a fundamental property of the OMD, which is useful to construct OMDs for various values of k, is provided as follows.

Lemma 3: Any subarray obtained by deleting any k' $(k' < k)$ columns of an $\mathrm{OMD}_\lambda(k \times c, v)$ is an $\mathrm{OMD}_\lambda((k-k') \times c, v)$.

Proof: Since an $\mathrm{OMD}_\lambda(k \times c, v)$ satisfies the conditions (C1) and (C2), it is clear that any two columns of the $\mathrm{OMD}_\lambda((k-k') \times c, v)$ also satisfy (C1) and (C2). □

Now, a combinatorial design used in later sections is introduced. Let v, k, λ be positive integers. A *group divisible design*, denoted by (k, λ)-GDD, is a triplet $(V, \mathcal{G}, \mathcal{B})$, where V is a set of v points, $\mathcal{G}$ is a partition of V into subsets (called groups) and $\mathcal{B}$ $(|\mathcal{B}| = b)$ is a family of subsets (called blocks) of size k each of V such that

(G1) every pair of distinct points $x, y \in V$ in different groups occurs in exactly λ blocks, and
(G2) every pair of distinct points $x, y \in V$ in the same group does not occur in any block.

The *group type* of a (k, λ)-GDD is a multi-set $\{|G| \mid G \in \mathcal{G}\}$. The usual exponential notation is used to describe group types. Thus the notation $h_1^{t_1} h_2^{t_2} \cdots h_n^{t_n}$ means that there are t_i groups of size h_i for $1 \leq i \leq n$ (cf. Ge, 2007).

The following proposition on GDDs is known.

Lemma 4 (Ge, 2007)**:** Let g, u and m be non-negative integers. Then there exists a $(3, 1)$-GDD of type $g^u m^1$ if and only if the following conditions are all satisfied:

(a) if $g > 0$, then $u \geq 3$, or $u = 2$ and $m = g$, or $u = 1$ and $m = 0$, or $u = 0$;
(b) $m \leq g(u-1)$ or $gu = 0$;
(c) $g(u-1) + m \equiv 0 \pmod 2$ or $gu = 0$;
(d) $gu \equiv 0 \pmod 2$ or $m = 0$; and
(e) $\frac{1}{2}g^2u(u-1) + gum \equiv 0 \pmod 3$.

The GDD will be utilized for a method of constructing OMDs discussed in Section 7.

3. Necessary Conditions

Necessary conditions for the existence of an $\mathrm{OMD}_\lambda(k\times c, v)$ are considered. It is obvious by the conditions (C1) and (C2) that for any $\mathrm{OMD}_\lambda(k \times c, v)$ of size $N \times k$

$$v \geq kc \tag{1}$$

holds. Since N is a positive integer,

$$c^2 \mid \lambda v(v-1) \tag{2}$$

holds. Furthermore, every point must occur equally $r\ (= cN/v)$ times in each column. Hence it is seen that

$$c \mid \lambda(v-1) \tag{3}$$

holds.

The sufficiency of these necessary conditions (1), (2), (3) for the existence when $(c, v) = (2, q)$ with any prime power q and $(k, c) = (3, 2)$, will be proved with some exceptions as in Theorems 1 and 2, respectively, in Section 7.

Furthermore, another necessary condition for the existence of an $\mathrm{OMD}_\lambda(k \times c, v)$ of size $N \times k$ can be presented by use of the following result.

Lemma 5 (Matsubara and Kageyama, 2021)**:** In a $\mathrm{PMA}_\lambda(k \times c, v)$ of size $N \times k$, it holds that

$$N \geq v - 1.$$

In particular, $N = v - 1$ implies $v = 2c$.

Theorem 3: In an $\mathrm{OMD}_\lambda(k \times c, v)$ of size $N \times k$, it holds that

$$N \geq v. \tag{4}$$

Proof: Since any $\mathrm{OMD}_\lambda(k \times c, v)$ is a $\mathrm{PMA}_{2\lambda}(k \times c, v)$, $N \geq v-1$ holds. For the proof, it is sufficient to show the non-existence of an $\mathrm{OMD}_\lambda(2 \times c, v)$ with $N = v-1$. When $N = v-1$, Lemma 5 implies $v = 2c$, that is, v is even and N is odd. On the other hand, $v = 2c$ and (1) imply that $k = 2$ holds and each point appears in all of N rows of the $\mathrm{OMD}_\lambda(2 \times c, v)$. Hence, each point cannot occur equally in each of the two columns. □

The existence of an $\mathrm{OMD}_1(2 \times c, c^2 + 1)$, which satisfies $N = v = c^2 + 1$ and $k = 2$, for any $c \geq 2$ is known in Matsubara and Kageyama (2021) as a $\mathrm{PMA}_2(2 \times c, c^2 + 1)$. Hence the inequality (4) is best possible when $k = 2$. However, any existence result on an $\mathrm{OMD}_\lambda(k \times c, v)$ with $N = v, k \geq 3$ and $c \geq 2$ is not known in literature as far as the authors know.

The minimality of λ is also discussed here. An $\mathrm{OMD}_\lambda(k \times c, v)$ is said to be *minimal* if there exists no $\mathrm{OMD}_{\lambda'}(k \times c, v)$ for any $\lambda' < \lambda$. Especially, it is clear that any OMD with $N = v$ and any OMD with $\lambda = 1$ are minimal. On the other hand, by taking u copies of each row of $\mathcal{A}$, it is clear that the existence of an $\mathrm{OMD}_\lambda(k \times c, v)$ implies the existence of an $\mathrm{OMD}_{\lambda u}(k \times c, v)$. In fact, the existence of a minimal $\mathrm{OMD}_\lambda(3 \times 2, v)$ plays an important role in Section 7. Some minimal $\mathrm{OMD}_\lambda(k \times 2, v)$ are exhaustively listed within the scope of $4 \leq v \leq 20$ in Table 1 of Appendix.

4. OMD with a Cyclic Automorphism

Combinatorial multi-arrays (OMA, PMA, OMD) are regarded as a pair $(V, \mathcal{R})$ of a point set V and a set $\mathcal{R}$ of rows. When $V = \mathbb{Z}_v$ (or $V = \mathbb{Z}_{v-1} \cup \{\infty\}$) and $\mathcal{R} = \{\boldsymbol{R} + t \mid \boldsymbol{R} \in \mathcal{R}\}$ with $\boldsymbol{R} + t = (a_{11} + t, \ldots, a_{1c} + t \mid \ldots \mid a_{k1} + t, \ldots, a_{kc} + t)$ for any $t \in \mathbb{Z}_v$ (or any $t \in \mathbb{Z}_{v-1}$), the array is said to be *cyclic* (or *1-rotational*, where ∞ is a fixed point with $\infty + t = \infty$ for any $t \in \mathbb{Z}_{v-1}$). Then a *row orbit* of $\boldsymbol{R} \in \mathcal{R}$ is defined by $\{\boldsymbol{R} + t \mid t \in \mathbb{Z}_v\}$ (or $\{\boldsymbol{R} + t \mid t \in \mathbb{Z}_{v-1}\}$). Note that the length of any row orbit on $\mathbb{Z}_v$ is assumed to be v in this paper. Choose an arbitrary row from each row orbit and call it a *base row*. Hence, for a cyclic multi-array, the array can be represented simply by displaying base rows. For example, the $\mathrm{OMD}_2(3 \times 2, 6)$ given in Example 1 is presented by

$$(\infty, 0 \mid 1, 4 \mid 2, 3), (2, 3 \mid \infty, 0 \mid 1, 4), (1, 4 \mid 2, 3 \mid \infty, 0) \mod 5.$$

For two points x and y in the j_1th and the j_2th $(1 \le j_1 < j_2 \le k)$ entries, respectively, of each base row, $x - y \equiv d \pmod v$ implies that in the orbit of the base row there exists a row containing x' and y' in the j_1th and the j_2th entries, respectively, for any distinct points x', y' in $\mathbb{Z}_v$ with $x' - y' \equiv d \pmod v$. Hence, it is seen that the multi-array obtained from orbits on $\mathbb{Z}_v$ of m base rows $(A_{i1}^* \mid \ldots \mid A_{ik}^*)$, $1 \le i \le m$, satisfies the condition (C2) of an $\mathrm{OMD}_\lambda(k \times c, v)$ and the condition (C3) of a $\mathrm{PMA}_\lambda(k \times c, v)$ if

$$\bigcup_{1 \le i \le m} \{d - d' \mid d \in A_{ij_1}^*, d' \in A_{ij_2}^*\} = \lambda\left(\mathbb{Z}_v \backslash \{0\}\right) \tag{5}$$

and

$$\bigcup_{1 \le i \le m} \{\pm(d - d') \mid d \in A_{ij_1}^*, d' \in A_{ij_2}^*\} = \lambda\left(\mathbb{Z}_v \backslash \{0\}\right) \tag{6}$$

holds, respectively, for any j_1, j_2 with $1 \le j_1 < j_2 \le k$, where λS means a multi-set containing each element of the set S exactly λ times. Furthermore, m base rows with a 1-rotational automorphism on $\mathbb{Z}_{v-1} \cup \{\infty\}$ yield a multi-array satisfying the condition (C2) if

$$\bigcup_{1 \le i \le m} \{d - d' \mid d \in A_{ij_1}^*, d' \in A_{ij_2}^*\} = \lambda\left((\mathbb{Z}_{v-1} \cup \{\infty\}) \backslash \{0\}\right), \tag{7}$$

where $\infty - t = t - \infty = \infty$ for any $t \in \mathbb{Z}_{v-1}$.

In fact, it can be checked that the base rows given in Examples 7, 8 (for cyclic OMDs), Examples 3 to 6 (for cyclic PMAs) and Examples 1 and 9 to 12 (for 1-rotational OMDs) satisfy the conditions (5), (6) and (7), respectively.

At first, a direct construction of an $\mathrm{OMD}_2(k \times 2, v)$ is provided as follows.

Lemma 6: Let v be odd and p be the smallest prime factor of v. Then there exists a cyclic $\mathrm{OMD}_2(k \times 2, v)$ with $k = (p-1)/2$.

Proof: Let $\mathcal{R}^*$ be a set of the following $(v-1)/2$ rows:

$$\boldsymbol{R}_t^* = (t, -t \mid 2t, -2t \mid \ldots \mid \frac{p-1}{2}t, -\frac{p-1}{2}t), \ 1 \le t \le \frac{v-1}{2}.$$

Since p is the smallest prime factor of the odd v, $\boldsymbol{R}_t^*$ contains $p-1$ different elements in $\mathbb{Z}_v$ for each t. Moreovre, both $\gcd(j_2-j_1,v)=1$ and $\gcd(j_1+j_2,v)=1$ hold for each j_1,j_2 with $1\le j_1<j_2\le (p-1)/2$. Hence it also holds that

$$\{\pm(j_1t-j_2t)\mid 1\le t\le \frac{v-1}{2}\}=\{\pm(j_1t+j_2t)\mid 1\le t\le \frac{v-1}{2}\}=\mathbb{Z}_v\backslash\{0\}.$$

Since two entries $\{j_1t,-j_1t\}$ and $\{j_2t,-j_2t\}$ yield four differences $\pm(j_1t-j_2t)$ and $\pm(j_1t+j_2t)$, it is shown that $\mathcal{R}^*$ yields the required cyclic $\mathrm{OMD}_2(k\times 2,v)$ with $k=(p-1)/2$. □

Next another method of constructing a cyclic $\mathrm{OMD}_\lambda(k\times c,v)$ from a cyclic $\mathrm{PMD}_\lambda(k\times c,v)$ is presented as follows.

Lemma 7: The existence of a cyclic $\mathrm{PMA}_\lambda(k\times c,v)$ implies the existence of a cyclic $\mathrm{OMD}_\lambda(k\times c,v)$.

Proof: Let a set of m base rows of the cyclic $\mathrm{PMA}_\lambda(k\times c,v)$ be

$$\mathcal{R}^*=\{(A_{i1}^*\mid\ldots\mid A_{ik}^*)\mid 1\le i\le m\}.$$

Then take the set $\mathcal{R}^*\cup\mathcal{R}^{**}$ of rows with

$$\mathcal{R}^{**}=\{(-A_{i1}^*\mid\ldots\mid -A_{ik}^*)\mid 1\le i\le m\}.$$

Since $\mathcal{R}^*$ satisfies (6), $\mathcal{R}^*\cup\mathcal{R}^{**}$ satisfies (5). Hence $\mathcal{R}^*\cup\mathcal{R}^{**}$ yields the required cyclic $\mathrm{OMD}_\lambda(k\times c,v)$. □

For an odd prime p, a cyclic $\mathrm{OMD}_1(k\times 2,p)$ can be constructed when there exists a point set satisfying the following condition on $\mathbb{Z}_p$:

(L) for any distinct points x,y in the set,

$$\left(\frac{x+y}{p}\right)\left(\frac{x-y}{p}\right)=-1,$$

where $\left(\frac{a}{p}\right)$ is the Legendre symbol of a at p.

Lemma 8: Let $p\equiv 1\pmod 4$ be an odd prime and α be a primitive element of $\mathbb{Z}_p$. If there exists a k-set S on $\mathbb{Z}_p$ satisfying the condition (L), then a cyclic $\mathrm{OMD}_1(k\times 2,p)$ exists.

Proof: Let $S=\{a_1,a_2,\ldots,a_k\}$ be a set satisfying (L) on $\mathbb{Z}_p$. Then, for any x,y satisfying (L), it is seen that $\pm(x+y)\alpha^{2t}$ $(1\le t\le (p-1)/4)$ yield a set of quadratic residues or a set of non-quadratic residues, according as $x+y$ is a quadratic residue or not. The same holds for the case of $\pm(x-y)\alpha^{2t}$.

Hence, for any j_1,j_2 with $1\le j_1<j_2\le k$, it holds that

$$\bigcup_{1\le t\le\frac{p-1}{4}}\{\pm(a_{j_1}+a_{j_2})\alpha^{2t},\pm(a_{j_1}-a_{j_2})\alpha^{2t}\}=\mathbb{Z}_p\backslash\{0\}.$$

On the other hand, two entries $\{a_{j_1}\alpha^{2t}, -a_{j_1}\alpha^{2t}\}$ and $\{a_{j_2}\alpha^{2t}, -a_{j_2}\alpha^{2t}\}$ yield four differences $\pm(a_{j_1}+a_{j_2})\alpha^{2t}$ and $\pm(a_{j_1}-a_{j_2})\alpha^{2t}$ for any t with $1 \leq t \leq (p-1)/4$ and $1 \leq j_2 < j_2 \leq k$. Therefore the base rows

$$(a_1\alpha^{2t}, -a_1\alpha^{2t} \mid \cdots \mid a_k\alpha^{2t}, -a_k\alpha^{2t})$$

with $1 \leq t \leq (p-1)/4$ can yield the required $\mathrm{OMD}_1(k \times 2, p)$. □

Examples of such k-set S are presented as follows.

Example 2: The following sets on $\mathbb{Z}_p$ satisfy the condition (L)

$$\{1,3,4\} \text{ on } \mathbb{Z}_{13},\ \{1,2,7\} \text{ on } \mathbb{Z}_{17},\ \{1,2,4\} \text{ on } \mathbb{Z}_{29},\ \{1,4,17\} \text{ on } \mathbb{Z}_{37},\ \{1,7,8\} \text{ on } \mathbb{Z}_{41}.$$

In the case where $\lambda = 2$ and even v, the 1-rotational automorphism is useful to construct an $\mathrm{OMD}_2(k \times c, v)$. Examples 9 to 12 (for 1-rotational OMDs) are used for the proof of Theorem 2.

5. GDD Construction

For combinatorial multi-arrays with fixed k and c, the GDD construction in the literature (*e.g.*, Li *et al.*, 2018; Matsubara and Kageyama, 2021) is useful to show the complete existence of multi-arrays for any v. Now, the GDD construction of an $\mathrm{OMD}_\lambda(k \times c, v)$ is presented.

Lemma 9: The existence of a (k,λ)-GDD of type $h_1^{t_1}h_2^{t_2}\cdots h_n^{t_n}$, an $\mathrm{OD}_1(k,k)$ and an $\mathrm{OMD}_\lambda(k \times c, h_i c+1)$ for each i $(1 \leq i \leq n)$ implies the existence of an $\mathrm{OMD}_\lambda(k \times c, v^*)$ with $v^* = c(h_1t_1 + \cdots + h_nt_n) + 1$.

Proof: Let G_ℓ be a group of a (k,λ)-GDD of type $h_1^{t_1}h_2^{t_2}\cdots h_n^{t_n}$ on $\mathbb{Z}_v$ with $1 \leq \ell \leq u$, $v = \sum_{i=1}^n h_i t_i$ and $u = \sum_{i=1}^n t_i$. Then, we take the direct product $\mathbb{Z}_v \times \mathbb{Z}_c$, and let $V = (\mathbb{Z}_v \times \mathbb{Z}_c) \cup \{\infty\}$ be a point set of the required $\mathrm{OMD}_\lambda(k \times c, v^*)$.

Further let the ith block of the (k,λ)-GDD of type $h_1^{t_1}h_2^{t_2}\cdots h_n^{t_n}$ be

$$\{v_{i1}, v_{i2}, \ldots, v_{ik}\},\ 1 \leq i \leq b,$$

where b is the number of blocks of the (k,λ)-GDD. Let the jth row of an $\mathrm{OD}_1(k,k)$ on $\mathbb{Z}_k$ be

$$(a_{j1}, a_{j2}, \ldots, a_{jk}),\ 1 \leq j \leq k(k-1).$$

Then replace each point $v_{ii'} \in \mathbb{Z}_v$ with a subset $B_{ii'} = \{(v_{ii'}, e) \mid e \in \mathbb{Z}_c\}$ for $1 \leq i \leq b$ and $1 \leq i' \leq k$. In this case the following row set:

$$\mathcal{R}_0 = \{\left(B_{ia_{j1}} \mid B_{ia_{j2}} \mid \ldots \mid B_{ia_{jk}}\right) \mid 1 \leq i \leq b, 1 \leq j \leq k(k-1)\}$$

is at first considered.

Furthermore, let $\mathcal{R}_\ell$ on $(G_\ell \times \mathbb{Z}_c) \cup \{\infty\}$ with $1 \leq \ell \leq u$ be the row sets obtained from the $\mathrm{OMD}_\lambda(k \times c, h_i c+1)$ with $|G_\ell| = h_i$ for some i $(1 \leq i \leq n)$. Then, any ordered pair of

two points (x, c_1) and (y, c_2) with $x, y \in G_\ell$ and $c_1, c_2 \in \mathbb{Z}_c$ for any ℓ appears in λ rows of any ordered two columns of $\mathcal{R}_\ell$, and does not appear in different entries of any row of other row sets. Moreover, any ordered pair of two points (x, c_1) and (y, c_2) with $x \in G_\ell, y \in G_{\ell'}$ and $c_1, c_2 \in \mathbb{Z}_c$ for any ℓ, ℓ' $(\ell \neq \ell')$ appears in λ rows of any ordered two columns of $\mathcal{R}_0$, while it does not appear in different entries of any row of other row sets.

Hence, the union of these row sets $\mathcal{R}_0 \cup \mathcal{R}_1 \cup \cdots \cup \mathcal{R}_u$ can yield the required $\mathrm{OMD}_\lambda(k \times c, v^*)$. □

Moreover, the following result can also be obtained.

Lemma 10: The existence of a (k, λ)-GDD of type $h_1^{t_1} h_2^{t_2} \cdots h_n^{t_n}$, an $\mathrm{OD}_1(k, k)$ and an $\mathrm{OMD}_\lambda(k \times c, h_i c)$ for each i $(1 \leq i \leq n)$ implies the existence of an $\mathrm{OMD}_\lambda(k \times c, v^*)$ with $v^* = c(h_1 t_1 + \cdots + h_n t_n)$.

Proof: Let G_ℓ $(1 \leq \ell \leq u)$ and $\mathcal{R}_0$ be the same as in the proof of Lemma 9. Moreover, let $\mathcal{R}_\ell$ on $G_\ell \times \mathbb{Z}_c$ with $1 \leq \ell \leq u$ be the row sets obtained from the $\mathrm{OMD}_\lambda(k \times c, h_i c)$ with $|G_\ell| = h_i$ for some i $(1 \leq i \leq n)$.

By discussion similar to the proof of Lemma 9, the union of these row sets $\mathcal{R}_0 \cup \mathcal{R}_1 \cup \cdots \cup \mathcal{R}_u$ can yield the required $\mathrm{OMD}_\lambda(k \times c, v^*)$. □

The following existence results on GDDs are obtained by checking that the parameters satisfy the conditions described in Lemma 4.

Lemma 11: There exist a $(3, 1)$-GDD of type $6^u 6^1$, a $(3, 1)$-GDD of type $6^u 8^1$ and a $(3, 1)$-GDD of type $6^u 10^1$ for any $u \geq 3$.

Lemma 12: There exist a $(3, 1)$-GDD of type 3^3, a $(3, 1)$-GDD of type 4^3, a $(3, 1)$-GDD of type 5^3 and a $(3, 1)$-GDD of type $3^4 5^1$.

Note that a (k, λ)-GDD with $\lambda \geq 1$ can be obtained from a $(k, 1)$-GDD by taking λ copies of each block.

6. Construction from k-SOLS(v)

Let $L = (a_{ij})$ and $L' = (a'_{ij})$ are two latin squares of order v. The latin squares L and L' are said to be *orthogonal* if all ordered pairs (a_{ij}, a'_{ij}) are distinct. A set of latin squares $L_1, \ldots, L_s$ is called *mutually orthogonal latin squares* of order v, denoted by s-MOLS(v), if they are orthogonal in each pair. A *self-orthogonal latin square* of order v is a latin square that is orthogonal to its transpose. A set $\{L_1, \ldots, L_s\}$ of self-orthogonal latin squares of order v is denoted by s-SOLS(v), if $\{L_1, L_2^T, \ldots, L_s, L_s^T\}$ is a $2s$-MOLS(v). Without loss of generality, any latin square in an s-SOLS(v) can be replaced by a latin square with $a_{ii} = i$, by renaming the symbols.

Lemma 13 (Abel and Bennet, 2012)**:** There exists a 2-SOLS(v) for any positive integer v, except for $v \in \{2, 3, 4, 5, 6\}$ and possibly for $v \in \{10, 12, 14, 18, 21, 22, 24, 30, 34\}$.

Lemma 14 (Finizio and Zhu, 2007)**:** There exists a $(2^{n-1} - 1)$-SOLS(2^n) for any $n \geq 2$.

It is well known (see Bierbrauer, 2007) that the existence of a k-MOLS(v), all of whose squares satisfy $a_{ii} = i$ with $1 \leq i \leq v$, is equivalent to the existence of an $\mathrm{OD}_1(k+2, v)$. Moreover, in Matsubara and Kageyama (2015) and Sawa *et al.* (2007), some type of combinatorial designs, called pairwise additive BIB designs, are constructed by use of a k-SOLS(v). In a manner similar to Matsubara and Kageyama (2015) and Sawa *et al.* (2007), the following construction is presented.

Lemma 15: The existence of a k-SOLS(v) implies the existence of an $\mathrm{OMD}_2((k+1)\times 2, v)$.

Proof: Let a set of $2k$-MOLS(v) derived from the k-SOLS(v) be $\{L_h, L_h^T \mid 1 \leq h \leq k\}$, where $L_h = (a_{ij}^{(2h-1)})$, $L_h^T = (a_{ij}^{(2h)}) = (a_{ji}^{(2h-1)})$ and $a_{ii}^{(2h-1)} = a_{ii}^{(2h)} = i$ $(1 \leq i \leq v)$. Further let $\mathcal{R}$ be a set of the following $v(v-1)/2$ rows:

$$(i, j \mid a_{ij}^{(1)}, a_{ij}^{(2)} \mid a_{ij}^{(3)}, a_{ij}^{(4)} \mid \cdots \mid a_{ij}^{(2k-1)}, a_{ij}^{(2k)})$$

with $1 \leq i < j \leq v$.

Then $(a_{ij}^{(2h_1-1)}, a_{ij}^{(2h_2-1)})$ and $(a_{ij}^{(2h_1)}, a_{ij}^{(2h_2)})$, for $1 \leq i < j \leq v$ and each h_1, h_2 of $1 \leq h_1 < h_2 \leq k$, yield all of pairs of distinct points in V, since L_{h_1} and L_{h_2} are orthogonal. Moreover, $(a_{ij}^{(2h_1)}, a_{ij}^{(2h_2-1)})$ and $(a_{ij}^{(2h_1-1)}, a_{ij}^{(2h_2)})$ for $1 \leq i < j \leq v$ also yield all of pairs of distinct points in V, since $L_{h_1}^T$ and L_{h_2} are orthogonal. Hence it is seen that the above-mentioned $\mathcal{R}$ yields an $\mathrm{OMD}_2((k+1)\times 2, v)$. □

Now, two families of an $\mathrm{OMD}_2(k \times 2, v)$ can be constructed by taking Lemma 15 with Lemmas 13 and 14 as the following shows.

Lemma 16: There exists an $\mathrm{OMD}_2(3 \times 2, v)$ for any $v \geq 7$ except for $v \in \{10, 12, 14, 18, 21, 22, 24, 30, 34\}$.

Lemma 17: There exists an $\mathrm{OMD}_2(2^{n-1} \times 2, 2^n)$ for any $n \geq 2$.

7. Proof of Main Results

We are now in a position to prove Theorems 1 and 2.

Proof of Theorem 1: For an odd prime power q, the existence of the required $\mathrm{OMD}_\lambda(k \times 2, q)$ with $2 \leq k \leq (q-1)/2$ is shown by taking Lemmas 2 and 3 with some copies of rows. On the other hand, the existence of the required $\mathrm{OMD}_\lambda(k \times 2, 2^n)$ with $n \geq 2$ and $2 \leq k \leq 2^{n-1}$ is shown by use of Lemmas 3 and 17 and taking copies of rows. □

Proof of Theorem 2: For the complete proof, it is enough to show the existence of the following cases:

(I) $v \equiv 1 \pmod 4$ and $v \neq 9$ when $\lambda \geq 1$,
(II) $v \equiv 0, 2, 3 \pmod 4$ when $\lambda \equiv 0 \pmod 2$,
(III) $v = 9$ and $\lambda \geq 2$.

In Cases (I) and (II), minimal $\mathrm{OMD}_\lambda(3 \times 2, v)$, *i.e.*, $\lambda = 1$ and $\lambda = 2$, respectively, are firstly constructed and then the existence for any λ is shown by taking copies of rows

of the OMD. Since the existence of a minimal $\mathrm{OMD}_\lambda(3\times 2,9)$, *i.e.*, $\lambda=1$, is unknown, the existence of $\mathrm{OMD}_\lambda(3\times 2,9)$ for any $\lambda\geq 2$ is shown in Case (III) by using examples with $\lambda=2,3$.

Case (I): Lemma 7 with Examples 5 and 6 shows the existence of an $\mathrm{OMD}_1(3\times 2,v)$ with $v=25,33$. Lemma 8 with Example 2 shows the existence of an $\mathrm{OMD}_1(3\times 2,v)$ with $v=13,17,29,37,41$. Moreover, Examples 7 and 8 show the existence of an $\mathrm{OMD}_1(3\times 2,v)$ with $v=21,45$.

On the other hand, by Lemma 11, there exist a $(3,1)$-GDD of type 6^u6^1, a $(3,1)$-GDD of type 6^u8^1 and a $(3,1)$-GDD of type 6^u10^1 for any $u\geq 3$. Now consider the $\mathrm{OD}_1(3,3)$ given in Lemma 1 and the $\mathrm{OMD}_1(3\times 2,v)$ with $v=6\cdot 2+1,8\cdot 2+1,10\cdot 2+1=13,17,21$ given above. Then Lemma 9 yields (i) an $\mathrm{OMD}_1(3\times 2,v)$ with $v\geq 49$ and $v\equiv 1 \pmod{12}$ from the $(3,1)$-GDD of type 6^u6^1, (ii) an $\mathrm{OMD}_1(3\times 2,v)$ with $v\geq 53$ and $v\equiv 5 \pmod{12}$ from the $(3,1)$-GDD of type 6^u8^1, and (iii) an $\mathrm{OMD}_1(3\times 2,v)$ with $v\geq 57$ and $v\equiv 9 \pmod{12}$ from the $(3,1)$-GDD of type 6^u10^1.

Hence, for Case (I), the required multi-arrays are constructed by taking copies of rows of the $\mathrm{OMD}_1(3\times 2,v)$.

Case (II): Lemma 16 gives an $\mathrm{OMD}_2(3\times 2,v)$ with $v\equiv 0,2,3 \pmod 4$ except for $v\in\{6,10,12,14,18,22,24,30,34\}$. Examples 1 and 9 to 12 yield an $\mathrm{OMD}_2(3\times 2,v)$ with $v\in\{6,10,12,14,22\}$.

On the other hand, by Lemma 12 with use of two copies of rows, there exist a $(3,2)$-GDD of type 3^3, a $(3,2)$-GDD of type 4^3, a $(3,2)$-GDD of type 5^3 and a $(3,2)$-GDD of type 3^45^1. Now consider the $\mathrm{OD}_1(3,3)$ and the $\mathrm{OMD}_2(3\times 2,v)$ with $v=3\cdot 2,4\cdot 2,5\cdot 2=6,8,10$ given above. Then Lemma 10 yields an $\mathrm{OMD}_2(3\times 2,v)$ with $v\in\{18,24,30,34\}$. Thus, for Case (II), the required multi-arrays are constructed by taking copies of rows of the $\mathrm{OMD}_2(3\times 2,v)$.

Case (III): Lemma 7 with Examples 3 and 4 shows the existence of an $\mathrm{OMD}_\lambda(3\times 2,9)$ with $\lambda=2,3$. Hence, for Case (III), the required multi-arrays are constructed by combining u copies and u' copies of rows of the $\mathrm{OMD}_2(3\times 2,9)$ and the $\mathrm{OMD}_3(3\times 2,9)$, respectively, with $\lambda=2u+3u'$ $(u\geq 0,u'\geq 0)$. □

8. Concluding Remark

Theorem 1 shows the existence of an $\mathrm{OMD}_\lambda(k\times 2,q)$ for any prime power q except possibly for $q\equiv 1 \pmod 4$ and $\lambda\equiv 1 \pmod 2$. Moreover, Theorem 2 shows that the necessary conditions (1) (2) and (3) are also sufficient for the existence of an $\mathrm{OMD}_\lambda(3\times 2,v)$ except possibly for an $\mathrm{OMD}_1(3\times 2,9)$. Unfortunately, the existence of the $\mathrm{OMD}_1(k\times 2,q)$ with $k\geq 4, q\equiv 1 \pmod 4$ and the $\mathrm{OMD}_1(3\times 2,9)$ cannot be proved by any method in this paper.

Lemma 7 together with the asymptotic existence results on a cyclic $\mathrm{PMA}_1(k\times 2,v)$ given in Li *et al.* (2018) and Matsubara and Kageyama (2021) can provide some asymptotic existence of a cyclic $\mathrm{OMD}_1(k\times 2,v)$ which is minimal. However, it seems difficult to show both of the exact and asymptotic existence of an $\mathrm{OMD}_\lambda(k\times c,v)$ with $N=v$, $k\geq 3$ and $c\geq 2$.

Finally, though we can find some applications of combinatorial structures (OMA, PMA, OD) related to the OMD as stated in Sections 1 and 2, any application of the OMD is not presented anywhere, including this paper. It will be discussed in a forthcoming paper.

References

Abel, R. J. R. and Bennet, F. E. (2012). Existence of 2 SOLS and 2 ISOLS. *Discrete Mathematics*, **312**, 854–867.

Bierbrauer, J. (2007). Ordered designs, perpendicular arrays, and permutation sets. In: Colbourn, C. J., Dinitz, J. H. (Eds.). *The CRC Handbook of Combinatorial Designs (2nd ed.)*. CRC Press, Boca Raton, 543–547.

Bierbrauer, J. and Edel, Y. (1994). Theory of perpendicular arrays. *Journal of Combinatorial Designs*, **2**, 375–406.

Brickell, E. F. (1984). A few results in message authentication. *Congressus Numerantium*, **43**, 141–154.

Finizio, N. J. and Zhu, L. (2007). Self-orthogonal latin squares. In: Colbourn, C. J., Dinitz, J. H. (Eds.). *The CRC Handbook of Combinatorial Designs (2nd ed.)*. CRC Press, Boca Raton, 211–219.

Ge, G. (2007). Group divisible designs. In: Colbourn, C. J., Dinitz, J. H. (Eds.). *The CRC Handbook of Combinatorial Designs (2nd ed.)*. CRC Press, Boca Raton, 255–260.

Kunert, J. and Martin, R. J. (2000). Optimality of type I orthogonal arrays for cross-over models with correlated errors. *Journal of Statistical Planning and Inference*, **87**, 119–124.

Li, M., Liang, M. and Du, B. (2015). A construction of t-fold perfect splitting authentication codes with equal deception probabilities. *Cryptography and Communications*, **7**, 207–215.

Li, M., Liang, M., Du, B. and Chen, J. (2018). A construction for optimal c-splitting authentication and secrecy codes. *Designs, Codes and Cryptography*, **86**, 1739–1755.

Majumdar, D. and Martin, R. J. (2004). Efficient designs based on orthogonal arrays of type I and type II for experiments using units ordered over time or space. *Statistical Methodology*, **1**, 19–35.

Matsubara, K. and Kageyama, S. (2015). The existence of 3 pairwise additive B$(v, 2, 1)$ for any $v \geq 6$. *Journal of Combinatorial Mathematics and Combinatorial Computing*, **95**, 27–32.

Matsubara, K. and Kageyama, S. (2021). The existence of perpendicular multi-arrays. In: Arnold, B. C., Balakrishnan, N., Coelho, C. A. (Eds.). *Contributions to Statistical Distribution Theory and Inference – Festschrift in Honor of C. R. Rao on the Occasion of His 100th Birthday*. Contributions to Statistics, Springer, to appear.

Mukerjee, R. (1998). On balanced orthogonal multi-arrays: Existence, construction and application to design of experiments. *Journal of Statistical Planning and Inference*, **73**, 149–162.

Rao, C. R. (1961). Combinatorial arrangements analogous to orthogonal arrays. *Sankhyā*, **A23**, 283–286.

Sawa, M., Matsubara, K., Matsumoto, D., Kiyama, H. and Kageyama, S. (2007). The spectrum of additive BIB designs. *Journal of Combinatorial Designs*, **15**, 235–254.

Sitter, R. R. (1993). Balanced repeated replications based on orthogonal multi-arrays. *Biometrika*, **80**, 211–221.

Appendix

Some individual examples which can be found by use of a computer are presented. Note that each of such examples cannot be presented by use of the construction methods provided in this paper.

Example 3: A cyclic $\mathrm{PMA}_2(3 \times 2, 9)$ on $\mathbb{Z}_9$ is given by

$$(0, 1 \mid 2, 4 \mid 3, 6), (0, 7 \mid 1, 2 \mid 5, 8) \mod 9.$$

Example 4: A cyclic $\mathrm{PMA}_3(3 \times 2, 9)$ on $\mathbb{Z}_9$ is given by

$$(0, 8 \mid 2, 3 \mid 1, 5), (0, 7 \mid 1, 2 \mid 3, 6), (0, 6 \mid 5, 7 \mid 2, 8) \mod 9.$$

Example 5: A cyclic $\mathrm{PMA}_1(3 \times 2, 25)$ on $\mathbb{Z}_{25}$ is given by

$$(0, 12 \mid 3, 23 \mid 17, 18), (0, 22 \mid 12, 21 \mid 13, 24), (0, 24 \mid 5, 7 \mid 3, 14) \mod 25.$$

Example 6: A cyclic $\mathrm{PMA}_1(3 \times 2, 33)$ on $\mathbb{Z}_{33}$ is given by

$$(0, 16 \mid 17, 27 \mid 12, 13), (0, 1 \mid 14, 24 \mid 2, 16), (0, 1 \mid 8, 30 \mid 7, 24),$$
$$(0, 3 \mid 15, 31 \mid 22, 28) \mod 33.$$

Example 7: A cyclic $\mathrm{OMD}_1(3 \times 2, 21)$ on $\mathbb{Z}_{21}$ is given by

$$(10, 11 \mid 5, 16 \mid 9, 12), (9, 12 \mid 8, 13 \mid 2, 19), (10, 11 \mid 8, 13 \mid 5, 16),$$
$$(2, 19 \mid 9, 12 \mid 10, 11), (10, 11 \mid 2, 19 \mid 7, 14) \mod 21.$$

Example 8: A cyclic $\mathrm{OMD}_1(3 \times 2, 45)$ on $\mathbb{Z}_{45}$ is given by

$$(19, 26 \mid 17, 28 \mid 3, 42), (20, 25 \mid 21, 24 \mid 6, 39), (17, 28 \mid 6, 39 \mid 11, 34),$$
$$(11, 34 \mid 6, 39 \mid 10, 35), (9, 36 \mid 21, 24 \mid 11, 34), (21, 24 \mid 11, 34 \mid 13, 32),$$
$$(17, 28 \mid 9, 36 \mid 10, 35), (4, 41 \mid 10, 35 \mid 1, 44), (22, 23 \mid 2, 43 \mid 10, 35),$$
$$(17, 28 \mid 10, 35 \mid 13, 32), (13, 32 \mid 16, 29 \mid 22, 23) \mod 45.$$

Example 9: A 1-rotational $\mathrm{OMD}_2(3 \times 2, 10)$ on $\mathbb{Z}_9$ is given by

$$(0, \infty \mid 1, 5 \mid 6, 8), (2, 7 \mid 0, \infty \mid 1, 4), (2, 4 \mid 5, 7 \mid 0, \infty), (0, 6 \mid 4, 8 \mid 1, 7),$$
$$(0, 7 \mid 4, 6 \mid 2, 3) \mod 9.$$

Example 10: A 1-rotational $\mathrm{OMD}_2(3 \times 2, 12)$ on $\mathbb{Z}_{11}$ is given by

$$(0, \infty \mid 4, 7 \mid 1, 10), (2, 9 \mid 0, \infty \mid 5, 6), (4, 7 \mid 1, 10 \mid 0, \infty), (2, 9 \mid 3, 8 \mid 1, 10),$$
$$(4, 7 \mid 5, 6 \mid 2, 9), (2, 9 \mid 5, 6 \mid 4, 7) \mod 11.$$

Example 11: A 1-rotational $\mathrm{OMD}_2(3 \times 2, 14)$ on $\mathbb{Z}_{13}$ is given by

$$(0, \infty \mid 1, 12 \mid 6, 7), (5, 8 \mid 0, \infty \mid 3, 10), (4, 9 \mid 3, 10 \mid 0, \infty), (4, 9 \mid 6, 7 \mid 5, 8),$$
$$(6, 7 \mid 2, 11 \mid 4, 9), (2, 11 \mid 4, 9 \mid 5, 8), (3, 10 \mid 6, 7 \mid 2, 11) \mod 13.$$

Example 12: A 1-rotational $\mathrm{OMD}_2(3 \times 2, 22)$ on $\mathbb{Z}_{21}$ is given by

$$(0, \infty \mid 7, 14 \mid 5, 16), (7, 14 \mid 0, \infty \mid 10, 11), (5, 16 \mid 8, 13 \mid 0, \infty), (9, 12 \mid 8, 13 \mid 1, 20),$$
$$(7, 14 \mid 9, 12 \mid 5, 16), (5, 16 \mid 1, 20 \mid 7, 14), (7, 14 \mid 4, 17 \mid 1, 20), (5, 16 \mid 4, 17 \mid 2, 19),$$
$$(9, 12 \mid 3, 18 \mid 8, 13), (7, 14 \mid 9, 12 \mid 8, 13), (9, 12 \mid 1, 20 \mid 2, 19) \mod 21.$$

Finally, a table of the existence of a minimal $\mathrm{OMD}_\lambda(k \times 2, v)$ shown by our methods is presented for $4 \leq v \leq 20$. When $c = 2$ is fixed, N and λ are uniquely determined by v. For v, N and λ, values of k are indicated about known or unknown existence of the OMD. Note that values of bold k represent the upper bound of k obtained from (1) and "–" in the column of unknown implies that the complete existence of an $\mathrm{OMD}_\lambda(k \times c, v)$ is shown. Moreover, for two minimal OMDs of Nos. 2 and 6 which cannot be obtained by Theorems 1 and 2, base rows are newly given.

Table 1: Minimal $\mathrm{OMD}_\lambda(k \times c, v)$ with $4 \leq v \leq 20, c = 2$

No	v	N	λ	known	unknown	Source
1	4	6	2	$k = \mathbf{2}$	–	Theorem 1
2	5	5	1	$k = \mathbf{2}$	–	$(1, 4 \mid 2, 3) \bmod 5$
3	6	15	2	$2 \leq k \leq \mathbf{3}$	–	Theorem 2
4	7	21	2	$2 \leq k \leq \mathbf{3}$	–	Theorem 1
5	8	28	2	$2 \leq k \leq \mathbf{4}$	–	Theorem 1
6	9	18	1	$k = 2$	$3 \leq k \leq \mathbf{4}$	$(0, 1 \mid 2, 4), (2, 4 \mid 0, 1) \bmod 9$
7	10	45	2	$2 \leq k \leq 3$	$4 \leq k \leq \mathbf{5}$	Theorem 2
8	11	55	2	$2 \leq k \leq \mathbf{5}$	–	Theorem 1
9	12	66	2	$2 \leq k \leq 3$	$4 \leq k \leq \mathbf{6}$	Theorem 2
10	13	39	1	$2 \leq k \leq 3$	$4 \leq k \leq \mathbf{6}$	Theorem 2
11	14	91	2	$2 \leq k \leq 3$	$4 \leq k \leq \mathbf{7}$	Theorem 2
12	15	105	2	$2 \leq k \leq 3$	$4 \leq k \leq \mathbf{7}$	Theorem 2
13	16	120	2	$2 \leq k \leq \mathbf{8}$	–	Theorem 1
14	17	68	1	$2 \leq k \leq 3$	$4 \leq k \leq \mathbf{8}$	Theorem 2
15	18	153	2	$2 \leq k \leq 3$	$4 \leq k \leq \mathbf{9}$	Theorem 2
16	19	171	2	$2 \leq k \leq \mathbf{9}$	–	Theorem 1
17	20	190	2	$2 \leq k \leq 3$	$4 \leq k \leq \mathbf{10}$	Theorem 2

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 187–198

A Note on the Folklore of Free Independence

Arijit Chakrabarty[1], **Sukrit Chakraborty**[1] **and Rajat Subhra Hazra**[1]
[1] *Theoretical Statistics and Mathematics Unit*
Indian Statistical Institute, 203 B.T. Road, Kolkata 700108

Dedicated to the memory of Professor Aloke Dey

Received: 12 November 2020; Revised: 22 January 2021; Accepted: 25 January 2021

Abstract

It is shown that a Wishart matrix of standard complex normal random variables is asymptotically freely independent of an independent random matrix, under minimal conditions, in two different sense of asymptotic free independence.

Key words: Voiculescu's theorem; Random matrix theory; Asymptotic free independence; Wishart matrix.

AMS Subject Classifications: Primary 60B20; Secondary 46L54.

1. Introduction

Since the seminal discovery of [10], there have been several folklores regarding free independence. For example, one such folklore is that any two independent Wigner matrices are asymptotically freely independent, and another is that any Wishart matrix is asymptotically freely independent of a deterministic matrix. While such folklores are true, more often than not, there are a few problems. The first and foremost problem is that the meaning of the phrase "asymptotically freely independent" varies with context. A widely used definition is in terms of the normalized expected trace (or without the expectation). Unfortunately, with this definition, the claim of asymptotic free independence can easily fail, in the absence of any other assumption. The counter example in [7] is noteworthy. This articulates the second problem with the folklore, which is that the required assumptions are usually missing. Nevertheless, in the literature, there are several rigorous proofs of various versions of Voiculescu's theorem; see, for example, the monographs [9], [2] and [8]. The reader will notice that the versions in the above references are not monotonic in strength, that is, one version does not necessarily imply another. In other words, there is no general theorem regarding asymptotic free independence from which most results of interest follow.

This note is a modest attempt at settling some of the issues mentioned above in a specific example. Theorems 1 and 2 claim asymptotic free independence of a Wishart matrix W_N of standard complex normal random variables and an independent matrix Y_N, under two different definitions of asymptotic free independence. The former is the usual definition, in terms of normalized expected trace, while the latter is in terms of the limiting spectral distribution of random matrices, which is weaker than the former. In both the above mentioned

Corresponding Author: Arijit Chakrabarty
Email: arijit.isi@gmail.com

theorems, the limiting spectral distribution of Y_N is assumed to be compactly supported, at the least. This assumption is relaxed in Theorem 3, a consequence of which is that the claim is also significantly weakened. The proofs of Theorems 2 and 3 are based on truncation arguments.

We choose to work with the complex normal distribution because they yield the strongest results in that the assumptions on Y_N become minimal. This is why, for example, Theorem 22.35 of [9] assumes the distribution to be complex normal. It is worth noting that Theorem 2 of [1] and the results in [6] are similar in spirit. Although the results are stated for a Wishart matrix, they hold for a Wigner matrix as well.

2. The Results

Let $(Z_{i,j} : i, j \in \mathbb{N})$ be a family of i.i.d. standard complex Normal random variables. That is, $(\Re(Z_{i,j}) : i, j \geq 1)$ and $(\Im(Z_{i,j}) : i, j \geq 1)$ are independent families of i.i.d. real $N(0, 1/2)$ random variables. Suppose that $(M_N : N \geq 1)$ is a sequence of positive integers such that

$$\lim_{N\to\infty} \frac{N}{M_N} = \lambda \in (0, \infty)\,. \tag{1}$$

For each $N \geq 1$, let X_N be the $M_N \times N$ random matrix defined by

$$X_N(i, j) := Z_{i,j},\ 1 \leq i \leq M_N,\ 1 \leq j \leq N\,.$$

For $N \geq 1$, define an $N \times N$ random Hermitian matrix by

$$W_N := \frac{1}{M_N} X_N^* X_N\,.$$

Notice that for $1 \leq i, j \leq N$,

$$W_N(i, j) = \frac{1}{M_N} \sum_{k=1}^{M_N} \overline{Z_{k,i}} Z_{k,j}\,.$$

Hence W_N is a Wishart matrix.

For a random Hermitian $N \times N$ matrix Z, its "empirical spectral distribution" and "expected empirical spectral distribution", denoted by $\mathrm{ESD}(Z)$ and $\mathrm{EESD}(Z)$, respectively, are probability measures on $\mathbb{R}$, defined as

$$\begin{aligned} \mathrm{ESD}(Z) &= \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}(\lambda_i \in \cdot)\,, \\ \mathrm{EESD}(Z) &= \frac{1}{N} \sum_{i=1}^{N} P(\lambda_i \in \cdot)\,, \end{aligned}$$

where $\lambda_1, \ldots, \lambda_N$ are the eigenvalues of Z, counted with multiplicity.

It is well known that as $N \to \infty$,

$$\mathrm{ESD}(W_N) \to \nu_\lambda\,,$$

weakly in probability, where ν_λ, with λ as in (1), is the Marčenko-Pastur distribution, defined by

$$\nu_\lambda(dx) = \begin{cases} \left(1-\frac{1}{\lambda}\right)\mathbf{1}(0\in dx) + \frac{1}{2\pi}\frac{\sqrt{(\lambda_+-x)(x-\lambda_-)}}{\lambda x}\mathbf{1}_{[\lambda_-,\lambda_+]}(x)\,dx, & \lambda > 1\,, \\ \frac{1}{2\pi}\frac{\sqrt{(\lambda_+-x)(x-\lambda_-)}}{\lambda x}\mathbf{1}_{[\lambda_-,\lambda_+]}(x)\,dx, & \lambda \le 1\,, \end{cases}$$

with $\lambda_\pm = (1\pm\sqrt{\lambda})^2$.

For each $N \ge 1$, Y_N is an $N\times N$ random complex Hermitian matrix, **independent** of $(Z_{i,j} : i,j \in \mathbb{N})$. The exact assumption on the spectrum of Y_N will vary from result to result, and hence will be mentioned in the statements of the respective results. However, at the very least, there exists a (non-random) probability measure μ on $\mathbb{R}$ such that

$$\mathrm{ESD}(Y_N) \to \mu\,, \tag{2}$$

weakly in probability, as $N\to\infty$.

The statements of the following results are based on the theory of C^*-probability spaces. A reader unacquainted with this may look at [9]. It is known that given probability measures μ_1 and μ_2 which are supported on a compact subset of $\mathbb{R}$, there exist a C^*-probability space $(\mathcal{A},\varphi)$, and two freely independent self-adjoint elements $a_1, a_2 \in \mathcal{A}$ such that

$$\varphi\left(a_i^n\right) = \int_{-\infty}^{\infty} x^n \mu_i(dx),\ n\in\mathbb{N},\ i = 1,2\,.$$

The probability measures μ_1 and μ_2 are called the distributions of a_1 and a_2, and denoted by $\mathcal{L}(a_1)$ and $\mathcal{L}(a_2)$, respectively.

The first result shows asymptotic free independence between W_N and Y_N in the sense of normalized expected trace.

Theorem 1: Assume that μ is compactly supported, and that for each $n\in\mathbb{N}$,

$$\lim_{N\to\infty} \mathrm{E}\left[\frac{1}{N}\mathrm{Tr}(Y_N^n)\right] = \int_{-\infty}^{\infty} x^n\mu(dx)\,,\ \text{and} \tag{3}$$

$$\lim_{N\to\infty} \mathrm{Var}\left[\frac{1}{N}\mathrm{Tr}(Y_N^n)\right] = 0\,. \tag{4}$$

Then, there exists a C^*-probability space $(\mathcal{A},\varphi)$, in which there are two freely independent self-adjoint elements w and y, having distributions ν_λ and μ, respectively, and satisfying the following: For every polynomial p in two variables having complex coefficients,

$$\lim_{N\to\infty}\frac{1}{N}\mathrm{ETr}\left[p\left(W_N,Y_N\right)\right] = \varphi\Big(p(w,y)\Big)\,. \tag{5}$$

Consequently, if $p\left(W_N,Y_N\right)$ has real eigenvalues, a.s., for all N, then as $N\to\infty$,

$$\mathrm{EESD}\left(p\left(W_N,Y_N\right)\right) \xrightarrow{w} \mathcal{L}\left(p\left(w,y\right)\right)\,. \tag{6}$$

Remark 1: When Y_N is deterministic, the assumptions of Theorem 1 just mean that

$$\lim_{n\to\infty}\frac{1}{N}\mathrm{Tr}(Y_N^n) = \int_{-\infty}^{\infty} x^n\mu(dx)\,,$$

which is stronger than (2). In general, (3) and (4) together imply (2) whenever μ is determined by its moments, which is necessarily the case if μ is compactly supported.

Remark 2: The claim (6) is an immediate consequence of (5), whenever p is such that the eigenvalues of $p(W_N, Y_N)$ are a.s. real. For example, if W_N is non-negative definite, then the above holds for

$$p(x, y) = xy\,.$$

In the next result, both the hypotheses and the claim are weakened to (2) and (6), respectively. In other words, this results proves asymptotic free independence in the sense of (6) as opposed to (5).

Theorem 2: If μ, as in (2), is compactly supported, then for every polynomial p in two variables having complex coefficients such that $p\,(W_N, Y_N)$ has real eigenvalues, a.s., for all N, (6) holds.

The last result deals with the case when the support of μ is possibly unbounded. For measures with possibly unbounded support, '$\boxplus$' and '$\boxtimes$' denote their free additive and multiplicative convolutions, respectively. For the latter, at least of one of the two measures has to be supported on the non-negative half line. See [5] for the details.

Theorem 3: If (2) holds for a probability measure μ which is not necessarily compactly supported, then

$$\begin{aligned} \mathrm{EESD}(Y_N + W_N) &\xrightarrow{w} \mu \boxplus \nu_\lambda\,, \text{and} \\ \mathrm{EESD}(Y_N W_N) &\xrightarrow{w} \mu \boxtimes \nu_\lambda\,, \end{aligned}$$

as $N \to \infty$.

Remark 3: Theorems 1 - 3 hold true, if the Wishart matrix is replaced by a Wigner matrix with standard complex normal entries, and the Marčenko-Pastur distribution is replaced by the semicircle law.

3. Some Facts

For the proofs of the results mentioned in Section 2, a few facts will be needed, which are stated here. The proofs are omitted because the results are either elementary or can be found in a cited reference.

The first one is a comparison between ranks of deterministic matrices.

Fact 3.1: Let p be a polynomial in two variables, with complex coefficients. Then, there exists a finite constant C, depending only on the polynomial p, such that

$$\mathrm{Rank}\,(p(A, B) - p(A', B)) \leq C\mathrm{Rank}(A - A')\,,$$

for square matrices A, A', B of the same order.

The next result, which is also based on rank, follows from Theorem A.43, page 503, of [4].

Fact 3.2: For probability measures μ_1 and μ_2 on $\mathbb{R}$, let $d(\mu_1, \mu_2)$ denote their Lévy distance, defined by

$$d(\mu_1, \mu_2) := \inf \{\varepsilon > 0 : \mu_1((-\infty, x-\varepsilon]) \leq \mu_2((-\infty, x]) \leq \mu_1((-\infty, x+\varepsilon])\} .$$

For $N \times N$ random Hermitian matrices A and B, it holds that

$$d(\mathrm{EESD}(A), \mathrm{EESD}(B)) \leq \frac{1}{N}\mathrm{E}[\mathrm{Rank}(A-B)] .$$

The following fact essentially follows from uniform integrability. Nonetheless, a proof is given.

Fact 3.3: For each $N \geq 1$, suppose that Y_N is an $N \times N$ random Hermitian matrix satisfying (3) and (4). Then it holds that for any $n \geq 1$ and $k_1, \ldots, k_n \geq 0$,

$$\lim_{N\to\infty} N^{-n}\mathrm{E}\left(\prod_{i=1}^{n}\mathrm{Tr}\left(Y_N^{k_i}\right)\right) = \prod_{i=1}^{n}\alpha_{k_i} ,$$

where α_n denotes the right hand side of (3).

Proof: Fix $n \geq 1$ and $k_1, \ldots, k_n \geq 0$. A consequence of (3) and (4) is that for all fixed $k \geq 1$,

$$\frac{1}{N}\mathrm{Tr}\left(Y_N^k\right) \xrightarrow{P} \alpha_k , N \to \infty .$$

Therefore,

$$N^{-n}\prod_{i=1}^{n}\mathrm{Tr}\left(Y_N^{k_i}\right) \xrightarrow{P} \prod_{i=1}^{n}\alpha_{k_i} , N \to \infty . \tag{7}$$

Let

$$k = \sum_{i=1}^{n} k_i ,$$

which we assume without loss of generality to be at least 1, and observe that

$$\begin{aligned} N^{-n}\left|\prod_{i=1}^{n}\mathrm{Tr}\left(Y_N^{k_i}\right)\right| &= \prod_{i=1}^{n}\left|\int_{-\infty}^{\infty} x^{k_i}\left(\mathrm{ESD}(Y_N)\right)(dx)\right| \\ &\leq \prod_{i=1}^{n}\int_{-\infty}^{\infty}|x|^{k_i}\left(\mathrm{ESD}(Y_N)\right)(dx) \\ &\leq \prod_{i=1}^{n}\left(\int_{-\infty}^{\infty} x^{2k}\left(\mathrm{ESD}(Y_N)\right)(dx)\right)^{k_i/2k} \\ &= \left(\frac{1}{N}\mathrm{Tr}\left(Y_N^{2k}\right)\right)^{1/2} , \end{aligned}$$

the penultimate line following from the Lyapunov inequality. Thus,

$$\limsup_{N\to\infty}\mathrm{E}\left[\left(N^{-n}\prod_{i=1}^{n}\mathrm{Tr}\left(Y_N^{k_i}\right)\right)^2\right] \leq \lim_{N\to\infty}\mathrm{E}\left(\frac{1}{N}\mathrm{Tr}\left(Y_N^{2k}\right)\right) = \alpha_{2k} < \infty ,$$

the equality being implied by (3). Hence,

$$\left(N^{-n} \prod_{i=1}^{n} \mathrm{Tr}\left(Y_N^{k_i}\right) : N \geq 1 \right)$$

is an uniformly integrable family, which in conjunction with (7) completes the proof. □

The next fact is elementary.

Fact 3.4: Let $Z_1, \ldots, Z_N$ be i.i.d. standard complex normal, that is for each $i = 1, \ldots, N$, the real and imaginary parts of Z_i are independent $N(0, 1/\sqrt{2})$. If Z denotes the column vector whose i-th component is Z_i, and U is an $N \times N$ deterministic unitary matrix, then the components of UZ are also i.i.d. standard complex normal.

The next fact has essentially been proved in page 386 of [9]. As mentioned therein, an $N \times N$ Haar unitary matrix is a random matrix distributed according to the Haar measure on the group of $N \times N$ unitary matrices. Before stating the fact, we need to introduce a few notations. Let S_n denote the group of permutations on $\{1, \ldots, n\}$ for $n \geq 1$. A permutation is identified with the partition of $\{1, \ldots, n\}$, induced by the cyclic decomposition. For $\alpha \in S_n$, $\#\alpha$ denotes the number of blocks in α, that is the number of cycles. For any block $\theta \in \alpha$, $\#\theta$ denotes the length of the cycle θ. For example, for

$$\alpha \in S_4,$$

defined by

$$\alpha(1) = 2,\ \alpha(2) = 4,\ \alpha(3) = 3,\ \alpha(4) = 1\,,$$

we write

$$\alpha = \{(1, 2, 4), (3)\}\,,$$

and hence $\#\alpha = 2$. If the elements of α, as listed above, are labelled as θ_1 and θ_2, respectively, then

$$\#\theta_1 = 3,\ \#\theta_2 = 1\,.$$

Fact 3.5: For a fixed N, let A and B be deterministic $N \times N$ Hermitian matrices. If U is an $N \times N$ Haar unitary matrix, then for any $1 \leq n \leq N$ and $k_1, \ldots, k_n \geq 0$,

$$\begin{aligned} &\mathrm{ETr}\left[\prod_{i=1}^{n}\left(UA^{k_i}U^*B\right)\right] \\ &= \sum_{\alpha,\beta\in S_n} \mathrm{Wg}(N, \alpha^{-1}\beta) \left(\prod_{\theta\in\alpha} \mathrm{Tr}\left(A^{\sum_{i\in\theta} k_i}\right)\right)\left(\prod_{\theta\in\beta^{-1}\gamma} \mathrm{Tr}\left(B^{\#\theta}\right)\right)\,, \end{aligned}$$

where Wg is the Weingarten function defined by

$$\mathrm{Wg}(N, \alpha) = \mathrm{E}\left[U(1,1)\ldots U(n,n)\overline{U(1,\alpha(1))}\ldots\overline{U(n,\alpha(n))}\right]\,,$$

for $\alpha \in S_n$, $N \geq n$ and

$$\gamma = \{(1, \ldots, n)\} \in S_n\,.$$

The following has essentially been proved in the course of the proof of Theorem 23.14 of [9].

Fact 3.6: For a fixed $n \geq 1$ and $\alpha \in S_n$,

$$\phi(\alpha) := \lim_{N\to\infty} N^{2n-\#\alpha}\mathrm{Wg}(N,\alpha) \text{ exists and is real.}$$

Furthermore, if $(\mathcal{A},\varphi)$, w and y are as in the statement of Theorem 1, then for $n \geq 1$ and $k_1,\ldots,k_n \geq 0$,

$$\begin{aligned}&\varphi\left(w^{k_1}y\ldots w^{k_n}y\right)\\ &= \sum_{\substack{\alpha,\beta\in S_n:\\ \#(\alpha^{-1}\beta)+\#\alpha+\#(\beta^{-1}\gamma)=2n+1}} \left[\phi(\alpha^{-1}\beta)\left(\prod_{\theta\in\alpha}\varphi\left(w^{\sum_{i\in\theta}k_i}\right)\right)\left(\prod_{\theta\in\beta^{-1}\gamma}\varphi\left(y^{\#\theta}\right)\right)\right].\end{aligned}$$

The following result is Corollary 2 of [3].

Fact 3.7: For a fixed $N \in \mathbb{N}$, there exists a measurable map

$$\psi : \mathbb{C}^{N\times N} \to \mathbb{C}^{N\times N},$$

where $\mathbb{C}^{N\times N}$ is the space of all $N \times N$ matrices with complex entries, such that $\psi(M)$ is an unitary matrix for every $M \in \mathbb{C}^{N\times N}$, and

$$\psi(M)^* M \psi(M)$$

is upper triangular for every M.

4. Proofs

Proof of Theorem 1: Let $(\mathcal{A},\varphi)$, w and y be as in the statement. In order to prove the claim, all that needs to be shown is that

$$\lim_{N\to\infty} N^{-1}\mathrm{E}\left[\mathrm{Tr}\left(W_N^{k_1}Y_N\ldots W_N^{k_n}Y_N\right)\right] = \varphi\left(w^{k_1}y\ldots w^{k_n}y\right), \tag{8}$$

for fixed $n \geq 1$ and $k_1,\ldots,k_n \geq 0$.

The foremost task is to show that the expectation on the left hand side of (8) exists. To that end, it suffices to show that there exists N_0 such that

$$\mathrm{E}\left[|Y_N(i,j)|^n\right] < \infty \text{ for all } N \geq N_0,\ 1 \leq i,j \leq N. \tag{9}$$

Fix $N \geq 1$ and enumerate the eigenvalues of Y_N in ascending order by $\lambda_1,\ldots,\lambda_N$. Notice that

$$\begin{aligned}\sum_{i,j=1}^N |Y_N(i,j)|^{2n} &\leq \left[\sum_{i,j=1}^N |Y_N(i,j)|^2\right]^n\\ &= \left(\sum_{i=1}^N \lambda_i^2\right)^n\\ &\leq N^{n-1}\sum_{i=1}^N \lambda_i^{2n}\\ &= N^{n-1}\mathrm{Tr}(Y_N^{2n}).\end{aligned}$$

Since (3) implies that the expectation of the right hand side is finite for N large, an N_0 satisfying (9) exists.

Proceeding towards (8), fix $N \geq N_0$, and let

$$\mathcal{F} := \sigma\left(X_N, Y_N\right) ,$$

that is $\mathcal{F}$ is the smallest σ-field with respect to which the entries of X_N and Y_N are measurable. Let U_N be a Haar unitary matrix independent of $\mathcal{F}$. Fact 3.4 implies that conditioned on U_N, the entries of $U_N X_N$ are i.i.d. standard complex Normal. That is, the conditional joint distribution of the entries of $U_N X_N$, given U_N, is the same as that of X_N. Therefore

$$\left(U_N W_N U_N^*, Y_N\right) \stackrel{d}{=} \left(W_N, Y_N\right) .$$

As a result,

$$\begin{aligned} C_N &:= \mathrm{E}\left[\mathrm{Tr}\left(W_N^{k_1} Y_N \ldots W_N^{k_n} Y_N\right)\right] \\ &= \mathrm{E}\left[\mathrm{Tr}\left(\left(U_N W_N U_N^*\right)^{k_1} Y_N \ldots \left(U_N W_N U_N^*\right)^{k_n} Y_N\right)\right] \\ &= \mathrm{E}\left[\mathrm{Tr}\left(U_N W_N^{k_1} U_N^* Y_N \ldots U_N W_N^{k_n} U_N^* Y_N\right)\right] \\ &= \mathrm{E}\,\mathrm{E}_{\mathcal{F}}\left[\mathrm{Tr}\left(U_N W_N^{k_1} U_N^* Y_N \ldots U_N W_N^{k_n} U_N^* Y_N\right)\right] , \end{aligned}$$

where $\mathrm{E}_{\mathcal{F}}$ is the conditional expectation given $\mathcal{F}$. By an appeal to Fact 3.5,

$$\begin{aligned} &\mathrm{E}_{\mathcal{F}}\left[\mathrm{Tr}\left(U_N W_N^{k_1} U_N^* Y_N \ldots U_N W_N^{k_n} U_N^* Y_N\right)\right] \\ = &\sum_{\alpha,\beta \in S_n} \mathrm{Wg}(N, \alpha^{-1}\beta) \left(\prod_{\theta \in \alpha} \mathrm{Tr}\left(W_N^{\sum_{i \in \theta} k_i}\right)\right) \left(\prod_{\theta \in \beta^{-1}\gamma} \mathrm{Tr}\left(Y_N^{\#\theta}\right)\right) . \end{aligned}$$

Taking the unconditional expectation of both sides, and using the independence of W_N and Y_N, we get that

$$C_N = \sum_{\alpha,\beta \in S_n} \mathrm{Wg}(N, \alpha^{-1}\beta) \mathrm{E}\left(\prod_{\theta \in \alpha} \mathrm{Tr}\left(W_N^{\sum_{i \in \theta} k_i}\right)\right) \mathrm{E}\left(\prod_{\theta \in \beta^{-1}\gamma} \mathrm{Tr}\left(Y_N^{\#\theta}\right)\right) . \tag{10}$$

It is well known that for all $k \in \mathbb{N}$,

$$\begin{aligned} \lim_{N\to\infty} \mathrm{E}\left(N^{-1}\mathrm{Tr}(W_N^k)\right) &= \varphi(w^k) , \\ \lim_{N\to\infty} \mathrm{Var}\left(N^{-1}\mathrm{Tr}(W_N^k)\right) &= 0 . \end{aligned}$$

Combining the above with Fact 3.3 yields that

$$\lim_{N\to\infty} \mathrm{E}\left(\prod_{\theta \in \alpha} N^{-1}\mathrm{Tr}\left(W_N^{\sum_{i \in \theta} k_i}\right)\right) = \prod_{\theta \in \alpha} \varphi\left(w^{\sum_{i \in \theta} k_i}\right) . \tag{11}$$

Similarly, (3), (4) and Fact 3.3 together imply that

$$\lim_{N\to\infty} \mathrm{E}\left(\prod_{\theta\in\beta^{-1}\gamma} N^{-1}\mathrm{Tr}\left(Y_N^{\#\theta}\right)\right) = \prod_{\theta\in\beta^{-1}\gamma} \varphi\left(y^{\#\theta}\right) . \tag{12}$$

Rewrite (10) as

$$\sum_{\alpha,\beta\in S_n}^{N^{-1}C_N} N^{\#\alpha+\#(\beta^{-1}\gamma)-1}\mathrm{Wg}(N,\alpha^{-1}\beta) \\ \mathrm{E}\left(\prod_{\theta\in\alpha} N^{-1}\mathrm{Tr}\left(W_N^{\sum_{i\in\theta}k_i}\right)\right)\mathrm{E}\left(\prod_{\theta\in\beta^{-1}\gamma} N^{-1}\mathrm{Tr}\left(Y_N^{\#\theta}\right)\right) .$$

The first claim of Fact 3.6 implies that for fixed $\alpha,\beta\in S_n$,

$$\begin{aligned} N^{\#\alpha+\#(\beta^{-1}\gamma)-1}\mathrm{Wg}(N,\alpha^{-1}\beta) &= O\left(N^{\#(\alpha^{-1}\beta)+\#\alpha+\#(\beta^{-1}\gamma)-2n-1}\right) \\ &= O(1) , \end{aligned}$$

because

$$\#\alpha+\#(\alpha^{-1}\beta)+\#(\beta^{-1}\gamma)\le 2n+1 ,$$

as shown in (23.4) and the following display on page 387 in [9]. Therefore, letting $N\to\infty$ in (13) and using the first claim of Fact 3.6 along with (11) and (12), we get that

$$\begin{aligned} &\lim_{N\to\infty} N^{-1}C_N \\ &= \sum_{\substack{\alpha,\beta\in S_n:\\ \#(\alpha^{-1}\beta)+\#\alpha+\#(\beta^{-1}\gamma)=2n+1}} \left[\phi(\alpha^{-1}\beta)\left(\prod_{\theta\in\alpha}\varphi\left(w^{\sum_{i\in\theta}k_i}\right)\right)\left(\prod_{\theta\in\beta^{-1}\gamma}\varphi\left(y^{\#\theta}\right)\right)\right] . \end{aligned}$$

The second claim of Fact 3.6 shows that the right hand side of the above equation is the same as that of (8). Thus the latter follows, which completes the proof. □

Proof of Theorem 2: Since μ is compactly supported, let $M>1$ be such that

$$\mu\left([-(M-1),M-1]\right)=1 .$$

Letting ψ be as in Fact 3.7, define

$$P_N=\psi(Y_N) ,$$

and

$$T_N := P_N^* Y_N P_N ,$$

which is an upper triangular matrix. Define an $N\times N$ matrix T_N' by

$$T_N'(i,j) := \begin{cases} T_N(i,j), & i\ne j , \\ T_N(i,i)\mathbf{1}(|T_N(i,i)|\le M), & i=j , \end{cases}$$

and let

$$Y'_N := P_N T'_N P_N^* \,. \tag{13}$$

In order to complete the proof, it suffices to show that for a fixed polynomial p satisfying the hypothesis,

$$\mathrm{EESD}\left(p(W_N, Y'_N)\right) \xrightarrow{w} \mathcal{L}\left(p(w,y)\right) \,, \tag{14}$$

and

$$\lim_{N\to\infty} d\left(\mathrm{EESD}\left(p(W_N, Y_N)\right), \mathrm{EESD}\left(p(W_N, Y'_N)\right)\right) = 0 \,, \tag{15}$$

where d is the Lévy metric, convergence in which is equivalent to weak convergence.

We start by showing (15). To that end, note that

$$\begin{aligned} N^{-1}\mathrm{Rank}(Y_N - Y'_N) &= N^{-1}\mathrm{Rank}(T_N - T'_N) \\ &= N^{-1}\#\{1 \le i \le N : |T_N(i,i)| > M\} \\ &= \left(\mathrm{ESD}(Y_N)\right)\left([-M, M]^c\right) \,, \end{aligned}$$

the inequality in the second line being based on the fact that $T_N - T'_N$ is a diagonal matrix, and hence

$$N^{-1}\mathrm{Rank}(Y_N - Y'_N) \xrightarrow{P} 0 \tag{16}$$

as $N \to \infty$. Fact 3.1 and the bounded convergence theorem show that

$$\lim_{N\to\infty} \mathrm{E}\left[\frac{1}{N}\mathrm{Rank}(p(W_N, Y_N) - p(W_N, Y'_N))\right] = 0 \,.$$

An appeal to Fact 3.2 establishes (15).

Proceeding towards (14), in view of Theorem 1 and Remark 2, it suffices to show that (3) and (4) hold with Y_N replaced by Y'_N. Equation (16) and the hypotheses imply that

$$\mathrm{ESD}(Y'_N) \to \mu \,,$$

weakly in probability, as $N \to \infty$. Since

$$\left(\mathrm{ESD}(Y'_N)\right)\left([-M, M]^c\right) = \left(\mathrm{ESD}(T'_N)\right)\left([-M, M]^c\right) = 0, \; N \ge 1 \,,$$

and

$$\mu\left([-M+1, M-1]^c\right) = 0 \,,$$

it follows that for a fixed $n \ge 1$, as $N \to \infty$,

$$\int_{-\infty}^{\infty} x^n \left(\mathrm{ESD}(Y'_N)\right)(dx) \xrightarrow{P} \int_{-\infty}^{\infty} x^n \mu(dx) \,.$$

The observations that

$$\frac{1}{N}\mathrm{Tr}\left[(Y'_N)^n\right] = \int_{-\infty}^{\infty} x^n \left(\mathrm{ESD}(Y'_N)\right)(dx) \,,$$

and that the modulus of the above quantity is bounded by M^n, show, by bounded convergence theorem, that (3) and (4) hold, with Y_N replaced by Y'_N. Theorem 1 now shows (14), which, in turn, completes the proof. □

Proof of Theorem 3: As in the preceding proof, let

$$P_N = \psi(Y_N),\ N \geq 1\,.$$

Fix $M > 0$ and let Y'_N be as in (13), M being suppressed in the notation. Theorem 2 implies that

$$\mathrm{EESD}(Y'_N + W_N) \xrightarrow{w} \mu_M \boxplus \nu_\lambda\,,$$

and

$$\mathrm{EESD}(Y'_N W_N) \xrightarrow{w} \mu_M \boxtimes \nu_\lambda$$

as $N \to \infty$, where

$$\mu_M(B) = \mu(B \cap [-M, M]) + \mu\left([-M, M]^c\right) \mathbf{1}(0 \in B)$$

for every Borel set $B \subset \mathbb{R}$. Proposition 4.13 and Corollary 6.7 of [5] imply, respectively, that as $M \to \infty$,

$$\begin{aligned} \mu_M \boxplus \nu_\lambda &\xrightarrow{w} \mu \boxplus \nu_\lambda\,, \text{ and} \\ \mu_M \boxtimes \nu_\lambda &\xrightarrow{w} \mu \boxtimes \nu_\lambda\,. \end{aligned}$$

In view of Facts 3.1 and 3.2, and recalling that convergence in the Lévy metric defined in the latter is equivalent to weak convergence, it suffices to show that

$$\lim_{M\to\infty} \limsup_{N\to\infty} \frac{1}{N} \mathrm{E}\left[\mathrm{Rank}(Y_N - Y'_N)\right] = 0\,.$$

However, arguments as in the proof of Theorem 2 show that for M such that

$$\mu(\{-M, M\}) = 0\,,$$

it holds that

$$\limsup_{N\to\infty} \frac{1}{N} \mathrm{E}\left[\mathrm{Rank}(Y_N - Y'_N)\right] = \mu\left([-M, M]^c\right)\,.$$

Hence the proof follows. □

Acknowledgement

The authors are grateful to an anonymous referee for a careful reading of the manuscript and suggesting changes that helped in improving it.

References

Adhikari, K. and Bose, A. (2019). Brown measure and asymptotic freeness of elliptic and related matrices. *Random Matrices: Theory and Applications*, **8(2)**, 1950007.

Anderson, G. W., Guionnet, A. and Zeitouni. O. (2010). *An Introduction to Random Matrices*. Cambridge University Press.

Azoff, E. A. (1974). Borel measurability in linear algebra. *Proceedings of the American Mathematical Society*, **42(2)**, 346 – 350.

Bai, Z. and Silverstein, J. W. (2010). *Spectral Analysis of Large Dimensional Random Matrices*. Springer Series in Statistics, New York, second edition.

Bercovici, H. and Voiculescu, D. (1993). Free convolution of measures with unbounded support. *Indiana University Mathematics Journal*, **42**, 733–773.

Hiai, F. and Petz, D. (2000). Asymptotic freeness almost everywhere for random matrices. *Acta Scientiarum Mathematicarum (Szeged)*, **66**, 809–834.

Male, C. (2017). The limiting distributions of large heavy Wigner and arbitrary random matrices. *Journal of Functional Analysis*, **272**, 1 – 46.

Mingo, J. A. and Speicher, R. (2017). *Free Probability and Random Matrices*. Springer.

Nica, A. and Speicher, R. (2006). *Lectures on the Combinatorics of Free Probability*. Cambridge University Press, New York.

Voiculescu, D. (1991). Limit laws for random matrices and free products. *Inventiones Mathematicae*, **104(1)**, 201–220.

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 199–208

On Three-Level A-Optimal Designs for Test-Control Discrete Choice Experiments

Rakhi Singh[1], **Ashish Das**[2] **and Feng-Shun Chai**[3]
[1] *University of North Carolina at Greensboro, Greensboro, NC, USA*
[2] *Indian Institute of Technology Bombay, Mumbai, India*
[3] *Academia Sinica, Taipei, Taiwan*

Received: 21 December 2020; Revised: 24 January 2021; Accepted: 27 January 2021

Abstract

Choice experiments are conducted when it is important to study the importance of different factors based on the perceived utility of choice options. We study the optimality of discrete choice experiments under a newly introduced inference problem of test-control discrete choice experiments; it is akin to the test-control inference problem in factorial experiments. For each factor, we have one control level and this control level is then compared with all the test levels of the same factor. For three-level choice designs with multiple factors, we first obtain a lower bound to the A-values for estimating the two test-control contrasts for each factor. We then provide some A-optimal designs for a small number of factors obtained through a complete search. For practical use with a somewhat large number of factors, we then provide some highly efficient designs.

Key words: Choice set; Test-control contrast effects; Hadamard matrix; Multinomial logit model; Linear paired comparison model.

AMS Subject Classifications: 62K05, 05B05

1. Introduction

Discrete choice experiments are used for quantifying the influence of the attributes which characterize the choice options. They are useful in many applied sciences, for example, psychology, marketing research, etc., where options (or, products) have to be judged with respect to a subjective criterion like preference or taste. For a latest application, see Ong *et al.* (2020). In choice experiments, respondents are shown a collection of choice sets and each of these choice sets consists of several options. Respondents are then asked to select one preferred option from each of the choice sets. We consider choice experiments with N choice sets each having two options (referred to as choice pairs hereafter); so, N choice pairs are shown to respondents and they are asked to pick one of the two options that they prefer from each of the N pairs. Each option is described by the same k factors, with each factor having two or more levels. We consider each factor at three levels. A choice design d then is a collection of these N choice pairs. Excellent reviews of the choice designs are provided in Street and Burgess (2012) and Großmann and Schwabe (2015), and a recent paper (Das and

Corresponding Author: Rakhi Singh
Email: r_singh5@uncg.edu

Singh, 2020) provides a unified theory on optimal choice experiments connecting different approaches to choice experiments.

Discrete choice experiments (DCEs) have usually been studied under the multinomial logit model (El-Helbawy and Bradley, 1978; Street and Burgess, 2007). Under the multinomial logit model, D-optimal designs have been studied for several situations, and orthogonal contrasts of main effects and two-factor interactions for the k factors are usually of interest. The multinomial logit model is non-linear, hence, the information matrix is a function of unknown parameters. The locally optimal designs are therefore obtained, and under the indifference assumption (that all treatment combinations have equal utility) of choice experiments, these locally optimal designs have been just termed as optimal designs. DCEs with only two options in each choice set can be equivalently studied under the traditional linear paired comparison model (McFadden, 1974; Huber and Zwerina, 1996; Großmann and Schwabe, 2015). The relationship between the two approaches (MNL models and linear paired comparison models) for studying DCEs has been studied in Das and Singh (2020). They also obtained the information matrices under different inference problems including briefly introducing the test-control inference problem in DCEs. So far, the inference problems that have been studied in choice experiments focus on comparing all levels of each factor with equal importance (Street and Burgess, 2007; Großmann and Schwabe, 2015; Chai *et al.*, 2017). We focus on the test-control inference problem for paired choice DCEs with each factor at three levels. The same setup with a traditional inference problem (of equal focus on all pairwise comparisons) was studied in Chai *et al.* (2017). The difference between the current paper and Chai *et al.* (2017) lies only in the studied inference problem, which ultimately leads to obtaining different optimal designs. We defer most of the technical details until the next section.

The primary goal in a test-control inference problem is to compare the test levels to a (pre-specified) control level. Here, we are not interested in making all pairwise comparisons, we are only interested in a subset of those comparisons. To the best of our knowledge, no one has worked on finding optimal choice designs when the interest might lie in making test-control comparisons. We are also not aware of any practical choice experiment which was conducted with this intention, however, it is not too hard to imagine that such an inference problem will find its use with practitioners. This is useful when manufacturers/service providers or policymakers want to study the effect of new test levels as against the existing control levels. Test-control inference problem has been studied by several authors; see, for example, Hedayat *et al.* (1988) and Majumdar (1996) for block designs, and Gupta (1995) and Gupta (1998) for multiple factors.

D-optimality is invariant to reparameterizations, and thus, D-optimal designs remain optimal even when the inference problem is changed (Großmann and Schwabe, 2015). On the other hand, A-optimal designs change with the inference problem which is one of the reasons behind us studying the A-optimal designs under the inference problem of test-control experiments. For linear models, it has been shown that when the inference problem is test-control, one often benefits by using the A-optimal designs specially designed for catering to this problem (see Banerjee and Mukerjee, 2008, for example). By studying optimal designs for the test-control inference problem for DCEs, we intend to do the same for DCEs (results in Table 2 and the final paragraph). A-optimal designs are the designs that minimize the sum of variances of the treatment contrasts of interest. For example, if the information matrix

for treatment contrasts of interest is M_d, then the design d^* which minimizes $trace(M_d^{-1})$ among all designs is called an A-optimal design. We provide constructions of A-optimal and A-efficient designs for estimating the test-control contrasts under the indifference assumption of the multinomial logit model. We also provide designs having high A-efficiencies.

2. Background

We only present here the details relevant to the current problem, and for more details, Das and Singh (2020) is suggested to be consulted. With each of the k factors at three levels, there are a total of 3^k options. Let the systematic component of the utility for options be denoted by a 3^k-tuple vector τ. Without loss of generality, let the options be arranged lexicographically. For example, for $k = 2$, the systematic component of the utility vector is $\tau = (\tau_{00}, \tau_{01}, \tau_{02}, \tau_{10}, \tau_{11}, \tau_{12}, \tau_{20}, \tau_{21}, \tau_{22})$. The coding that we use is more commonly known as effects coding, see Großmann and Schwabe (2015), for example. In effects coding, for one factor at three levels, level 0 is coded as (1 0), level 1 is coded as (0 1) and level 2 is coded as $(-1\ -1)$; here, level 2 is the control level, and levels 0 and 1 are test levels.

The nth choice pair is denoted by $T_n = (t_{(n1)}, t_{(n2)})$, with $t_{(nj)}$ is the jth option in the nth choice pair, $n = 1, \ldots, N$, $j = 1, 2$. Corresponding to the jth option in N choice sets, $A_j = (t_{(1j)}^T\ t_{(2j)}^T \cdots\ t_{(Nj)}^T)^T$ is a $N \times k$ matrix representing the levels of the k attributes. Let a $N \times 2k$ matrix X_j denote the effects coded matrix corresponding to A_j implying that 0, 1 and 2 in A_j is replaced by the vectors $(1, 0)$, $(0, 1)$ and $(-1, -1)$, respectively, in X_j. Then, the effects coded difference matrix for the first and second option is

$$X = X_1 - X_2. \tag{1}$$

Let B be a $2k \times 3^k$ matrix such that the ith column of B corresponds to the effects coding for the ith option, $i = 1, \ldots, 3^k$. It is assumed that the 3^k options are arranged lexicographically. For example, if $k = 3$, the 3rd column in B would correspond to the effects coding corresponding to the option (002) which is $(1\ 0\ 1\ 0\ -1\ -1)$, or that, the 7th column would correspond to option (020), that is, $(1\ 0\ -1\ -1\ 1\ 0)$. The matrix B has been called B_E in Das and Singh (2020). For simplicity, we drop the subscript E in the current work. This should not be confused with B used in Street and Burgess (2007), since, the matrix B has traditionally corresponded to the orthonormal coding.

The inference problem studied in the current paper is $B\tau$ which corresponds to the situations where the primary interest lies in making test-control comparisons which means that some new levels (called test levels) of factors are compared with an existing control level for the same factor. From Das and Singh (2020), the average information matrix for the inference problem $B\tau$ is $\mathcal{I}(B\tau) = \frac{1}{4N} M_d$ where

$$M_d = (BB^T)^{-1} X^T X (BB^T)^{-1}. \tag{2}$$

Note that the word average here comes from using N in the definition of the information matrix implying that the information considered here is per choice pair. Given the structure of B, it is easy to see that $(BB^T)^{-1} = (\frac{1}{3^k}) diag(V_1^{-1}, \ldots, V_k^{-1})$ where $V_i^{-1} = (3I_2 - J_2)$ for all $i = 1, \ldots, k$. For three-level factors, a choice design d is connected if all the test-control contrasts are estimable, and this happens if and only if M_d has rank $2k$. In what follows, the

class of all connected paired choice designs with k three-level factors and N choice pairs is denoted by $\mathcal{D}_{k,N}$. As stated before, we use the standard A-optimality criteria. The A-value of a design d in $\mathcal{D}_{k,N}$ is $4Ntrace(M_d^{-1})$. A design that minimizes the A-value among all designs in $\mathcal{D}_{k,N}$ is said to be A-optimal.

3. Lower Bounds to the A-value

To find the lower bound to the A-value under the inference problem $B\tau$, we adopt the same strategy as in Chai *et al.* (2017). We first find the lower bound of the A-value for designs with only one factor, and then use the same to find a naïve bound to the A-value when k factors are taken into consideration. Let the matrix X in (1) be partitioned as $(X_{(1)}|X_{(2)}|\cdots|X_{(k)})$, where $X_{(p)}$ is a $N \times 2$ matrix corresponding to the pth factor. Notice that rows in $X_{(p)}$ determine the corresponding options in a design for the pth factor. In $X_{(p)}$, rows $(+2,\ +1)$, $(-2,\ -1)$, $(+1,\ +2)$, $(-1,\ -2)$, $(+1,\ -1)$ and $(-1,\ +1)$ correspond to choice pairs $(0,\ 2)$, $(2,\ 0)$, $(1,\ 2)$, $(2,\ 1)$, $(0,\ 1)$ and $(1,\ 0)$ respectively. Similarly, row $(0,\ 0)$ in $X_{(p)}$ could correspond to any of the three choice pairs $(0,\ 0)$, $(1,\ 1)$ or $(2,\ 2)$.

Similarly, the matrix $M_d = (M_{dpq})$ for a design d can also be partitioned into 2×2 blocks such that the block corresponding to the pth and qth factor is $M_{dpq} = \frac{1}{3^{2k}}(3I_2 - J_2)X_{(p)}^T X_{(q)}(3I_2 - J_2)$; $p = 1, \ldots, k; q = 1, \ldots, k$. It can be shown that we always benefit by not considering the pairs corresponding to the type $(0, 0)$ in $X_{(p)}$ (for an explanation, see Chai *et al.* (2017)). Also, note that $(0, 0)$ in $X_{(p)}$ implies that the corresponding value for a factor in both the options are same. Let y be the number of rows of $X_{(p)}$ that are equal to either $(2,\ 1)$ or $(-2, -1)$ and z be the number of rows of $X_{(p)}$ that are equal to either $(1,\ 2)$ or $(-1, -2)$. Then the remaining $N - y - z$ $(= x$, say) rows of $X_{(p)}$ are necessarily equal to either $(1,\ -1)$ or $(-1,\ 1)$. It can then be shown that for the pth factor,

$$M_{dpp} = \frac{1}{3^{2(k-1)}} \begin{bmatrix} N - z & y + z - N \\ y + z - N & N - y \end{bmatrix} = \frac{1}{3^{2(k-1)}} C_{dpp}, \tag{3}$$

where

$$C_{dpp} = \begin{bmatrix} N - z & y + z - N \\ y + z - N & N - y \end{bmatrix}.$$

We need to obtain a lower bound to $trace(M_{dpp}^{-1}) = 3^{2(k-1)} trace(C_{dpp}^{-1})$, which is equivalent to obtaining a lower bound to $trace(C_{dpp}^{-1})$. The

$$trace(C_{dpp}^{-1}) = (2N - y - z)/h_N(y, z) = g_N(y, z),$$

where

$$h_N(y, z) = det(C_{dpp}) = yz + N(y + z) - (y + z)^2.$$

Note that both $h_N(y, z)$ and $g_N(y, z)$ are symmetric in y and z. We now find the values y and z for which $g_N(y, z)$ is minimized for $1 \leq y + z \leq N$, $y \neq N$, $z \neq N$. These conditions are required for every p so that the design d is connected, that is, $rank(M_d) = 2k$. We need these conditions because our eventual goal is to find a lower bound to the A-value for a design with k factors. Let $\lfloor x \rfloor$ denote the greatest integer contained in x. Let $L_a = min_{d \in \mathcal{D}_{1,N}} trace(C_{dpp}^{-1}) = min_{1 \leq y+z \leq N, y \neq N, z \neq N} g_N(y, z) = g_N(a^*, b^*)$.

Lemma 1: For the pth ($p = 1, \ldots, k$) factor in design $d \in \mathcal{D}_{1,N}$ with $N > 3$,

$$trace(C_{dpp}^{-1}) = g_N(y, z) \geq g_N(a^*, b^*)$$

where $g_N(a^*, b^*) = min\{g_N(a_1, b_1), g_N(a_2, b_2), g_N(a_3, b_3)\}$ with

(i) $a_1 = b_1 = t$,

(ii) $a_2 = b_2 = t + 1$,

(iii) $a_3 = t, b_3 = t + 1$

and $t = \lfloor N(3 - \sqrt{3})/3 \rfloor$.

Proof: The proof follows on similar lines as the proof of corresponding lemma in Chai *et al.* (2017). Treating y and z as continuous variables and adopting the usual derivative approach to minimize $g_N(y, z)$, we get $\partial g_N(y, z)/\partial y = (2N(y + z - N) - yz + y(2N - y - z))/h_N^2(y, z)$. Similarly, $\partial g_N(y, z)/\partial z = (2N(y + z - N) - yz + z(2N - y - z))/h_N^2(y, z)$.

Now, $\partial g_N(y, z)/\partial y = \partial g_N(y, z)/\partial z = 0$ implies that $(y - z)(2N - y - z) = 0$. In other words, y=z, since $2N - y - z > 0$.

Now, for $y = z$, it follows that $\partial g_N(y, z)/\partial y = 0$ implies that $3y^2 - 6Ny + 2N^2$=0 or $y = N(3 \pm \sqrt{3})/3$. However, since $y < N$, the only feasible solution of y is $N(3 - \sqrt{3})/3 = t_1$.

Similarly, checking the matrix of second derivatives, we see that the minimum of $g_N(y, z)$ is attained at $y = z = t_1$. Since t_1 is non-integer, $g_N(y, z) = L_a$ at one of the integer points nearest to (t_1, t_1). □

Using Lemma 1, we have computed the values of a^* and b^* for $4 \leq N \leq 64$ and summarize it in Table 1. Also, note that since $g_N(y, z)$ is symmetric in y and z, interchanging the values of a^* and b^* yield the same values of $g_N(a^*, b^*)$ and therefore from Table 1, we could either say that $y = a^*$ and $z = b^*$ or that $z = a^*$ and $y = b^*$. The optimal value for x can then be computed as $N - y - z$. Recall that we consider choice designs with y number of rows of $X_{(p)}$ equal to either $(2,\ 1)$ or $(-2, -1)$, z number of rows equal to either $(1,\ 2)$ or $(-1, -2)$, and the remaining x rows equal to either $(1,\ -1)$ or $(-1,\ 1)$.

Results in Table 1 are not surprising since we know that for block designs or for factorial experiments, when test-control inference problem is of interest then unequal replication of levels, often with control treatment being repeated more number of times than other treatments, is common. Let $L_A = min_{d \in \mathcal{D}_{k,N}} trace(M_d^{-1})$. We now give a lower bound of L_A for paired choice designs with k factors in N choice pairs.

Theorem 1: For a paired choice design $d \in \mathcal{D}_{k,N}$, $trace(M_d^{-1}) \geq L_A \geq k3^{2(k-1)}L_a = k3^{2(k-1)}g_N(a^*, b^*)$ where a^* and b^* are as in Lemma 1.

Proof: Similar to the proof of Theorem 2.1 in Chai *et al.* (2017), first we apply the inequality $trace(M_d^{-1}) \geq \sum_{p=1}^{k} trace(M_{dpp}^{-1})$ which, using Schur complement and the inverse of partitioned matrices, follows easily for $k = 2$. For example, for $k = 2$, let the 2×2 partitioned

Table 1: Values of a^* and b^* for $4 \leq N \leq 64$ for Lemma 1

N	a^*	b^*	N	a^*	b^*	N	a^*	b^*	N	a^*	b^*	N	a^*	b^*	N	a^*	b^*
4	2	2	14	6	6	24	10	10	34	14	15	44	18	19	54	23	23
5	2	2	15	6	6	25	10	11	35	15	15	45	19	19	55	23	23
6	2	3	16	7	7	26	11	11	36	15	15	46	19	20	56	24	24
7	3	3	17	7	7	27	11	12	37	15	16	47	20	20	57	24	24
8	3	4	18	7	8	28	12	12	38	16	16	48	20	20	58	24	25
9	4	4	19	8	8	29	12	12	39	16	17	49	21	21	59	25	25
10	4	4	20	8	9	30	13	13	40	17	17	50	21	21	60	25	26
11	4	5	21	9	9	31	13	13	41	17	17	51	21	22	61	26	26
12	5	5	22	9	9	32	13	14	42	18	18	52	22	22	62	26	26
13	5	6	23	10	10	33	14	14	43	18	18	53	22	23	63	26	27
															64	27	27

matrix M be $M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}$. Then

$$M^{-1} = \begin{bmatrix} (M_{11} - M_{12}M_{22}^{-1}M_{21})^{-1} & -M_{11}^{-1}M_{12}(M_{22} - M_{21}M_{11}^{-1}M_{12})^{-1} \\ -M_{22}^{-1}M_{21}(M_{11} - M_{12}M_{22}^{-1}M_{21})^{-1} & (M_{22} - M_{21}M_{11}^{-1}M_{12})^{-1} \end{bmatrix}.$$

Since $M_{12}M_{22}^{-1}M_{21}$ is non-negative definite, $(M_{11} - M_{12}M_{22}^{-1}M_{21}) \leq M_{11}$ and therefore $(M_{11} - M_{12}M_{22}^{-1}M_{21})^{-1} \geq M_{11}^{-1}$. Similarly, $(M_{22} - M_{21}M_{11}^{-1}M_{12})^{-1} \geq M_{22}^{-1}$. Therefore, $trace(M^{-1}) = trace(M_{11} - M_{12}M_{22}^{-1}M_{21})^{-1} + trace(M_{22} - M_{21}M_{11}^{-1}M_{12})^{-1} \geq trace(M_{11}^{-1}) + trace(M_{22}^{-1})$. Now, using the method of induction, one can see that the inequality holds for a general k, that is, the $trace(M^{-1}) \geq \sum_{p=1}^{t} trace(M_{pp}^{-1})$. Finally, using Lemma 1, the proof follows. □

In the next section, we provide some A-optimal designs attaining the lower bounds of Theorem 1. In some situations, since we are not able to provide designs attaining the A-lower bounds, A-efficiencies are given.

4. Design Constructions

A design $d \in \mathcal{D}_{k,N}$ would be A-optimal under the test-control inference problem if $trace(M_d^{-1})$ attains the bound obtained in Theorem 1. To attain this bound, the design should not only have the values of a^* and b^*, for each factor, as in Table 1 (or, Lemma 1) but should also satisfy the orthogonality property, that is, the blocks M_{dpq} for $p \neq q = 1, \ldots, k$ should be block matrices with all values equal to 0s. The closer these block matrices are to zero matrices, the higher is the efficiency expected to be. This is a somewhat hard combinatorial problem, less studied, and it is more difficult to deal with the problem as compared to finding designs for other inference problems. For finding optimal designs, an algorithm such as the one recently studied in case of factorial experiments (Chai and Das, 2020) would be more helpful. The A-efficiency of a design $d \in \mathcal{D}(k, N)$ is given by

$$\phi_A = \frac{min_{d_0 \in \mathcal{D}(k,N)} trace(M_{d_0}^{-1})}{trace(M_d^{-1})}.$$

It then follows from Theorem 1 that for $d \in \mathcal{D}_{k,N}$

$$\phi_A \geq \frac{k3^{2(k-1)}g_N(a^*, b^*)}{trace(M_d^{-1})}. \tag{4}$$

As an example, consider $N = 9$ and $k = 2$. Here, the optimal values for a^* and b^* are both equal to 4, and, from the result of complete search, we see that no design with $a^* = b^* = 4$ for $N = 9$ achieves the bound in Theorem 1. One of the designs d_9 with the smallest A-value (= 282.8769) among the designs having $a^* = b^* = 4$ is provided below. In fact, it is surprising to note that designs for $a^* = 3$ and $b^* = 4$ (or vice versa) while satisfying the orthogonality condition have A-value (= 274.1538) which is smaller than d_9. This design is provided below as d_9^+. The bound from Theorem 1 (= $k3^{2(k-1)}g_N(a^*, b^*)$) is 7.5, and, therefore, the bound for A-value is then $4N(7.5) = 270$. Thus, the design d_9 has an A-efficiency of at least 0.9545 whereas d_9^+ has an A-efficiency of at least 0.9848.

Similarly for $N = 7, k = 2$, no design with $a^* = b^* = 3$ achieves the bound in Theorem 1. Design d_7 with the smallest A-value (= 279.7321) among the designs having $a^* = b^* = 3$ is provided below. The designs having either $a^* = 2$, $b^* = 3$ or vice versa for one factor and $a^* = b^* = 3$ for another factor and additionally satisfying the orthogonality condition have A-value (276.1500) which is smaller than d_7. This design is also given below as d_7^+. The bound from Theorem 1 is 9.6, and, therefore the bound for A-value is then $4N(9.6) = 268.8$. Thus, the design d_7 has an A-efficiency of at least 0.9609 whereas d_7^+ has an A-efficiency of at least 0.9734.

$$d_9 = \begin{pmatrix} 00, & 22 \\ 01, & 22 \\ 02, & 11 \\ 02, & 20 \\ 02, & 21 \\ 10, & 22 \\ 11, & 20 \\ 12, & 20 \\ 12, & 21 \end{pmatrix}, \; d_9^+ = \begin{pmatrix} 00, & 22 \\ 01, & 10 \\ 01, & 22 \\ 02, & 11 \\ 02, & 21 \\ 10, & 22 \\ 11, & 20 \\ 12, & 20 \\ 12, & 21 \end{pmatrix} \text{ and } d_7 = \begin{pmatrix} 00, & 22 \\ 01, & 22 \\ 02, & 11 \\ 02, & 20 \\ 10, & 22 \\ 11, & 20 \\ 12, & 21 \end{pmatrix}, \; d_7^+ = \begin{pmatrix} 00, & 22 \\ 01, & 22 \\ 02, & 11 \\ 02, & 20 \\ 10, & 21 \\ 11, & 20 \\ 12, & 21 \end{pmatrix}$$

The designs obtained for $N = 7$ and $N = 9$ suggest that the lower bound in Theorem 1 is not tight. In fact, it suggests that orthogonality is somewhat more important than the designs satisfying the property in Lemma 1 for every factor. For $k = 2$ and $N = 4, 5, 6$ and 8, A-optimal designs have been obtained using complete search and reported below as d_4, d_5, d_6, and d_8, respectively. These designs satisfy the orthogonality property and satisfy the values of a^* and b^* in Lemma 1 thereby attaining the optimal bound in Theorem 1. In fact, the complete search result also shows that d_7^+ and d_9^+ have the smallest A-value and are, therefore, A-optimal. Note that there exists more than one design with the same A-values and only one of them is reported here. Besides, these complete searches are carried out within the class of designs having distinct choice pairs since that is more desirable in practice (Chai *et al.*, 2017).

$$d_4 = \begin{pmatrix} 01, & 22 \\ 02, & 21 \\ 10, & 22 \\ 12, & 20 \end{pmatrix}, d_5 = \begin{pmatrix} 00, & 22 \\ 01, & 10 \\ 02, & 21 \\ 11, & 22 \\ 12, & 20 \end{pmatrix}, d_6 = \begin{pmatrix} 00, & 12 \\ 01, & 20 \\ 02, & 21 \\ 10, & 22 \\ 11, & 22 \\ 12, & 21 \end{pmatrix} \text{ and } d_8 = \begin{pmatrix} 00, & 22 \\ 01, & 22 \\ 02, & 11 \\ 02, & 20 \\ 10, & 21 \\ 11, & 22 \\ 12, & 20 \\ 12, & 21 \end{pmatrix}.$$

A Hadamard matrix H_m is a $m \times m$ matrix with elements ± 1 such that $H_m^T H_m = H_m H_m^T = mI_m$. Using the construction in Chai *et al.* (2017) and designs reported above, called *base designs*, we now find designs with larger numbers of factors k and $N \geq 8$. Consider a base design $d_0 \in \mathcal{D}_{k_0,N_0}$ with the trace $(M_{d_0}^{-1}) = k_0 3^{2(k_0-1)} g_{N_0}(a_0, b_0)$. Using d_0, a paired choice design d_H with parameters $k = mk_0$, $N = mN_0$ is constructed with the corresponding design matrix $X_H = H_m \otimes X$, where X is the design matrix of d_0. This method of construction obtains a final design by performing Kronecker product of the small design with a Hadamard matrix. To find the A-value of design d_H, we first note that

$$X_H^T X_H = H_m^T H_m \otimes X^T X = mI_m \otimes X^T X.$$

Then, from (2), we have

$$\begin{aligned} M_{d_H} &= \frac{1}{3^{2k}} diag(V_1^{-1}, \ldots, V_k^{-1})(mI_m \otimes X^T X) diag(V_1^{-1}, \ldots, V_k^{-1}) \\ &= \frac{1}{3^{2mk_0}} diag(V_1^{-1}, \ldots, V_k^{-1})(mI_m \otimes X^T X) diag(V_1^{-1}, \ldots, V_k^{-1}) \\ &= \frac{m}{3^{2k_0(m-1)}} I_m \otimes M_{d_0}. \end{aligned} \tag{5}$$

Therefore, $trace(M_{d_H}^{-1}) = 3^{2k_0(m-1)} trace(M_{d_0}^{-1})$, and the A-efficiency of d_H is given by

$$\phi_A \geq \frac{k 3^{2(k-1)} g_N(a^*, b^*)}{3^{2k_0(m-1)} trace(M_{d_0}^{-1})} = \frac{k 3^{2(k_0-1)} g_N(a^*, b^*)}{trace(M_{d_0}^{-1})} = \phi_A^*, \tag{6}$$

where d_0 is a base design in $\mathcal{D}_{k_0,N_0}$ and a^* and b^* are as in Theorem 1 for a design with N runs and k factors. In Table 2, for $N \geq 4$ and $k \geq 2$, we provide A-optimal and A-efficient designs with efficiency bounds as in (6), and the corresponding methods of constructions. One of the designs with $k = 4$ and $N = 10$ is A-optimal. We denote designs d_H obtained using the Hadamard matrix H_m and a base design d_{N_0} by $H_m \otimes d_{N_0}$. A design with a smaller k retains its optimality property for given N when factors are deleted from a design with larger k. As is expected, designs that are obtained using d_7^+ and d_9^+ are better (higher efficiency) than the designs obtained using d_7 and d_9. Note that the A-efficiencies of designs reported in Table 2 could actually be higher than the reported lower bounds.

We have obtained optimality bounds for the test-control inference problem for DCEs. From Table 2, we see that the designs obtained for $k \geq 2$ are highly efficient. It is worth observing that corresponding designs obtained in Chai *et al.* (2017) are less efficient as

Table 2: A-optimal/efficient designs with distinct choice pairs ($N \geq 4$ and $k \geq 2$)

k	N	ϕ_A^*	Method	k	N	ϕ_A^*	Method	k	N	ϕ_A^*	Method
2	4	1	d_4	4	14	0.9609	$H_2 \otimes d_7$	8	36	0.9502	$H_4 \otimes d_9$
2	5	1	d_5	4	14	0.9734	$H_2 \otimes d_7^+$	8	36	0.9805	$H_4 \otimes d_9^+$
2	6	1	d_6	4	16	0.987	$H_2 \otimes d_8$	16	32	0.9338	$H_8 \otimes d_4$
2	7	0.9609	d_7	4	18	0.9526	$H_2 \otimes d_9$	16	40	0.9953	$H_8 \otimes d_5$
2	7	0.9734	d_7^+	4	18	0.9829	$H_2 \otimes d_9^+$	16	48	0.9778	$H_8 \otimes d_6$
2	8	1	d_8	8	16	0.9351	$H_4 \otimes d_4$	16	56	0.9609	$H_8 \otimes d_7$
2	9	0.9545	d_9	8	20	0.9973	$H_4 \otimes d_5$	16	56	0.9734	$H_8 \otimes d_7^+$
2	9	0.9848	d_9^+	8	24	0.9778	$H_4 \otimes d_6$	16	64	0.9849	$H_8 \otimes d_8$
4	8	0.9474	$H_2 \otimes d_4$	8	28	0.9609	$H_4 \otimes d_7$	16	72	0.9501	$H_8 \otimes d_9$
4	10	1	$H_2 \otimes d_5$	8	28	0.9734	$H_4 \otimes d_7^+$	16	72	0.9803	$H_8 \otimes d_9^+$
4	12	0.9778	$H_2 \otimes d_6$	8	32	0.9856	$H_4 \otimes d_8$				

compared to the designs reported in Table 2 under the current test-control inference problem. For example, the design for $k = 2, N = 4$ reported in Chai *et al.* (2017) (given below for convenience) is shown to be both A- and D-optimal under the traditional inference problem of pairwise comparisons, but it is only 83% efficient under the current test-control inference problem.

Design for $k = 4$ in Chai *et al.* (2017) is $\begin{pmatrix} 20, & 01 \\ 21, & 10 \\ 12, & 00 \\ 02, & 11 \end{pmatrix}$. Note that this design has with $a = b = 1$, whereas the optimal design for test-control inference problem should have $a^* = b^* = 2$ from Table 1.

Acknowledgements

We thank the reviewer for helping us improve the presentation in the paper. Ashish Das's work is supported by the grant MTR/2020/000209 of the Science and Engineering Research Board, India. The research of F.-S. Chai was supported by the grant MOST 106-2118-M-001-011 from the Ministry of Science and Technology of Taiwan.

References

Banerjee, T. and Mukerjee, R. (2008). Optimal factorial designs for c-DNA microarray experiments. *The Annals of Applied Statistics*, **2(1)**, 366–385.

Chai, F. S. and Das, A. (2020). A-optimal factorial designs for test versus control comparisons. *Journal of Statistical Theory and Practice*, **14(4)**, 1–8.

Chai, F. S., Das, A. and Singh, R. (2017). Three-level A- and D-optimal paired choice designs. *Statistics & Probability Letters*, **122**, 211–217.

Das, A. and Singh, R. (2020). Discrete choice experiments—a unified approach. *Journal of Statistical Planning and Inference*, **205**, 193–202.

El-Helbawy, A. T. and Bradley, R. A. (1978). Treatment contrasts in paired comparisons: Large-sample results, applications, and some optimal designs. *Journal of the American Statistical Association*, **73(364)**, 831–839.

Graßhoff, U., Großmann, H., Holling, H. and Schwabe, R. (2004). Optimal designs for main effects in linear paired comparison models. *Journal of Statistical Planning and Inference*, **126(1)**, 361–376.

Großmann, H. and Schwabe, R. (2015). Design for discrete choice experiments. In A Dean, M Morris, J Stufken, and D. Bingham, editors, *Handbook of Design and Analysis of Experiments*, pages 787–831. Chapman and Hall/CRC, Boca Raton.

Gupta, S. (1995). Multifactor designs for test versus control comparisons. *Utilitas Mathematica*, **47**, 199–210.

Gupta, S. (1998). A class of multi-factor designs for test versus control comparisons. *Journal of Statistical Planning and Inference*, **72(1)**, 291–302.

Hedayat, A. S., Jacroux, M. and Majumdar, D. (1988). Optimal designs for comparing test treatments with controls. *Statistica Sinica*, **4**, 462–491.

Huber, J. and Zwerina, K. (1996). The importance of utility balance in efficient choice designs. *Journal of Marketing Research*, **33**, 307–317.

Majumdar, D. (1996). Optimal and efficient treatment-control designs. In S Ghosh and C.R. Rao, editors, *Handbook of Statistics 13: Design and Analysis of Experiments*, pages 1007–1053. Elsevier, North-Holland, Amsterdam.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P Zarembka, editor, *Frontiers in Econometrics*, pages 105–142. Academic Press, New York.

Ong, J. J., Lourenco, R. D. A, Street, D. J., Smith, K., Jamil, M. S., Terris-Prestholt, F., Fairley, F. C., McNulty, A., Hynes, A., Johnson, K., *et al.* (2020). The preferred qualities of human immunodeficiency virus testing and self-testing among men who have sex with men: A discrete choice experiment. *Value in Health*, **23(7)**, 870–879.

Street, D. J. and Burgess, L. (2007). *The Construction of Optimal Stated Choice Experiments: Theory and Methods*. John Wiley & Sons, Hoboken, New Jersey.

Street, D. J. and Burgess, L. (2012). Designs for choice experiments for the multinomial logit model. In K Hinkelmann, editor, *Design and Analysis of Experiments, Special Designs and Applications*, **Volume 3**, pages 331–378. John Wiley & Sons, Inc, Hoboken, New Jersey.

Statistics and Applications ISSN 2454-7395 (online)
Volume 19, No. 1, 2021 (New Series), pp 209–221

Stratified Subsampling Based *p*-values for Hypothesis Tests in Genomics Research

Sudesh Pundir[1,2*], Yanrong Ji[1*], Arunima Shilpi[1] and Ramana V. Davuluri[3]
[1]*Division of Health and Biomedical Informatics, Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA.*
[2]*Department of Statistics, Pondicherry University, Pondicherry, India*
[3]*Department of Biomedical Informatics, Stony Brook University, Stony Brook, NY, USA.*

Received: 14 August 2020; Revised: 30 January 2021; Accepted: 04 February 2021

Abstract

Multiple testing, which refers to testing of more than one hypothesis in an experiment, is routinely performed in statistical analysis of genome-wide data, such as testing the association of single-nucleotide polymorphisms (SNPs) with a particular phenotype. A common practice is application of multiple-testing correction methods to exclude candidate SNPs that could otherwise be spuriously marked as statistically significant. However, in many cases such methods are overly conservative and often result in no significant SNPs at all. In this paper, we summarize commonly used multiple-testing correction procedures and Monte Carlo simulation-based methods. We propose a simple modification to subsampling-based simulation method to estimate empirical *p*-values by borrowing the principles of stratified sampling. Using real datasets from the cancer genome atlas (TCGA) data repository, we demonstrate that the traditional multiple testing correction methods yielded almost none or very few significant risks associated SNPs, whereas the proposed stratified subsampling successfully resulted in appropriate number of significant candidate SNPs. We also show that the proposed modification has provided meaningful *p*-values and made the test more powerful as compared to simple subsampling without stratification.

Key words: Multiple comparison test; Subsampling; Stratified sampling; *p*-value.

1. Introduction

With the exponential growth of the omics data, computational analysis of large datasets has become commonplace in the study of human biology and disease. The sampled subjects, on which the data is collected, usually differ by sex, race, age and ethnicity, leading to heterogeneous data. The research presented here is motivated by statistical analyses of such genome-scale data, *e.g.*, The Cancer Genome Consortium Data (Ding *et. al.*, 2018), involving multiple comparisons of thousands of genomic features between heterogeneous populations. While human genome sequences are mostly identical between different individuals, a small number of genetic differences exist that result in the striking phenotypic variation observed among individuals. Studying the association between genetic and phenotypic variation and identification of disease associated genetic variants and their prevalence across different

Corresponding Author: Ramana V. Davuluri;
Email: Ramana.Davuluri@stonybrookmedicine.edu
*Equal contribution

populations have been the subjects of numerous genome projects since the publication of the human genome (Lander *et. al.*, 2001; Landrum *et. al.*, 2018).

The most common genetic variation is single nucleotide polymorphism (SNP), which roughly occur every 1200 base pairs in comparisons of a pair of human chromosomes. For example, dbSNP database provides a general catalog of SNPs that are characterized according to frequency, distribution among populations and functional genomic regions, potential functional consequences, inferred mutation pattern, linkage, and organization within each chromosome in each individual (haplotype) (Sherry *et. al.*, 2001; Neykov *et. al.*, 2019). These SNPs can help discern small differences both within a population and among different populations, leading to the identification of population based risk genetic variants for common and complex diseases.

Motivating Example: Cancer is a complex genetic disease with significant heterogeneity across patients. Molecular understanding of tumor heterogeneity is key to effective cancer treatment and personalized medicine. High-grade serous ovarian carcinoma (HGSOC) accounts for 70 to 80 percent of ovarian cancer deaths, with little improvement in overall survival in recent years (Siegel *et. al.*, 2016). The standard therapy for HGSOC includes maximal cytoreductive surgery followed by platinum and taxane chemotherapy. While the majority of HGSOC patients respond to initial treatment, most tumors recur and become increasingly resistant to chemotherapy, with an overall 5-year survival rate of approximately 30 percent (Reid *et. al.*, 2017). As a heterogeneous disease, understanding how genetic differences in individuals contribute to their cancer susceptibility and response to therapy can help guide medical practitioners to give the best advice to achieve a favorable outcome for the patient. As genome technologies evolve, genotyping of individuals could be available to all patients using a simple saliva test. Large-scale genome-wide association studies and meta analyses have provided powerful insights into SNPs that may be predictive of disease and an individual's length of survival (or response to therapy). For example, The Cancer Genome Atlas (TCGA) data portal (https://portal.gdc.cancer.gov) provides multiple layers of -omics data (*e.g.* gene expression, methylation, SNPs) along with clinical/phenotypic information (*e.g.* cancer stage, survival information, drugs/treatment information) for more than 1500 ovarian cancer patients (Cancer Genome Atlas Research, 2011; GTExProject, 2017). These data provide an unprecedented opportunity for exploratory data analysis to identify SNPs that are associated with cancer, survival status and response therapy. It is expected that the catalogue of such SNPs will provide the foundation for tailored detection, prevention and treatment of diseases leading to the era of personalized cancer medicine (Dayem Ullah *et. al.*, 2018). One common goal in large genome-wide experiments is to identify the genomic markers (*e.g.* genes or SNPs) that are significantly different between different populations or associated with a response or covariate of interest. The response could be censored survival time or other clinical outcomes, the covariates could be either categorical (*e.g.* treatment/control status, cancer subtype) or continuous (*e.g.* dose of a drug).

In the above example of ovarian cancer data, our main goal is to identify the SNPs that are associated with patient survival. Log rank test is the most widely used test for testing the equality of survival distributions between different patient populations. However, a major challenge in the analysis and interpretation of such large-scale genome studies is the simultaneous handling of multiple comparisons, where a large number of genes or SNPs (or null hypotheses) are simultaneously tested. For example, let us suppose that an experiment involves 100 SNPs to be tested, each with a Type 1 error probability of 0.05, assuming the null hypothesis is true for each SNP the expected number of false significant SNPs is equal to 5.

Moreover, if all tests are mutually independent, then the probability that at least one true null hypothesis will be rejected is given by $1–0.95^{100} = 0.994$. Therefore, in any large genome-wide study involving large number of SNPs (usually more than a million), any truly significant calls will be accompanied by correspondingly large number of false findings.

2. Multiple-Testing Correction Methods

Multiple-testing correction methods adjust the significance level for each test to a value α such that the overall type I error for the study (the probability of rejecting a correct null hypothesis in at least one of the tests) will not exceed a predetermined acceptable level, often set to 0.05. Widely accepted approaches to deal with the multiple-testing problem control either the family wise error rate (FWER), which is the probability of at least one false rejection (Hochberg and Tamhane, 1987), or the false discovery rate (FDR), which is the expected proportion of falsely rejected null hypotheses (Reiner *et. al.*, 2003; Benjamini and Yekutieli, 2005).

For example, Bonferroni correction, which controls FWER, for testing the SNPs that are associated with survival, is performed as

(i) Compute *p*-values using log rank test.
(ii) Reject the null hypothesis for $p_l \leq \frac{\alpha}{m}$.

where m is the total number of comparisons/tests we are performing, or total number of hypotheses.

Similarly, Benjamini-Hochberg (BH) correction, which controls FDR, is performed by following step-wise procedure.

(i) Sort the *p*-values in increasing order.
(ii) For a given α, find the largest l such that $p_l \leq \frac{l}{m}\alpha$, where m is again the total number of hypotheses to be tested, and l is the rank of SNPs.
(iii) Reject the null hypotheses for all $H_{(m)}$, m=1, 2, …, l.

Resampling-based multiple-testing correction methods: Resampling-based multiple testing procedures are widely used in genome data analysis, especially when the sample size is small or the distribution of test statistic does not follow normality assumption or is unknown. Resampling-based multiple testing procedures can account for dependent structures among *p*-values or test statistics, resulting in lower type II errors. The commonly used resampling techniques include permutation tests and bootstrap methods.

In permutation tests, the distribution of the test statistics is constructed by calculating all possible values or a sufficiently large number of test statistics (usually 1000 or above) from permuted sampling observations under the null hypothesis. Permutation tests are distribution-free, which can provide exact *p*-values even when sample size is small. Bootstrap method finds an approximate distribution of the test statistic by taking many repeated samples with replacement from one random sample (Efron and Tibshirani, 1994). The bootstrap method provides an asymptotically unbiased estimator for the variance of a sample median and for error rates in a linear discrimination problem outperforming cross-validation (Efron, 1979).

The *p*-values obtained by the bootstrap method are less exact than *p*-values computed from the permutation method, and the bootstrap estimated *p*-values are asymptotically convergent to the true *p*-values (Pollard and van der Laan, 2004). Please refer (Farcomeni, 2008) for a review of multiple hypothesis testing procedures and applications in the analysis of DNA microarray data.

Subsampling-based multiple-testing correction methods: Subsampling procedure is different from resampling technique. While resampling scheme generates multiple samples (of size equal to the original sample size) by choosing the observations from the sample with replacement, subsampling scheme selects the observations from the sample without replacement. Statistical inference based on the samples of fixed size in resampling but in case of subsampling scheme, the inference is drawn on the samples of smaller size than fixed sample size. The samples are drawn in resampling technique by using simple random sampling with replacement (SRSWR), whereas in subsampling the samples are drawn by simple random sampling without replacement (SRSWOR). Subsampling (or Two stage sampling) at few places in the literature should not be confused with subsampling defined by (Politis and Romano, 1993). Technically, while two-stage sampling is a two-stage-sampling scheme, subsampling is resampling method without replacement by selecting a smaller size subsamples from the original sample. For example, (Nigam and Rao, 1996) constructed second order balanced designs when sample size (n) is a composite and prime number, and extended the results to stratified multistage samples and provided inferential procedures on balanced bootstrap for stratified multistage samples.

The distribution of Studentized statistic was estimated by subsampling by (Politis and Romano, 1993). They constructed confidence regions by approximating the sampling distribution of a statistic based on the values of the statistic computed over small subsets of the data, and showed their method works well under weak assumptions (Politis and Romano, 1994). In the subsequent publications, they approximated the sampling distribution of a statistic based on the values of the statistic computed over small subsets of the data, and illustrated its application on time series data (Politis and Romano, 1996). Their book provides some of the foundation for subsampling methodology and related methods (Politis *et. al.*, 1999). Further, the asymptotic theory of subsampling was discussed in (Politis *et. al.*, 2001), and *K*-sample subsampling for iid observations and time series data were discussed by (Politis and Romano, 2008). In a later publication, they constructed the confidence intervals and *p*-values for the tests based on subsampling by shortening the number of iterations (Berg *et. al.*, 2010). They showed that the new *p*-values were asymptotically uniform under the null hypothesis and converged to zero under alternative hypothesis, leading to improved power of the test and meaningful *p*-values.

The application of subsampling methods for assessing the significance of observations in large-scale genome studies was discussed in (Bickel *et. al.*, 2010). Recently, a subsampling without replacement-based normalization scheme was employed for identification of differentially expression that accounted for the hierarchy and amplitude of effect sizes within samples (Mohorianu *et. al.*, 2017). Xavier *et al.* (Xavier *et. al.*, 2017) proposed the use of subsampling bootstrap Markov chain in genomic prediction. The proposed method consists of fitting whole-genome regression models by subsampling observations in each round of a Markov Chain Monte Carlo. Further, the subsampling based approach was effectively used for determining appropriate sequencing depth trough efficient read subsampling of RNA-seq data (Robinson and Storey, 2014).

In this paper, we propose a modification to the subsampling scheme, by performing stratified sampling without replacement. Because of the complex and heterogeneous nature of disease population (cancer patients in the current study), there is a need to account for the heterogeneity; such as race, living status, cancer type *etc*. Using a real example of TCGA ovarian and brain cancer data, we demonstrate that dividing the heterogeneous data into strata and then applying subsampling approach leads to more meaningful empirical *p*-values for log rank test. We also show that traditional multiple testing correction methods seem to be too strict for studies on a genomics scale, whereas the proposed stratified subsampling approach can successfully result in appropriate number of significant observations. In the following section, we begin by introducing the basic principle of stratified subsampling.

3. Stratified Subsampling

In stratified subsampling, instead of drawing a subsample of size $b<<n$, we first partition the sample into non-overlapping groups, and then subsamples without replacement are drawn within each stratum as explained below. Strata are non-overlapping and homogeneous with respect to the characteristic under study. For example, in a survival analysis study based on genome sequencing data from cancer patients, the sample usually consists of both living patients and diseased, usually with varying proportions. If subsamples are drawn without accounting for this heterogeneity, the subsamples may disproportionately consist of one group versus the other, therefore, leading to spurious *p*-values. Here, we propose an approach to statistical significance in the analysis of genome-wide data sets, based on the concept of stratified sub-sampling *p*-values.

Procedure of stratified subsampling:

1. Divide the sample of N units into k strata. Let the i^{th} stratum have n_i, i=1,2, …, k, number of units, such that $N = \sum_{i=1}^{k} n_i$.
2. Draw a subsample of size b_i from sample of size n_i from i^{th} stratum using SRSWOR.
3. All the subsampling units drawn from each stratum will constitute a stratified sample of size b.

Let us define the following symbols as
k: Number of strata
n_i: Numbers of sampling units to be drawn from i^{th} stratum
b_i: Number of subsampling units to be drawn from i^{th} stratum
$n = \sum_{i=1}^{k} n_i$: Total sample size
$b = \sum_{i=1}^{k} b_i$: Total subsample size.

Let $x_n = (X_1, X_2, \ldots, X_n)$ be a sample of n independent and identically distributed (iid) random variables taking values in an arbitrary sample space S with unknown probability distribution P. P belongs to a class of distributions H *w*hich may be parametric, nonparametric or semiparametric. The idea is to approximate the sampling distribution of a statistic based on the values of the statistic computed over smaller subsets of the data.

Let $t(P)$ be the parameter and its estimator (or statistic) is given by

$$t_n = f(X_1,\ X_2,\ \ldots,\ X_n).$$

Then the sampling distribution of the statistic is given by

$$J_n(x,P) = P\{\tau_n[(t_n - t(p)) \le x]\},$$

where τ_n is a normalizing sequence.

The fundamental idea behind subsampling is that $J_n(x,P)$ can be accurately approximated by the normalized distribution of the same estimator calculated on appropriately data chosen subsets of data of size b ($b<<n$).

Let the statistic calculated on the i^{th} subset of size b is denoted by

$$t_{n,b,i} = t_b(X_{i1}, X_{i2}, \ldots, X_{ib}).$$

Let $N_n = \binom{n}{b}$ be the total number of available subsets of the data of size b. In this case, the i^{th} subsample is constructed by sampling without replacement from iid data with purpose of forming a subsample of size b.

The subsampling estimator of $J_n(x,P)$ is defined as follows

$$L_{n,b}(x) = \frac{1}{N_n}\sum_{i=1}^{N_n} I[\tau_b(t_{n,b,i} - t_n) \le x]$$

where I is the indicator function. Under general conditions

$$N_n \to \infty, b \to \infty, \quad \frac{b^k}{n} \to 0$$

and for the appropriate values of k and τ_n is such that $\frac{\tau_b}{\tau_n} \to 0$ whenever $\frac{b}{n} \to 0$. Politis and Romano (1994) showed that

$$L_{n,b}(x) - J_n(x,P) \xrightarrow{P} 0.$$

Let the hypotheses for testing the parameter be

$$H_0: t(P) = \theta_0, \qquad P \in P_0$$
$$H_1: t(P) > \theta_0, \qquad P \in P_1$$

The sampling distribution of the statistic under null hypothesis is given by

$$J_n(x,P_0) = P[\tau_n(t_n - \theta_0) \le x]$$

and its subsampling estimator is given by

$$L_{n,b}(x,P_0) = \tfrac{1}{N_n}\textstyle\sum_{i=1}^{N_n} I[\tau_b(t_{n,b,i} - \theta_0) \le x].$$

Politis *et al.* (1999) gave the proof of the consistency of the test. The test rejects H_0 when

$$\frac{1}{N_n}\sum_{i=1}^{N_n} I[\tau_n(t_n - \theta_0) \ge \tau_b(t_{n,b,i} - \theta_0)] > 1 - \alpha$$

$$\frac{1}{N_n}\sum_{i=1}^{N_n} I[T_n \ge T_{n,b,i}] > 1 - \alpha$$

Under null hypothesis, the subsampled distribution of $T_{n,b,I}$ approximates the sampling distribution of T_n .

Stratified subsampling and Log rank test: Log rank test is the most widely used test for testing the equality of survival distributions. Let

Y: Time until an event occurs where event is death of person
T: Failure time with distribution function $F(x)$ and probability density function $f(x)$
C: Censoring time with distribution function $G(x)$ and probability density function

$$\Delta = min(T, C) = \begin{cases} 1, & T \leq C \\ 0, & T > C \end{cases}$$

Survival function is defined as the probability that a person will survive beyond a time *t*. It is defined as

$$S(t) = P(Y > t) = \int_t^{\infty} f(x)dx = 1 - F(t), \qquad 0 < t < \infty.$$

Consider the following $q \times 2$ table classifying those with and without the event of interest

Group	Event		Total
	Dead at time T_i	Alive at time T_i	
0	D_{0i}	$N_0(T_i) - D_{0i}$	$N_0(T_i)$
1	D_{1i}	$N_1(T_i) - D_{1i}$	$N_1(T_i)$
2	D_{2i}	$N_2(T_i) - D_{2i}$	$N_2(T_i)$
.	.	.	.
.	.	.	.
.	.	.	.
q	D_{qi}	$N_q(T_i) - D_{qi}$	$N_q(T_i)$
Total	D_i	$N(T_i) - D_i$	$N(T_i)$

where $T_1, \ldots, T_i, \ldots, T_k$ are distinct failure times
$N_g(T_i)$: Number of persons in group g at risk at T_i
D_{gi}: Number of persons in group g who fail at T_i, $g = 0,1, 2, \ldots, q$, $i = 1, 2, \ldots, k$.

D_{gi} follows hypergeometric distribution.

The hypotheses for testing the survival functions of different groups are given as

$$H_0: S_0(t) = S_1(t) = \cdots = S_q(t)$$

H_1: Two or more Survival functions are different from others.

The log rank test statistic for testing the above hypotheses is defined as

$$\chi = \frac{\sum_{i=1}^{k}(O_i - E_i)}{\sqrt{\sum_{i=1}^{k} V_i}}$$

where O_i: Observed number of failures, E_i: Expected number of failures, V_i: Variance of observed number of failures. Under H_0, χ (or χ^2) follows standard normal (or chi-square) distribution approximately. This approximation is generally used to obtain an approximate test

for H_0 by comparing the observed value of χ (or χ^2) to the tail area of the standard normal (or chi-square) distribution.

Empirical *p*-values based on Monte Carlo simulations: Monte Carlo simulations are routinely applied in permutation and resampling based methods to estimate the *p*-values. Suppose χ^2_{obs} is the observed Chi-squared test-statistic value for a given random sample from log rank test. In Monte Carlo simulations, independent random datasets are generated using pseudo-random number either by resampling or subsampling methods. Assuming *m* such data sets are simulated under the null hypothesis, each yielding a distinct test statistic χ^2_{sim}, the ideal *p*-value is $p_\infty = P(\chi^2_{sim} > \chi^2_{obj})$. However, p_∞ is unknown, because generating infinite number of datasets is not possible and only a finite number (*m*) of datasets are available. Let *B* be the number of times out of m that $\chi^2_{sim} > \chi^2_{obs}$. It was previously shown that the unbiased estimator $\hat{p}_\infty = B/m$ leads to an invalid test that does not correctly control the type I error rate at the required level (Phipson and Smyth, 2010), therefore, computing the tail probability directly for the Monte Carlo results was suggested as a valid approach. Therefore, in a randomization test, the test statistic is *B* rather than χ^2_{obs}, and the required tail probability is *P(B ≤ b)*. It was shown by (Phipson and Smyth, 2010) that, under the null hypothesis, the marginal distribution of *B* over all possible data sets is discrete uniform on the integers from 0, …, *m*, and the exact Monte Carlo *p*-value is estimated as

$$P_u = P(B \leq b) = \frac{b+1}{m+1}.$$

While this is not an unbiased estimator, the amount of positive bias is just enough to allow for the uncertainty of estimation and to produce a test with the correct size. For further details about this *p*-value calculation, please refer (Edgington and Onghena, 2007; Phipson and Smyth, 2010).

4. Application on real-life datasets

In order to compare our stratified subsampling scheme with other multiple testing correction methods, we have applied our method on two real-life datasets: SNP array data of TCGA ovarian cancer (OV, 570 patients, 580,886 SNPs) and low-grade glioma (LGG, 505 patients, 251,258 SNPs). Each patient has three potential genotypes: AA (reference), Aa (heterozygous) and aa (alternative), for each SNP. We associated their survival functions with the genotypes and used log rank test to determine the statistical significance of the overall survival difference between 3 genotypes. For each genotype, we further stratified on the vital status of the patient, and drew random subsamples with different number of subsampling percentage (60%, 70%, 80%) for n = 500 and 1,000 iterations. We then compared our stratified subsampling scheme with other methods for multiple testing correction, including Bonferroni and Benjamini-Hochberg procedures, Bootstrapping method, as well as subsampling scheme without stratification, by plotting an empirical distribution of χ^2-statistic from log rank test. Specifically, we compared the χ_1^2, …, χ_m^2, …, χ_n^2 with the χ_0^2 obtained using the original un-permuted sample, and computed the empirical *p*-value based on the Monte Carlo empirical p-value formula below (and introduced in previous section):

$$P_{emp} = \frac{r+1}{M+1},$$

where *r* is the total number of iterations that $\chi_m^2 > \chi_0^2$.

Table 1 shows the comparison of number of significant SNPs declared at different thresholds for OV and LGG respectively.

Table 1: Number of significant SNPs for different α

Method	Ovarian (OV)			Brain (LGG)		
	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$
Single Sample Logrank test	8,228	35,257	66,855	5,275	18,917	33,645
Multiple comparison-Bonferroni Correction	1	2	2	4	4	8
Multiple comparison-Benjamini Hochberg Correction	2	2	12	4	67	234
Bootstrapping	0	0	0	0	0	0
Stratified Subsampling with size 60% (500 iterations)	2	542	5,943	44	2,164	11,872
Stratified Subsampling with size 70% (500 iterations)	0	46	1,396	4	630	5,298
Stratified Subsampling with size 80% (500 iterations)	0	1	107	0	71	1,022

Table 1 shows that traditional multiple testing correction methods, including Bonferroni and Benjamini-Hochberg procedures, as well as resampling-based method (Bootstrapping), all did not control number of significant findings to an appropriate level as they appear to be too stringent on a genomics scale, where hundreds of thousands of tests are performed simultaneously. This fact is more apparent when we compare across ovarian cancer (total 580,886 SNPs) and glioma (total 251,258 SNPs), where less SNPs results in more significant candidates after multiple testing correction due to the less total number of tests performed in LGG. Compare with the methods above, stratified subsampling provided more candidates at different levels, across the two datasets. Moreover, decrease of subsampling percentage seems to be able to provide additional relaxation, allowing number of candidates to be controlled by adjusting the subsampling parameters.

It is also noteworthy that bootstrapping gives no significant candidates in our case no matter what cutoff we chose. To potentially elucidate why this happens, as well as why a larger subsample size results in smaller number of significant candidates, we plotted the estimated sampling distributions in these cases for the particular SNP with lowest *p*-value from single log rank test (rs10824799 for OV, rs7754576 for LGG), with increased number (10,000) of iterations (Figures 1 and 2).

Table 2 shows that for both OV and LGG random subsampling tends to give less significant candidates as compared to stratified subsampling, indicating that it may again be too strict. Moreover, the random subsampling returns similar number of candidates as Benjamini-Hochberg approach in both cases. Since there is much more computation associated with subsampling approach compared to traditional multiple testing correction methods, applying simple random subsampling does not seem to offer any advantage. We can see from the examples and comparison that stratification can best capture the heterogeneity within the sample while not being too stringent. In this example, the number of strata is 2, with stratification based on living status – dead or living. However, the stratification and number of strata can be modified depending on other attributes, such as, race, ethnicity, sex, *etc.*, provided such information is available and the sample size is large enough to yield desired power.

Table 2: Number of significant SNPs for different α

Subsampling percentage	Ovarian (OV)			Brain (LGG)		
	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$
60% (random)	0	19	666	0	28	438
70% (random)	0	1	75	1	2	72
80% (random)	0	0	2	0	1	2
60% (stratified)	2	542	5,943	44	2,164	11,872
70% (stratified)	0	46	1,396	4	630	5,298
80% (stratified)	0	1	107	0	71	1,022

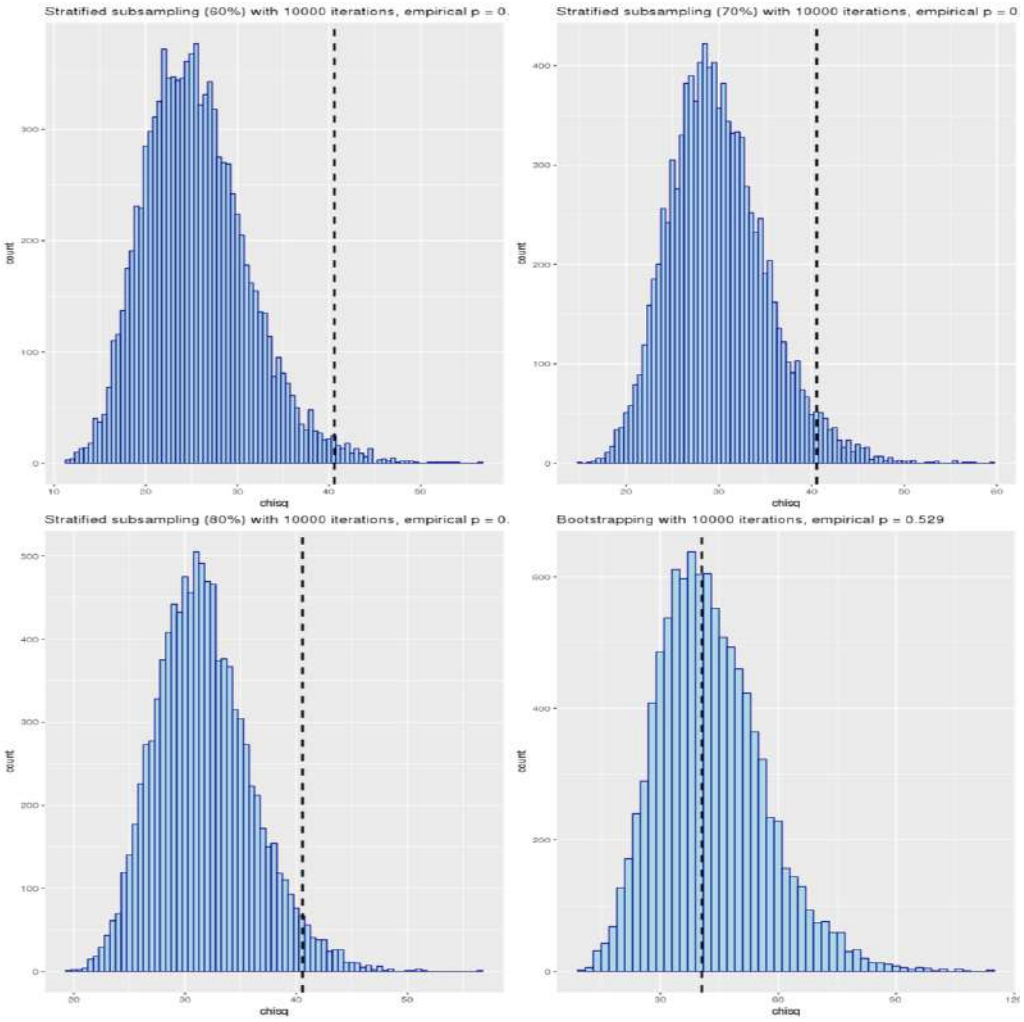


Figure 1: Distribution of simulated test-statistic (χ^2_{sim}) based on stratified subsampling with 60% (top left), 70% (top right), 80% (bottom left) and bootstrapping (bottom right) with 10,000 iterations for OV. Black line indicates the actual test statistic value on the overall sample (χ^2_{obs}).

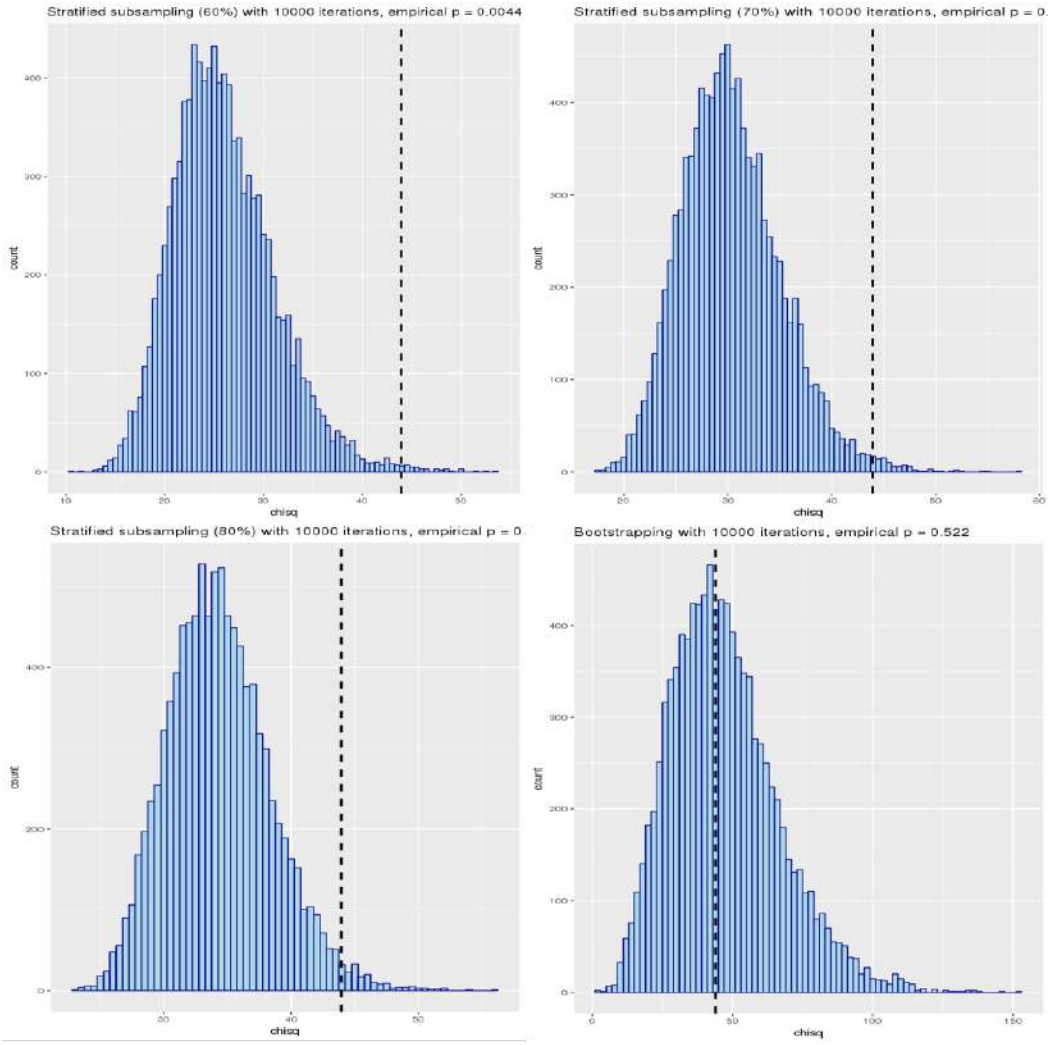


Figure 2: Distribution of simulated test-statistic (χ^2_{sim}) based on stratified subsampling with 60% (top left), 70% (top right), 80% (bottom left) and bootstrapping (bottom right) with 10,000 iterations for LGG. Black line indicates the actual test statistic value on the overall sample (χ^2_{obs}).

5. Conclusions

In this paper, we introduced the concept of stratified subsampling for constructing *p*-values for hypothesis tests in genomics research and showed that it can effectively handle the problem of multiple testing while not being too conservative. While the stratified subsampling based empirical *p*-values are proposed for the log rank test, the method can be generalized for any other statistical test. The proposed modification can be applied in case of heterogeneous data and when subsampling is performed to construct the *p*-values.

Based on the empirical evaluation, we found that the simple random subsampling returned much less significant SNPs than stratified subsampling, suggesting that the simple random subsampling is also too stringent, and considering the computational burden, subsampling based *p*-values (without stratification) do not have advantages over traditional multiple testing correction (*e.g.* BH, Bonferroni), as they similarly returned very few candidates. We are currently working on theoretical aspects of constructing the confidence

intervals and *p*-values based on the stratified subsampling procedure proposed here. In addition, further work is needed to derive and evaluate the asymptotic properties of the proposed test-statistic under the null and alternative hypotheses.

Acknowledgements

This work was supported by the National Library of Medicine of the NIH [R01LM011297 to RD]. We thank the reviewer and Dr. V.. Gupta for their suggestions and thoughtful comments, which substantially helped the revised version.

References

Benjamini, Y. and Yekutieli, D. (2005). Quantitative trait Loci analysis using the false discovery rate. *Genetics*, **171**, 783-790.

Berg, A., McMurry, T. L. and Politis, D. N. (2010). Subsampling *p* values. *Statistics and Probability Letters*, **80**, 1358-1364.

Bickel, P. J., Boley, N., Brown, J. B., Huang, H. and Zhang, N. R. (2010). Subsampling methods for genomic Inference. *The Annals of Applied Statistics*, **4**, 1660–1697.

Cancer Genome Atlas Research, N. (2011). Integrated genomic analyses of ovarian carcinoma. *Nature*, **474**, 609-615.

Dayem Ullah, A. Z., Oscanoa, J., Wang, J., Nagano, A., Lemoine, N. R. and Chelala, C. (2018). SNPnexus: assessing the functional relevance of genetic variation to facilitate the promise of precision medicine. *Nucleic Acids Research*, **46**, W109-W113.

Ding, L., Bailey, M. H., Porta-Pardo, E., Thorsson, V., Colaprico, A., Bertrand, D., Gibbs, D. L., Weerasinghe, A., Huang, K. L., Tokheim, C. *et al*. (2018). Perspective on Oncogenic Processes at the End of the Beginning of Cancer Genomics. *Cell*, **173**, 305-320 e310.

Edgington, E. S. and Onghena, P. (2007). *Randomization tests*. Chapman & Hall/CRC, Boca Raton, FL.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, **8**, 1-26.

Efron, B. and Tibshirani, R. (1994). *An Introduction to the Bootstrap.* CRC Press, New York.

Farcomeni, A. (2008). A review of modern multiple hypothesis testing, with particular attention to the false discovery proportion. *Statistical Methods in Medical Research*, **17**, 347-388.

GTExProject. (2017). Enhancing GTEx by bridging the gaps between genotype, gene expression, and disease. *Natural Genetics*, **49**, 1664-1670.

Hochberg, Y. and Tamhane, A. C. (1987). *Multiple comparison procedures*. Wiley, New York.

Lander, E. S. Linton, L. M. Birren, B. Nusbaum, C. Zody, M. C. Baldwin, J. Devon, K. Dewar, K. Doyle, M. FitzHugh, W. et al. (2001). Initial sequencing and analysis of the human genome. *Nature*, **409**, 860-921.

Landrum, M. J., Lee, J. M., Benson, M., Brown, G. R., Chao, C., Chitipiralla, S., Gu, B., Hart, J., Hoffman, D., Jang, W. *et al*. (2018). ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Research*, **46**, D1062-D1067.

Mohorianu, I., Bretman, A., Smith, D. T., Fowler, E. K., Dalmay, T. and Chapman, T. (2017). Comparison of alternative approaches for analysing multi-level RNA-seq data. *PLoS One*, **12**, e0182694.

Neykov, M., Lu, J. and Liu, H. (2019). Combinatorial Inference for Graphical Models. *The Annals of Statistics*, **47**, 795-827.

Nigam, A. K. and Rao, J. N. K. (1996). On balanced bootstrap for stratified multistage samples. *Statistica Sinica*, **6**, 199-214.

Phipson, B. and Smyth, G. K. (2010). Permutation P-values should never be zero: calculating exact P-values when permutations are randomly drawn. *Statistical Applications in Genetics and Molecular Biology*, **9**, Article39.

Politis, D. N. and Romano, J. P. (1993). Estimating the distribution of studentized statistic by subsampling. *Bulletin of the International Statistical Institute*, **49**, 315-316.

Politis, D. N. and Romano, J. P. (1994). Large sample confidence regions based on subsamples under minimal assumptions. *The Annals of Statistics*, **22**, 2031-2050.

Politis, D. N. and Romano, J. P. (1996). Subsampling for econometrics models- Comments on bootstrapping time series models. *Econometric Review*, **15**,169-176.

Politis, D. N. and Romano, J. P. (2008). K-sample subsampling in general spaces: the case of independent time series. *Journal of Multivariate Analysis*, **101**, 316-326.

Politis, D. N., Romano, J. P. and Wolf, M. (1999). *Subsampling*. Springer-Verlag, New York.

Politis, D. N., Romano, J. P. and Wolf, M. (2001). On the asymptotic theory of subsampling. *Statistica Sinica*, **11**, 1105-1124.

Pollard, K. S. and van der Laan, M. K. (2004). Choice of a null distribution in resampling-based multiple testing. *Journal of Statistical Planning and Inference*, **125**, 85-100.

Reid, B. M., Permuth, J. B. and Sellers, T. A. (2017). Epidemiology of ovarian cancer: a review. *Cancer Biology and Medicine*, **14**, 9-32.

Reiner, A., Yekutieli, D. and Benjamini, Y. (2003). Identifying differentially expressed genes using false discovery rate controlling procedures. *Bioinformatics*, **19**, 368-375.

Robinson, D. G. and Storey, J. D. (2014). subSeq: determining appropriate sequencing depth through efficient read subsampling. *Bioinformatics*, **30**, 3424-3426.

Sherry, S. T., Ward, M. H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M. and Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. *Nucleic Acids Research*, **29**, 308-311.

Siegel, R. L., Miller, K. D. and Jemal, A. (2016). Cancer statistics, 2016. *CA - A Cancer Journal for Clinicians*, **66**, 7-30.

Xavier, A., Xu, S., Muir, W. and Rainey, K. M. (2017). Genomic prediction using subsampling. *BMC Bioinformatics*, **18**, 191.

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 223–246

Locally Optimal Binary Crossover Designs

Siuli Mukhopadhyay[1], Satya Prakash Singh[2] and Arpan Singh[2]
[1]*Department of Mathematics, Indian Institute of Technology Bombay, Mumbai 400 076, India*
[2]*Department of Mathematics, Indian Institute of Technology Hyderabad, Telangana 502285, India*

Received: 11 December 2020; Revised: 03 February 2021; Accepted: 05 February 2021

Abstract

Optimal two-treatment, p period crossover designs for binary responses are determined. The optimal designs are obtained by minimizing the variance of the treatment contrast estimator over all possible allocations of n subjects to 2^p possible treatment sequences. An appropriate logistic regression model is postulated and the within subject covariances are modeled through a working correlation matrix. The marginal mean of the binary responses are fitted using generalized estimating equations. The efficiencies of some crossover designs for $p = 2, 3, 4$ periods are calculated. An equivalence theorem is provided to verify optimality of numerically obtained locally optimal designs.

Key words: Binary response; Generalized estimating equations; Logistic regression; Efficiency.

AMS Subject Classifications: MSC: 62K05

1. Introduction

In crossover trials, every experimental unit receives a sequence of treatments over different time periods. For the real life applications of crossover trials see *e.g.*, Jones and Kenward (2014) and Senn (2003). The problem of determining optimal designs for crossover trials has been studied quite extensively in recent years and we refer to Bose and Dey (2009) for a review of results on optimal crossover designs. However, most of the available results on optimal crossover designs relate to situations where the response variable is continuous (see Kershner and Federer (1981), Laska and Meisner (1985), Matthews (1987) and Carriere and Huang (2000) and the references therein). In clinical or pharmaceutical research, the outcome of interest is often binary in nature. While methods for analyzing binary data arising from crossover trials are available in Jones and Kenward (2014) and Senn (2003), the question of designing such studies in an optimal manner does not seem to have been addressed much in the literature. Waterhouse *et al.* (2006) considered crossover designs for binary response, where the treatments were taken to be continuous in nature and no period effects were considered in the model. Singh and Mukhopadhyay (2016) proposed optimal

Correponding Author: Siuli Mukhopadhyay
Email: siuli@math.iitb.ac.in

crossover Bayesian designs for the generalized linear models (GLMs). One of their case study was based on a four periods ($p = 4$) binary crossover design for four periods. Recently, Singh *et al.* (2020) proposed min–max crossover designs for the GLMs. Following the methodology proposed in Singh and Mukhopadhyay (2016), Jankar *et al.* (2020) proposed locally D-optimal designs for the GLMs. In comparison aforementioned references, the present article discusses the binary crossover design in greater details for $p = 2, 3, 4$ periods and two treatments. We discuss optimal designs for binary responses in a logistic regression framework.

Since the main interest lies in the estimation of the treatment effect, the designs proposed, minimize the variance of the estimator associated with the treatment effect. For a binary logistic model, the variance of the treatment effect estimator depends on the model parameters, to address the issue of parameter dependence, various intervals of model parameter are assumed and a subset of parameter values are selected from these intervals. In crossover studies the response at the current time period may have the effect of the treatment from the previous time period. This effect is refer to as the "carryover effect". Often the interest lies in estimating the carryover effect. Optimal crossover designs to estimate the carryover effect for the normal response are discussed in Laska and Meisner (1985) and Gondaliya and Divecha (2015). In our setting assuming that the carryover effect of a treatment lasts only to the next succeeding period, optimal designs for estimating the carryover effect are also discussed. A population average approach is utilized for the estimation of the model parameters. In this approach we treat the subject effects as a nuisance parameter and use the generalized estimating equations of Liang and Zeger (1986) to estimate the marginal means. The observations from each subject over different time points are assumed to be mutually correlated while the observations from different subjects are uncorrelated. The correlation between observations within subjects are modeled using a "working correlation structure". We study the effect of three working correlation structures, uncorrelated, equi-correlated and autoregressive (AR) on the designs chosen. The rest of the article is organised as foloows. In Section 2, we define the crossover logistic model for a binary response and discuss the estimation of the crossover model using generalized estimating equations. In Section 3, results on optimal two-treatment designs for 2, 3 and 4 periods are given. The optimally of numerically obtained locally optimal designs is verified using an equivalence theorem given in Section 3.5.

2. The Model and Estimation

Consider a crossover trial involving t treatments, n subjects and p periods. Suppose the response obtained from the jth subject is $\mathbf{Y}_j = (Y_{1j}, \ldots, Y_{pj})'$, where a prime denotes transposition. Instead of specifying a joint distribution of the repeated measurements we use a working GLM to describe the marginal distribution of Y_{ij} as (Liang and Zeger, 1986)

$$f(y_{ij}) = exp[\{y_{ij}\phi_{ij} - b(\phi_{ij}) + c(y_{ij})\}\psi].$$

For a binary random variable Y_{ij}, $\phi_{ij} = \log \dfrac{\mu_{ij}}{1-\mu_{ij}}$, $b(\phi_{ij}) = \log[1 + \exp\{\phi_{ij}\}]$, $c(y_{ij}) = 0$, and the scale parameter ψ is 1 (Robinson and Khuri, 2003). The mean of Y_{ij} is μ_{ij} and variance $\mu_{ij}(1 - \mu_{ij})$.

In a crossover setup, we model the marginal mean μ_{ij} using the population-average model

$$logit(\mu_{ij}) = \eta_{ij} = \mu + \beta_i + \tau_{d(i,j)} + \rho_{d(i-1,j)};\ i = 1\ \ldots, p,\ j = 1, \ldots, n, \tag{1}$$

where μ is the overall mean, β_i represents the effect of the ith period, τ_s is the treatment effect due to treatment s and ρ_s is the carryover effect due to treatment s, $s = 1, \ldots, t$. Throughout, $\mathbf{1}_u$ is a $u \times 1$ vector of all ones, I_u is the identity matrix of order u and 0_{ab} is an $a \times b$ null matrix. Also, we write $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)'$, $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_t)'$ and $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_t)'$. Since there is no carryover effect in the first period, we set $\rho_{d(0,j)} = 0$ for all j.

In matrix notation, the linear predictor corresponding to the jth subject, $\boldsymbol{\eta}_j = (\eta_{1j}, \ldots, \eta_{pj})'$, can be written as

$$\boldsymbol{\eta}_j = \mathbf{X}_j\boldsymbol{\theta}, \tag{2}$$

where $\boldsymbol{\theta} = (\mu, \boldsymbol{\beta}', \boldsymbol{\tau}', \boldsymbol{\rho}')'$. The design matrix is $\mathbf{X}_j = [\mathbf{1}_p\ P_j\ T_j\ F_j]$, where $P_j = I_p$; T_j is a $p \times t$ matrix with its (i, s)th entry equal to 1 if subject j receives the treatment effect of the treatment s in the ith period and zero otherwise; F_j is a $p \times t$ matrix with its (i, s)th entry equal to 1 if subject j receives the carryover effect of the treatment s in the ith period and zero otherwise. Note that the first row of F_j consists of all zeros since $\rho_{d(0,j)} = 0$ for all j.

Since we are working with the population–average model, the estimation of the model parameters can be done using the generalized estimating equation (GEE) approach proposed by Liang and Zeger (1986) and Zeger et al. (1988). The GEEs are utilized to estimate the parameters of GLM with a possible unknown correlation between outcomes. The resulting estimators are referred to as the GEE estimators. The GEE estimators are consistent even if the correlation structure is misspecified. It is assumed that measurements from the same subject in the p periods are correlated while observations from different subjects are uncorrelated. The dependencies between repeated observations from a subject are modeled using a "working correlation" matrix $C(\boldsymbol{\alpha})$ where $\boldsymbol{\alpha}$ is a vector of length s. If $C(\boldsymbol{\alpha})$ is the true correlation matrix of $\mathbf{Y}_j$, then

$$Cov[\mathbf{Y}_j] = D_j^{1/2}C(\boldsymbol{\alpha})D_j^{1/2}, \tag{3}$$

where $D_j = diag(\mu_{1j}(1 - \mu_{1j}), \ldots, \mu_{pj}(1 - \mu_{pj}))$. Let $W_j = D_j^{1/2}C(\boldsymbol{\alpha})D_j^{1/2}$.

For a repeated-measures model, Zeger *et al.* (1988, equation (3.1)) derived the generalized estimating equations (GEE) to be

$$\sum_{j=1}^{n} \frac{\partial \boldsymbol{\mu}_j'}{\partial \boldsymbol{\theta}} W_j^{-1}(\mathbf{Y}_j - \boldsymbol{\mu}_j) = 0,$$

where $\boldsymbol{\mu}_j = (\mu_{1j}, \ldots, \mu_{pj})'$. The asymptotic variance for the GEE estimator $\widehat{\boldsymbol{\theta}}$ (see Zeger *et al.*, 1988, equation (3.2)) is

$$Var(\widehat{\boldsymbol{\theta}}) = \left[\sum_{j=1}^{n} \frac{\partial \boldsymbol{\mu}_j'}{\partial \boldsymbol{\theta}} W_j^{-1} \frac{\partial \boldsymbol{\mu}_j}{\partial \boldsymbol{\theta}}\right]^{-1}, \tag{4}$$

if $Cov(\mathbf{Y}_j) = W_j$, *i.e.* the working correlation is same as the true correlation. However, if the true correlation structure varies from the "working correlation" structure, then $Var(\widehat{\boldsymbol{\theta}})$ is given by the sandwich formula (Zeger *et al.*, 1988, equation (3.2))

$$Var(\widehat{\boldsymbol{\theta}}) = A^{-1}BA^{-1},$$

where

$$A = \sum_{j=1}^{n} \frac{\partial \boldsymbol{\mu}_j'}{\partial \boldsymbol{\theta}} W_j^{-1} \frac{\partial \boldsymbol{\mu}_j}{\partial \boldsymbol{\theta}}, \quad \text{and} \quad B = \sum_{j=1}^{n} \frac{\partial \boldsymbol{\mu}_j'}{\partial \boldsymbol{\theta}} W_j^{-1} Cov(\mathbf{Y}_j) W_j^{-1} \frac{\partial \boldsymbol{\mu}_j}{\partial \boldsymbol{\theta}}.$$

For the crossover model (1), the ith element of $\frac{\partial \boldsymbol{\mu}_j}{\partial \boldsymbol{\theta}}$ is $\mathbf{x}_{ij}'\mu_{ij}(1-\mu_{ij})$, where $\mathbf{x}_{ij}'$ is the ith row of $\mathbf{X}_j$ for $i = 1, \ldots, p$.

Before introducing the design selection criterion, we list the main objectives of the paper with the help of the following example. Consider a trial reported by Senn (2003, page 127) wherein it was desired to study the effect of two drugs on 24 children aged 7 to 13 suffering from exercise-induced asthma. The two treatments were, a single dose of $12\mu g$ formoterol solution aerosol (treatment A) and a single dose of $200\mu g$ of salbutamol solution aerosol (treatment B). Each child was given both the treatments either in the order, AB or BA. The response variable was binary, taking value 1 if the drug was effective and 0 otherwise. An equal number of children were allocated to each treatment sequence, AB or BA. Several questions arise about the design used:

- Is the design with equal allocation to sequences $\{AB, BA\}$ optimal for the binary model? If not which is the optimal design?
- For continuous responses [Laska and Meisner (1985)], in a 2–periods 2–treatments (2×2) crossover study, proved that the design with equal allocation to treatment sequences $\{AB, BA\}$ is optimal when there are no carryover effects in the model. If the same design is used for binary model what is the efficiency loss, if any?
- In binary models design selection depends on the model parameters. What will be the effect of these parameters when selecting a crossover design?
- Will the design change in a binary model if we include carryover effects in the model?

Finding an exact optimal design (optimal number of subjects to the treatment sequences) which is associated with the integer optimization problem of a non-linear function is mathematically intractable. Instead to find optimal crossover designs for the binary model we use the approximate theory as in Laska and Meisner (1985) and Kushner (1997, 1998). For a review of results on optimal crossover designs using the approximate theory, we refer to Bose and Dey (2009, Chapter 4). An approximate/continuous crossover design with k treatment sequences can be expressed in the form of a probability measure as follows:

$$\zeta = \begin{Bmatrix} \boldsymbol{\omega}_1 & \boldsymbol{\omega}_2 & \ldots & \boldsymbol{\omega}_k \\ p_{\omega_1} & p_{\omega_2} & \ldots & p_{\omega_k} \end{Bmatrix},$$

where $\boldsymbol{\omega}_i \in \Omega$ (set of all permutations of t treatments of length p). Observe that Ω denotes the set of all possible treatment sequences of length p. Here $p_{\boldsymbol{\omega}_i}$ is the proportion of subjects assigned to treatment sequence $\boldsymbol{\omega}_i$. Fixing the number of subjects to n and periods to p, we determine the proportion of subjects assigned to a particular treatment sequence. We denote by n_ω the number of subjects assigned to sequence $\boldsymbol{\omega}$. Then, $n = \sum_{\boldsymbol{\omega}\in\Omega} n_\omega, n_\omega \geq 0$, $p_{\boldsymbol{\omega}_i} = n_{\boldsymbol{\omega}_i}/n \geq 0$ and $\sum_{i=1}^{k} p_{\boldsymbol{\omega}_i} = 1$, for $i = 1, \cdots, k$. Once an approximate optimal design is obtained, an exact design can be found by efficient rounding (Kiefer (1971), Pukelsheim and Rieder (1992)).

It follows from Lemma 4.2.1 in Bose and Dey (2009) that $T_{\boldsymbol{\omega}} = T_j$ and $F_{\boldsymbol{\omega}} = F_j$ for all j subjects assigned to a treatment sequence $\boldsymbol{\omega}$. This implies that $\mathbf{X}_j = \mathbf{X}_{\boldsymbol{\omega}}$. Since $np_{\boldsymbol{\omega}}$ subjects are assigned to sequence $\boldsymbol{\omega}$, the variance of $\widehat{\boldsymbol{\theta}}$ in (4) can be expressed as

$$Var_\zeta(\widehat{\boldsymbol{\theta}}) = U^{-1} = \left[\sum_{\boldsymbol{\omega}\in\Omega} np_{\boldsymbol{\omega}} \frac{\partial \boldsymbol{\mu}'_{\boldsymbol{\omega}}}{\partial \boldsymbol{\theta}} W_{\boldsymbol{\omega}}^{-1} \frac{\partial \boldsymbol{\mu}_{\boldsymbol{\omega}}}{\partial \boldsymbol{\theta}}\right]^{-1}. \tag{5}$$

For the estimation of the treatment effect, instead of working with the full variance-covariance matrix of $\widehat{\boldsymbol{\theta}}$ we concentrate on $Var(\widehat{\boldsymbol{\tau}})$ where,

$$Var_\zeta(\widehat{\boldsymbol{\tau}}) = HVar_\zeta(\widehat{\boldsymbol{\theta}})H', \tag{6}$$

where H is a $t \times m$ matrix given by $[0_{t1}, 0_{tp}, I_t, 0_{tt}]$ and m is the total number of parameters in $\boldsymbol{\theta}$.

A locally optimal design (LOD) ζ^* is one which minimizes the $\log(Var_\zeta(\widehat{\boldsymbol{\tau}}))$ with respect to p_w, when $\sum_{w\in\Omega} p_w = 1$ and $\omega \in \Omega$. Similarly, an LOD associated with the carryover effect can be obtained by minimizing $\log(Var_\zeta(\widehat{\boldsymbol{\rho}}))$. As an example, consider the case when $p = 3$ and $\Omega = \{AAA, AAB, ABB, ABA, BAB\ BAA, BBA, BBB\}$. An optimal design (approximate/continuous) is specified by the optimal proportions p^*_ω for each $\omega \in \Omega$ for which $Var_\zeta(\widehat{\boldsymbol{\tau}})$ is minimized with respect to these proportions. In other words, ζ^* determines the optimal proportion p^*_ω of the total observations assigned to the treatment sequences ω. Suppose,

$$\zeta^* = \begin{Bmatrix} AAA & AAB & ABB & ABA & BAB & BAA & BBA & BBB \\ 0.40 & 0 & 0 & 0.15 & 0.10 & 0.25 & 0 & 0.10 \end{Bmatrix},$$

Then the optimal design allocates 40% of subjects to treatment AAA, 15% to treatment ABA, 10% subjects to BAB and BBB, and 25% to BAA. No subjects are allocated to the treament sequences $\{AAB, ABB, BBA\}$. In equation (8) we note that the variance of the treatment effect estimator depends on the model parameters. Thus, the optimal design found by minimizing the variance of the treatment effect is parameter dependent and actually an LOD.

3. Two Treatment Crossover Trials: Results and Discussion

With two treatments of interest, the problem simplifies to minimizing the variance of the treatment contrast $\tau_1 - \tau_2$ to obtain optimal crossover designs. Reparameterizing

model (1) as in Singh and Mukhopadhyay (2016) and Laska and Meisner (1985), using $\tau = (\tau_1 - \tau_2)/2$ and $\rho = (\rho_1 - \rho_2)/2$ we get

$$logit(\mu_{ij}) = \mu + \beta P + \tau \Phi_{d(i,j)} + \rho \Phi_{d(i-1),j)}, \tag{7}$$

where P takes value 0 for period 1 and 1 for period 2, $\Phi_A = 1$,$\Phi_B = -1$ and $\Phi_{d(0,j)} = 0$.

For illustration we go back to the example in Section 2, where there are two treatments, A and B applied in two periods to each child. The design used involved the treatment sequences AB and BA, with equal allocation to each treatment sequence. Thus, the matrix $\mathbf{X}_{\boldsymbol{\omega}}$ depends on the treatment sequence $\boldsymbol{\omega} \in \Omega = \{AB, BA\}$. If the treatment sequence, for example is $\boldsymbol{\omega} = AB$, then

$$\mathbf{X}_{\boldsymbol{\omega}} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & -1 & 1 \end{pmatrix}$$

In the following, we look at the performance of the design, $\{AB, BA\}$. Senn (2003) fitted a logistic model with no carryover effect to the data set and computed confidence intervals for the various components of $\boldsymbol{\theta}$. Using these intervals we investigate if the above two-period design is the best choice in the given situation. We also look at general situations for determining optimal designs when $p = 2, 3$ or 4 for two treatment case.

3.1. Designs compared

An optimal design obtained by considering all possible treatment sequences associated with a p period model is denoted by $D^{(p)}$. For example when $p = 2$, $D^{(2)}$ is consists of the treatment sequences $\{AA, AB, BA, BB\} = \Omega$, with optimal proportions p^*_ω associated to the treatment sequence $\omega \in \Omega$. The designs that we consider are similar to those discussed by Laska and Meisner (1985) and Carriere and Huang (2000) and are listed below for $p = 2, 3$ and 4. The notation $D_i^{(p)}$ denotes ith design considered for a model with p time periods.

(i) $p = 2$:

$D_1^{(2)}$: AB and BA; $D_2^{(2)}$: AB, AA, BA and BB, with equal number of subjects assigned to each sequence. For normal responses, when there is no carryover effect, $D_1^{(2)}$ is an optimal design [Grizzle (1965)]. Design $D_2^{(2)}$ is shown to be universally optimal [Carriere and Reinsel (1992)].

(ii) $p = 3$:

$D_1^{(3)}$: ABB and BAA;

$D_2^{(3)}$: ABB, AAB, BAA and BBA;

$D_3^{(3)}$: ABB, ABA, BAA and BAB.

In designs $D_1^{(3)} - D_3^{(3)}$, each treatment sequence is allocated equally. These designs are shown to be optimal under different scenarios for normal responses. Under appropriate assumption on the within-subject correlation, Kershner (1986) and Laska *et al.* (1983) shown that $D_1^{(3)}$ is an universally optimal design. Optimality of designs $D_2^{(3)}$ and $D_3^{(3)}$ was investigated by

Laska *et al.* (1983), Ebbutt (1984), Matthews (1987), Carriere (1994), and Carriere and Huang (2000) for normal responses.

(iii) $p = 4$:

$D_1^{(4)}$: $AABB$, $BBAA$, $ABBA$ and $BAAB$;

$D_2^{(4)}$: $AABB$, $BBAA$;

$D_3^{(4)}$: $ABBA$, $BAAB$;

$D_4^{(4)}$: $ABAB$, $BABA$.

In designs $D_1^{(4)} - D_4^{(4)}$, each treatment sequence is allocated equally. The performances of these designs are investigated in Gondaliya and Divecha (2018).

For evaluating and comparing the above designs we define an efficiency measure as

$$\Gamma(\zeta) = \left(\frac{Var_{\zeta^*}(\widehat{\tau})}{Var_{\zeta}(\widehat{\tau})}\right)^{1/m}, \tag{8}$$

where ζ^* is the locally optimal crossover design and m is the number of unknown regression parameters in the model. Note that the efficiency in (8) defined for designs associated with the estimation of the treatment effect. Efficiency of designs associated with the estimation of the carryover effect can be defined by replacing $\widehat{\tau}$ with $\widehat{\rho}$ in (8).

3.2. Working correlation structures

We consider the uncorrelated, compound symmetric or, equi-correlated and the AR(1) structures for the correlation matrix $C(\boldsymbol{\alpha})$. Under the equi-correlated covariance structure, $C(\boldsymbol{\alpha}) = (1-\alpha)I_p + \alpha J_p$, where α is a scalar.

Under the AR(1) assumption the (i,j)th element c_{ij} fo $C(\boldsymbol{\alpha})$ is,

$$c_{ij} = \alpha^{|i-j|},\ i \neq j.$$

A Working Example: Here we present an example to illustrate the proposed methodology for obtaining the optimal proportions and compute design efficiency. Consider the case $p = 3$. We have used the reparametrized version of the model as described in Singh and Mukhopadhyay (2016). Let the estimates of the parameters be $\widehat{\boldsymbol{\theta}} = [\widehat{\mu}, \widehat{\beta}_1, \widehat{\beta}_2, \widehat{\tau}, \widehat{\rho}] = [0.5, 1.0, -1.0, -2.0, 1]$ and $\widehat{\alpha} = 0.1$. A compound symmetric correlation structure is assumed and $\Omega = \{AAA, ABB, ABA, AAB, BAA, BAB, BBA, BBB\}$. Optimal design (proportions) is obtained by minimizing the variance function given in equation (6) with respect to ζ. For the given parameter estimates and treatment sequences, LOD is,

$$\zeta^* = \begin{Bmatrix} AAA & ABB & ABA & AAB & BAA & BAB & BBA & BBB \\ 0 & 0.1865 & 0 & 0.1317 & 0.2068 & 0.1105 & 0.3645 & 0 \end{Bmatrix}$$

Observe that the above design uses treatment sequences $\{ABB, AAB, BAA, BAB, BBA\}$ with proportions of subjects $\{0.1865, 0.1317, 0.2068, 0.1105, 0.3645\}$, respectively. The efficient conversion of approximate design to an exact design can be done using the methods

given in Pukelsheim and Rieder (1992). Nearest integer approach is also one of the methods used and works quite well in most of the cases.

Suppose we are interested in comparing ζ^* with another design say

$$\zeta_A = \begin{Bmatrix} ABB & AAB & BAA & BBA \\ 1/4 & 1/4 & 1/4 & 1/4 \end{Bmatrix}$$

Design ζ_A distributes equal proportion of subjects to each treatment sequence considered. The values of $Var_{\zeta^*}(\widehat{\boldsymbol{\tau}}) = 3.7263$ and $Var_{\zeta_A}(\widehat{\boldsymbol{\tau}}) = 3.9778$. The design efficiency calculated using the formula given in equation (8) for $m = 6$ is 0.987.

Remark 1: It is clear that the expression of $Var_{\zeta}(\widehat{\boldsymbol{\theta}})$ given in (5) is a scalar multiple of n and the design efficiency is independent of n, *i.e.*, to compute the efficiency based on the formula (8), total number of observations (n) is not required.

Remark 2: Here and later in this article, design optimization is done using numerical techniques. In particular, we have used *fmincon* function of MATLAB R2014a. The function *fmincon* is used for nonlinear optimization under a constraint. These programs are available from authors upon request.

3.3. Parameter dependence

The variance of the treatment effect estimator depends on the model parameters and the optimal design found by minimizing the variance of the treatment effect is actually locally optimal. We have assumed that historical data from same study are available. Based on the historical data, using GEE approach the point estimate $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is obtained. A parameter space $\mathcal{B}$ for $\boldsymbol{\theta}$ is set up by taking Cartesian product of the confidence intervals of the individual parameters. For each period size p, the efficiencies of the designs listed in Section 3.1 are calculated as follows:

(a) An LOD ζ^* is obtained using $\widehat{\boldsymbol{\theta}}$. Suppose ζ^N denotes a competitive design listed in Section 3.1.

(b) From the parameter space $\mathcal{B}$, 5000 values of $\boldsymbol{\theta}$ are randomly generated. For each value of $\boldsymbol{\theta}$, the efficiency based on ζ^* and ζ^N is computed using (8). Thus, we shall have 5000 efficiencies values corresponding to 5000 values of $\boldsymbol{\theta}$.

(c) The performances of ζ^* and ζ^N are assessed through the box–plot of 5000 efficiencies values calculated in (b).

This allows us to study the robustness of the designs selected to the changes in the parameter values.

3.4. Results

LODs are computed for the following scenarios:

Scenario ID1: LOD for the estimation of the treatment effect (minimize the variance of the estimate of the parameter associated treatment effect) in the model with no carryover effects (NC). The working correlation structure is assumed to be independent (ID).

Scenario ID2: LOD for the estimation of the treatment effect when the carryover effect is included in the model (WC). The working correlation structure is assumed to be ID.

Scenario ID3: LOD for the estimation of the carry effect when the carryover effect is included in the model (WC). The working correlation structure is assumed to be ID.

Scenarios **CS1**, **CS2**, and **CS3** are same as scenarios **ID1**, **ID2**, and **ID3** except that the ID structure is replaced by CS correlation structure. Similarly, scenarios **AR1**, **AR2**, and **AR3** are same as scenarios **CS1**, **CS2**, and **CS3** except that the CS structure is replaced by an auto regressive (AR) correlation structure. In the subsequent sections, optimal proportions are denoted by a vector $\mathbf{p}^* = (p^*_{\omega_1}, \ldots, p^*_{\omega_s})'$, where $\omega_i \in \Omega$ for $i = 1, \ldots, s$, and s is the cardinality of Ω. A LOD ζ^* can be identify by $\mathbf{p}^*$ and the associated treatment sequences.

3.4.1. Two periods two treatment (2×2) crossover trials

For 2×2 binary trial, we used the data from a study reported in Jones and Kenward (2014) (Page 97, Table 2.36). The experiment was designed to assess the cerebrovascular deficiency. Two drugs (placebo and an active drug) given to subjects based on the treatment sequences $\{AB, BA\}$. The responses are recorded as abnormal (0) and normal (1) electrocardiogram readings. Based on the data, the point estimates of the model parameters and the 95% estimated confidence intervals are reported in Table 1. The estimated value of α is 0.1.

Table 1: Estimated 95% confidence intervals of the model parameters in a 2×2 binary crossover trial. The point estimates are the middle points of the associated confidence intervals

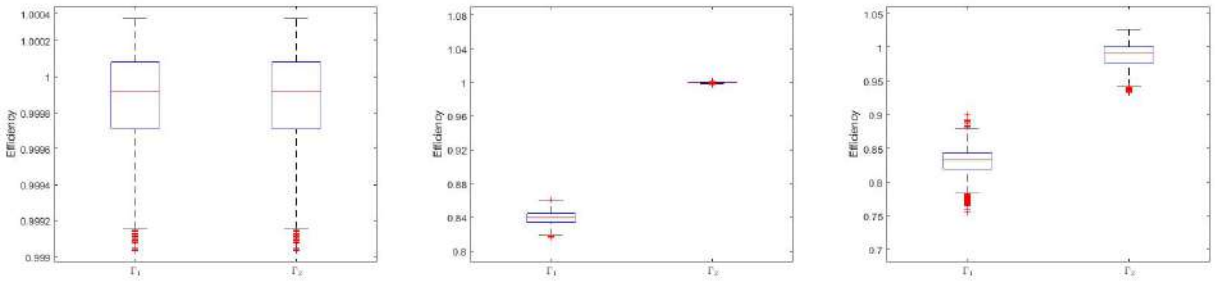
Model	μ	β	τ	ρ
NC	[0.2997, 1.1253]	[-0.5600, 0.2012]	[-0.0572, 0.3238]	·
WC	[0.2976, 1.1364]	[-0.5652, 0.1952]	[-0.1924, 0.6464]	[-0.5441, 0.9141]

LODs for the binary 2×2 trial under scenarios ID1, ID2, ID3, CS1, CS2 and CS3 are computed based the point estimate $\widehat{\boldsymbol{\theta}}$ reported in Table 1. The parameter space to compute the efficiencies is made up the Cartesian product of the interval estimates of the model parameters (given in Table 1). The optimal proportions are reported for the following sequence of treatments: $\{AB, AA, BA, BB\}$.

Scenario ID1: The optimal proportions are obtained as $\mathbf{p}^* = (0.2513, 0.2601, 0.2486, 0.2400)'$. Thus, LOD ζ^* is close to $D_2^{(2)}$. From Figure 1(a), it can be observed that the performances of ζ^*, $D_1^{(2)}$ and $D_2^{(2)}$ are similar. For normal responses, $D_2^{(2)}$ is an optimal design [see Laska and Meisner (1985)].

Scenario ID2: LOD consists of the optimal proportions $\mathbf{p}^* = (0.2436, 0.2633, 0.2514, 0.2418)'$. Design $D_2^{(2)}$ is as efficient as ζ^* whereas $D_1^{(2)}$ performs worst (see Figure 1(b)). Design $D_2^{(2)}$ is an optimal design for normal responses.

Scenario ID3: The vector of optimal proportions is $\mathbf{p}^* = (0.5070, 0, 0, 0.4930)'$. Observing Figure 1(c), it can be concluded that LOD ζ^* is slightly better than $D_2^{(2)}$ since the median efficiency of $D_2^{(2)}$ compared with ζ^* is less than 1. The performance of $D_1^{(2)}$ is worst.



(a) **(b)** **(c)**

Figure 1: Binary 2×2 crossover trials with independent correlation structure. The efficiencies of design $D^{(2)}$ when compared to $D_1^{(2)}$ and $D_2^{(2)}$ are denoted by "Γ_1" and "Γ_2" respectively and given as box–plots. The red line indicates the median and the red dots the outliers. (a) Scenario ID1 (b) Scenario ID2 (c) Scenario ID3

Scenario CS1: Optimal proportions assigned by LOD ζ^* are $\mathbf{p}^* = (0.5011, 0, 0.4989, 0)'$. LOD is very close to $D_1^{(2)}$ which is optimal for normal responses. From the efficiency box–plots (see Figure 2 (a)) it is observed that ζ^* and $D_1^{(2)}$ are equally efficient. The performance of $D_2^{(2)}$ is not satisfactory.

Scenario CS2: LOD ζ^* assigns the proportions $\mathbf{p}^* = (0.2435, 0.2632, 0.2515, 0.2419)'$. LOD is very similar to $D_2^{(2)}$. The performance of ζ^* is similar to $D_2^{(2)}$ whereas $D_1^{(2)}$ perform poorly (see Figure 2 (b)).

Scenario CS3: Optimal proportions are $\mathbf{p}^* = (0.4763, 0.0319, 0, 0.4919)'$. Thus LOD assigns approximately all the proportions to the sequences AB and BB. The efficiency plot (Figure 2 (c)) shows that design $D_2^{(2)}$ is as efficient as ζ^*.

In the above three scenarios (CS1, CS2 and CS3) the correlation parameter α is assumed to take value 0.1. We have repeated the above exercise with $\alpha = 0.4$ and computed the efficiencies. The efficiency plots are depicted in Figure 2 (d), (e) and (f) for scenarios CS1,

CS2 and CS3 respectively. From these plots it is observed that the ranking of designs based on the efficiency remain unchanged for the higher correlation value.

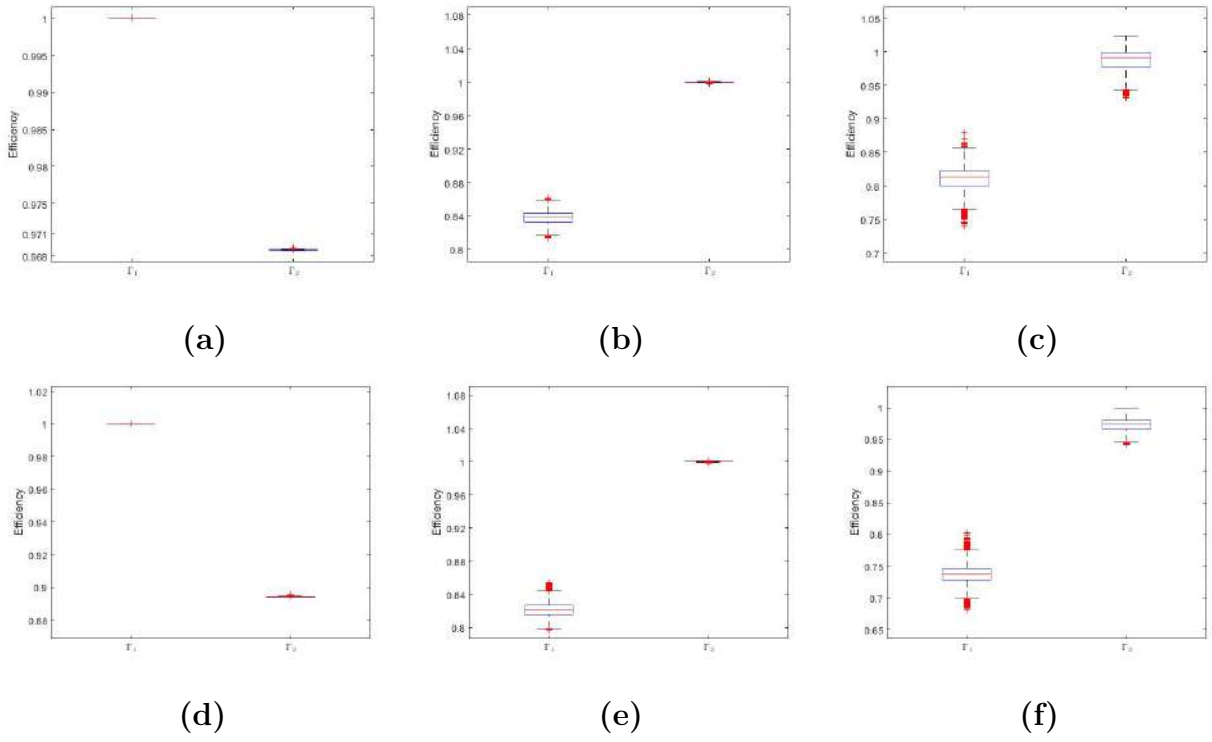


Figure 2: Binary 2×2 crossover trials with CS correlation structure. The efficiencies of design $D^{(2)}$ when compared to $D_1^{(2)}$ and $D_2^{(2)}$ are denoted by "Γ_1" and "Γ_2" respectively and given as boxplots. The red line indicates the median and the red dots the outliers. (a) Scenario CS1 with $\alpha = 0.1$ (b) Scenario CS2 with $\alpha = 0.1$ (c) Scenario CS3 with $\alpha = 0.1$ (d) Scenario CS1 with $\alpha = 0.4$ (e) Scenario CS2 with $\alpha = 0.4$ (f) Scenario CS3 with $\alpha = 0.4$

3.4.2. Three periods two treatment (3×2) crossover trials

The reparameterized version of model (1) for a 3×2 crossover trial is written as

$$logit(\mu_{ij}) = \mu + \beta_1 P_1 + \beta_2 P_2 + \tau \Phi_{d(i,j)} + \rho \Phi_{d(i-1,j)},$$

where where (P_1, P_2) takes values (0,0), (1,0), and (0,1) for the period 1, 2, and 3 respectively.

For the estimation of the confidence intervals we used the data set given in Example 3 of Morrey (1989). The estimated confidence intervals and the point estimates are presented in Table 2. The optimal proportions are reported in the following sequences of treatments: $\{ABB, ABA, AAB, BAA, BAB, BBA, AAA, BBB\}$. LODs are calculated based on the point estimates of the parameters reported in 2. In the computations of efficiency values, the parameter space is made up the Cartesian product of the confidence intervals of the parameters given in Table 2.

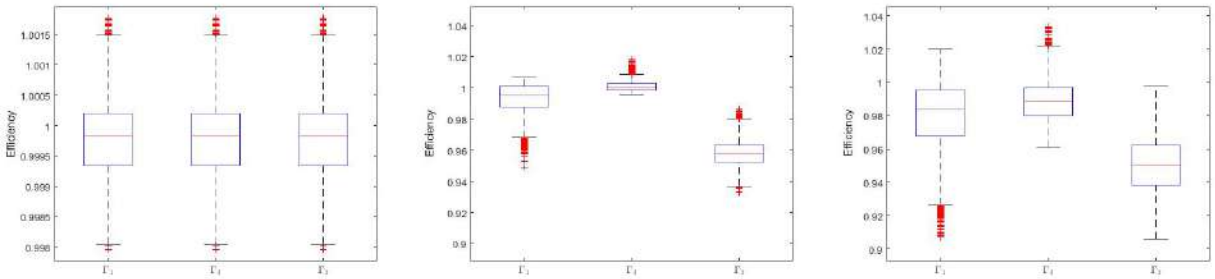
Table 2: Estimated 95% confidence intervals of the model parameters in a 3×2 binary crossover trial. The point estimates are the middle points of the associated confidence intervals

Correlation	Model	ν	β_1	β_2	τ
CS	NC	[-0.8185, 0.4045]	[-0.6396, 1.0616]	[-1.1237, 0.4717]	[0.1021, 0.7959]
	WC	[-0.8210, 0.3978]	[-0.5991, 1.0233]	[-1.3311, 0.4557]	[0.1178, 0.8488]
AR	NC	[-0.8231, 0.4019]	[-0.6334, 1.0620]	[-1.1162, 0.4764]	[0.0722, 0.7814]
	WC	[-0.8216, 0.3976]	[-0.5984, 1.0244]	[-1.3308, 0.4568]	[0.1175, 0.8505]
		ρ			
CS	NC	·			
	WC	[0.0976, 0.9564]			
AR	NC	·			
	WC	[0.0988, 0.9652]			

Scenario ID1: When there is no carryover effect, LOD ζ^* assigns the optimal proportions $\mathbf{p}^* = (0.1288, 0.1154, 0.1289, 0.1203, 0.1356, 0.1202, 0.1155, 0.1354)'$. Thus, ζ^* utilizes all the treatment sequences. Observing the box-plots depicted in Figure 3 (a), it is concluded that ζ^* is as efficient as $D_i^{(3)}$, for $i = 1, 2, 3$.

Scenario ID2: The vector of the optimal proportions is $\mathbf{p}^* = (0.3716, 0, 0.0428, 0.3398, 0.0791, 0, 0.0751, 0.0916)'$. More than 70% observations are assigned the sequence ABB and its dual. In terms of efficiency, design $D_2^{(3)}$ is as efficient as ζ^* followed by $D_1^{(3)}$ (see Figure 3 (b)). Design $D_3^{(3)}$ performs worst.

Scenario ID3: The optimal proportions are $\mathbf{p}^* = (0, 0.2558, 0, 0.2458, 0, 0.2549, 0.2435, 0)'$. In this case, the conclusion is same as in scenario ID2 (see Figure 3 (c)).



(a) **(b)** **(c)**

Figure 3: Binary 3×2 crossover trials with independent correlation structure. The efficiencies of design $\zeta^* = D^{(3)}$ when compared to $D_1^{(3)}$, $D_2^{(3)}$ and $D_3^{(3)}$ are denoted by "Γ_1", "Γ_2" and "Γ_3" respectively and given as boxplots. The red line indicates the median and the red dots the outliers. (a) Scenario ID1 (b) Scenario ID2 (c) Scenario ID3

Scenario CS1: LOD ζ^* chooses the treatment sequences with weights $\mathbf{p}^* = (0.0319, 0.4553, 0, 0, 0.4980, 0.0145, 0, 0)'$. Thus, more than 95% of subjects are assigned to the sequence ABA and its dual BAB. In terms of efficiency, ζ^* is as efficient as $D_3^{(3)}$ (less wider spread of the box–plot) followed by $D_1^{(3)}$ and $D_2^{(3)}$ (see Figure 4 (a)).

Scenario CS2: The optimal proportions are $\mathbf{p}^* = (0.3344, 0, 0.1891, 0.3588, 0, 0.1177, 0, 0)'$. Thus, LOD is positively supported only on the treatment sequences ABB, ABA and their duals BAA and BAB with more than 68% proportion only to ABB and its dual. For normal responses, optimal design equally assigns subjects to ABB and its dual. From Figure 4 (b), it is observed that $D_2^{(3)}$ is best in terms of the efficiency closely followed by $D_1^{(3)}$. Design $D_3^{(3)}$ performs worst.

Scenario CS3: LOD ζ^* is consist of the optimal proportions $\mathbf{p}^* = (0.4219, 0.0845, 0, 0.4441, 0, 0.0223, 0, 0.0272)'$. Thus ζ^* assigns more than 85% subjects to the sequence ABB and its dual. The efficiency plot (Figure 4 (c)) shows that only $D_1^{(3)}$ has satisfactory performance.

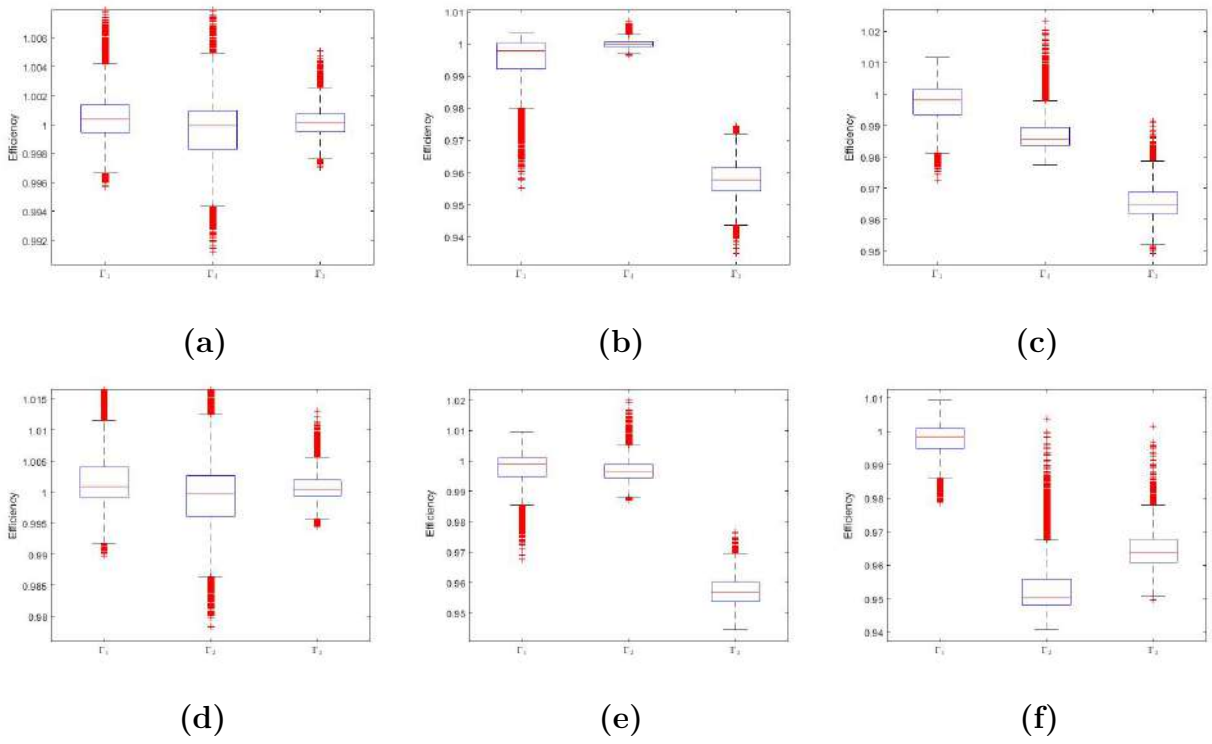


Figure 4: Binary 3×2 crossover trials with CS correlation structure. The efficiencies of design $\zeta^* = D^{(3)}$ when compared to $D_1^{(3)}$, $D_2^{(3)}$ and $D_3^{(3)}$ are denoted by "Γ_1", "Γ_2" and "Γ_3" respectively and given as boxplots. The red line indicates the median and the red dots the outliers. (a) Scenario CS1 with $\alpha = 0.1$ (b) Scenario CS2 with $\alpha = 0.1$ (c) Scenario CS3 with $\alpha = 0.1$ (d) Scenario CS1 with $\alpha = 0.4$ (e) Scenario CS2 with $\alpha = 0.4$ (f) Scenario CS3 with $\alpha = 0.4$

The above three scenarios (CS1, CS2 and CS3) are done based on the correlation

parameter $\alpha = 0.1$. We have repeated the above exercise with $\alpha = 0.4$ and computed the efficiencies. The efficiency plots are depicted in Figure 4 (d), (e) and (f) for scenarios CS1, CS2 and CS3 respectively. From these plots it is observed that the ranking of designs based on the efficiency remain unchanged for the higher correlation expect in scenario CS3. In Scenario CS3 an increase in α worsen the the performance of $D_2^{(3)}$.

Scenario AR1: In this case LOD utilizes only two treatment sequences ABA and BAB with approximately equal proportion of subjects. A design with equal proportions of subjects in ABA and BAB is optimal for normal responses. The efficiencies of all designs $D_i^{(3)}$, $i = 1, 2, 3$ compared with LOD are less than 1 (see Figure 5 (a)). Thus, the performances of any of $D_i^{(3)}$'s are not satisfactory.

Scenario AR2: Optimal proportions are obtained as $\mathbf{p}^* = (0.3922, 0, 0.1012, 0.4153, 0.0229, 0, 0, 0.0684)'$. LOD ζ^* assigns more than 80% subjects to the sequence ABB and its dual. Design $D_2^{(3)}$ is comparably as efficient as ζ^* followed by $D_1^{(3)}$ (see Figure 5 (b)). The performance of design $D_3^{(3)}$ is not satisfactory.

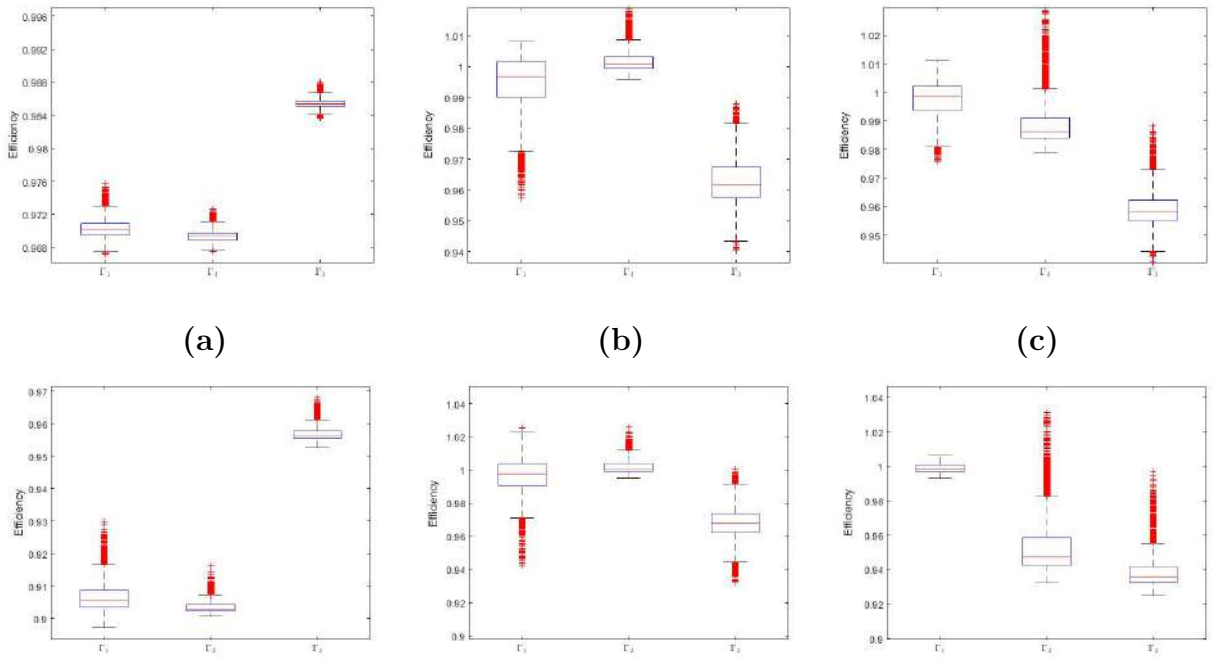


Figure 5: Binary 3×2 crossover trials with AR correlation structure. The efficiencies of design $\zeta^* = D^{(3)}$ when compared to $D_1^{(3)}$, $D_2^{(3)}$ and $D_3^{(3)}$ are denoted by "Γ_1", "Γ_2" and "Γ_3" respectively and given as boxplots. The red line indicates the median and the red dots the outliers. (a) Scenario AR1 with $\alpha = 0.1$ (b) Scenario AR2 with $\alpha = 0.1$ (c) Scenario AR3 with $\alpha = 0.1$ (d) Scenario AR1 with $\alpha = 0.4$ (e) Scenario AR2 with $\alpha = 0.4$ (f) Scenario AR3 with $\alpha = 0.4$

Scenario AR3: LOD ζ^* is consist of the optimal proportions $\mathbf{p}^* = (0.4663, 0.0370, 0, 0.4349, 0,$

$0.0217, 0, 0.0401)'$. LOD assigns more than 90% subjects to the sequence ABB and its dual BAA. Efficiency plots (see Figure 5 (c)) shows that only $D_1^{(3)}$ is as efficient as ζ^*.

When $\alpha = 0.4$, the efficiency plots are depicted in Figure 5 (d), (e) and (f) for scenarios AR1, AR2 and AR3 respectively. It is observed that the efficiency wise ranking of designs remain unchanged.

3.4.3. Four periods two treatment (4×2) crossover trials

We analyse the two treatment double blinded crossover data reported in McKnight and Van Den Eeden (1993). The study was designed to examine whether aspartame causes headaches in subjects who believe they experience aspartame-induced headaches. The run-in period was 7 days followed by a wash-out day repeated for 4 periods. Three doses per day of Both aspartame (A) 30 $mg/kg/day$, and placebo (B) were given to the subjects. There were four possible ordering of the treatments ($ABAB$, $ABBA$, $BABA$, $BAAB$). The response y takes values 0 if the number of days with headache is less than 2 and is equal to 1 if the number of days with headache is greater than or equal to 2.After removing the dropouts, the data for 21 subjects is given in Table 3.

The reparameterized version of model (1) for a 4×2 crossover trial is written as

$$logit(\mu_{ij}) = \mu + \beta_1 P_1 + \beta_2 P_2 + \beta_3 P_3 + \tau \Phi_{d(i,j)} + \rho \Phi_{d(i-1,j)},$$

where where (P_1, P_2, P_3) takes values (0,0,0), (1,0,0), (0,1,0) and (0,0,1) for the period 1, 2, 3 and 4 respectively.

The estimated confidence intervals and the point estimates based on the data from Table 3 are presented in Table 4. LODs are calculated based on the point estimates of the parameters reported in 4. Optimal proportions ($\mathbf{p}^*$) for all scenarios are reported in Table 5. In the computation of the efficiency, the parameter space is made of the Cartesian products of the confidence intervals of the parameters given in Table 4.

Table 3: Treatment order and corresponding Response of each period

Treatment order	period 1	period 2	period 3	period 4
$ABAB$	0	1	0	0
$ABAB$	1	1	1	1
$ABAB$	0	0	0	0
$ABAB$	1	1	1	1
$ABAB$	1	0	1	0
$ABBA$	0	1	0	0
$ABBA$	1	1	1	1
$ABBA$	0	0	1	0
$ABBA$	0	0	0	0
$ABBA$	1	1	1	1
$ABBA$	0	0	0	0
$BABA$	0	1	1	1
$BABA$	0	0	0	0
$BABA$	1	0	0	0
$BABA$	1	0	0	1
$BABA$	0	0	0	1
$BABA$	1	1	0	1
$BAAB$	0	0	0	0
$BAAB$	0	0	0	0
$BAAB$	0	1	0	0
$BAAB$	1	1	0	0

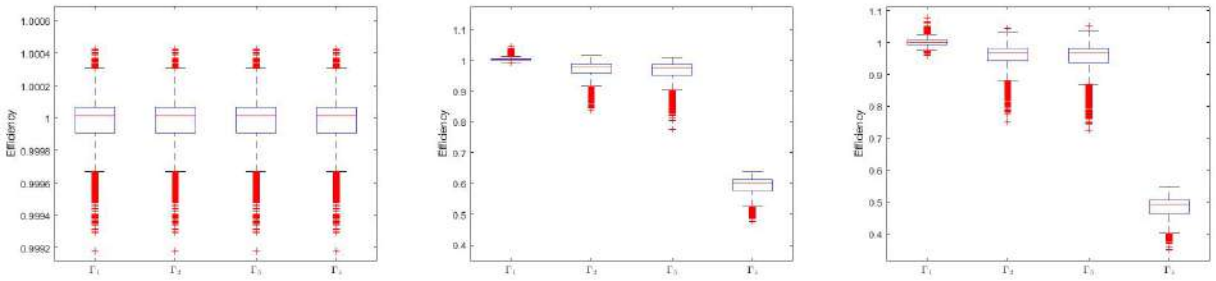
Table 4: Estimated 95% confidence intervals of the model parameters in a 4×2 binary crossover trial. The point estimates are the middle points of the associated confidence intervals

Correlation	Model	ν	β_1	β_2	β_3
CS	NC	[-1.160, 0.570]	[-0.683, 1.453]	[-1.381 , 0.615]	[-1.050, 0.640]
	WC	[-1.129, 0.623]	[-0.684, 1.492]	[-1.518, 0.466]	[-1.133, 0.603]
AR	NC	[-1.161, 0.563]	[-0.685, 1.455]	[-1.374, 0.618]	[-1.057, 0.641]
	WC	[-1.141, 0.615]	[-0.671, 1.493]	[-1.498, 0.462]	[-1.117, 0.615]
		τ	ρ		
CS	NC	[-0.227, 0.447]	·		
	WC	[-0.566, 0.280]	[-1.012, 0.045]		
AR	NC	[-0.242, 0.498]	·		
	WC	[-0.638, 0.318]	[-1.047, 0.109]		

Scenario ID1: In this case LOD ζ^* utilizes all the treatment sequences except $AAAA$ and its dual (see Table 5). Observing Figure 6 (a) it is concluded that all designs are equally efficient.

Scenario ID2: LOD ζ^* assigns more than 80% observations to the sequences $AABB$, $AABA$ and their duals. Design $D_1^{(4)}$ is as efficient as ζ^* whereas $D_4^{(4)}$ perform worst (see Figure 6 (b)).

Scenario ID3: More than 85% observations are assigned to the sequences $BAAB$ and $ABBB$. Design $D_1^{(4)}$ is as efficient as LOD (see Figure 6 (c)). Performance of $D_4^{(4)}$ is worst.



(a) **(b)** **(c)**

Figure 6: Binary 4×2 crossover trials with independent correlation structure. The efficiencies of design $\zeta^* = D^{(4)}$ when compared to $D_1^{(4)}$, $D_2^{(4)}$, $D_3^{(4)}$ and $D_4^{(4)}$ are denoted by "Γ_1", "Γ_2", "Γ_3" and "Γ_4" respectively and given as boxplots. The red line indicates the median and the red dots the outliers. (a) Scenario ID1 (b) Scenario ID2 (c) Scenario ID3

Scenario CS1: LOD ζ^* equally assigns more than 94% observations to the sequence $ABBA$ and its dual. When $\alpha = 0.4$, then also, LOD utilizes $ABBA$ and its dual with more 96% observations assigned to them. Design $D_3^{(4)}$ is as efficient as ζ^* (see Figures 7 (a) and (d)).

Scenario CS2: In this case LOD assigns more than 90% observations to the sequences $AABB$, $ABAB$ and their duals. LOD is not affected by change in α from 0.1 to 0.4. Design $D_1^{(4)}$ is most efficient when compared to others with respect to LOD (see Figure 7 (b) and (e)). Note that $D_1^{(4)}$ is an optimal design for normal responses.

Scenario CS3: LOD ζ^* utilizes only the following sequences $\{AABB, BAAB, ABBB, BAAA\}$. SImilar to scenario CS2, in this case $D_1^{(4)}$ is as efficient as ζ^* (see Figure 7 (c) and (f)).

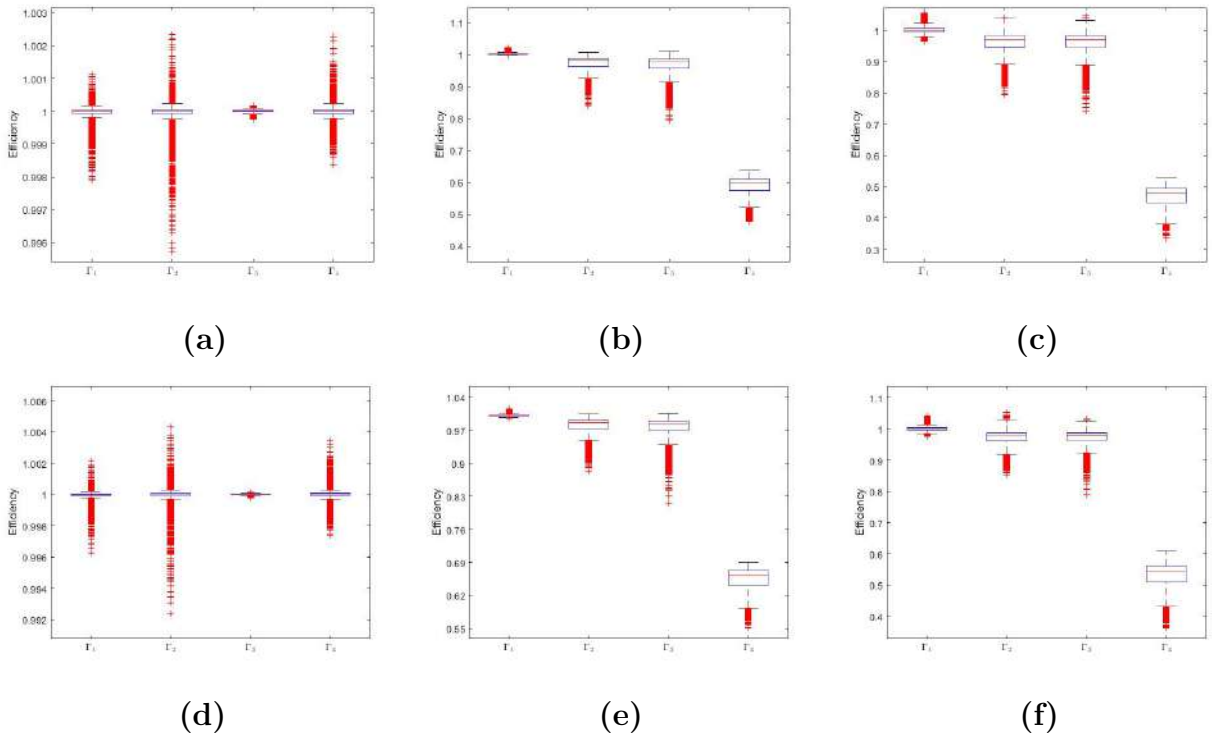


Figure 7: Binary 4×2 crossover trials with CS structure. The efficiencies of design $\zeta^* = D^{(4)}$ when compared to $D_1^{(4)}$, $D_2^{(4)}$, $D_3^{(4)}$ and $D_4^{(4)}$ are denoted by "Γ_1", "Γ_2", "Γ_3" and "Γ_4" respectively and given as boxplots. The red line indicates the median and the red dots the outliers. (a) Scenario CS1 with $\alpha = 0.1$ (b) Scenario CS2 with $\alpha = 0.1$ (c) Scenario CS3 with $\alpha = 0.1$ (d) Scenario CS1 with $\alpha = 0.4$ (e) Scenario CS2 with $\alpha = 0.4$ (f) Scenario CS3 with $\alpha = 0.4$

Scenario AR1: LOD assigns approximately equal observations only to the sequence $ABAB$ and its dual. LOD is as efficient as $D_4^{(4)}$ (see Figure 8 (a) and (d)) which is an optimal design for normal responses.

Scenario AR2: When $\alpha = 0.1$, LOD assigns more than 80% observations to the sequences $AABB$, $BABB$ and their duals. However, when $\alpha = 0.4$, ζ^* utilizes $AABB$, $ABBA$ and

their duals with approximately all the observations assigned to them. Design $D_1^{(4)}$ is as efficient as ζ^* (see Figure 8 (b) and (e)).

Scenario AR3: Approximately all the observations assigned to $AABB$, $BABB$ and their duals. Design $D_1^{(4)}$ is as efficient as ζ^* (see Figure 8 (c) and (f)).

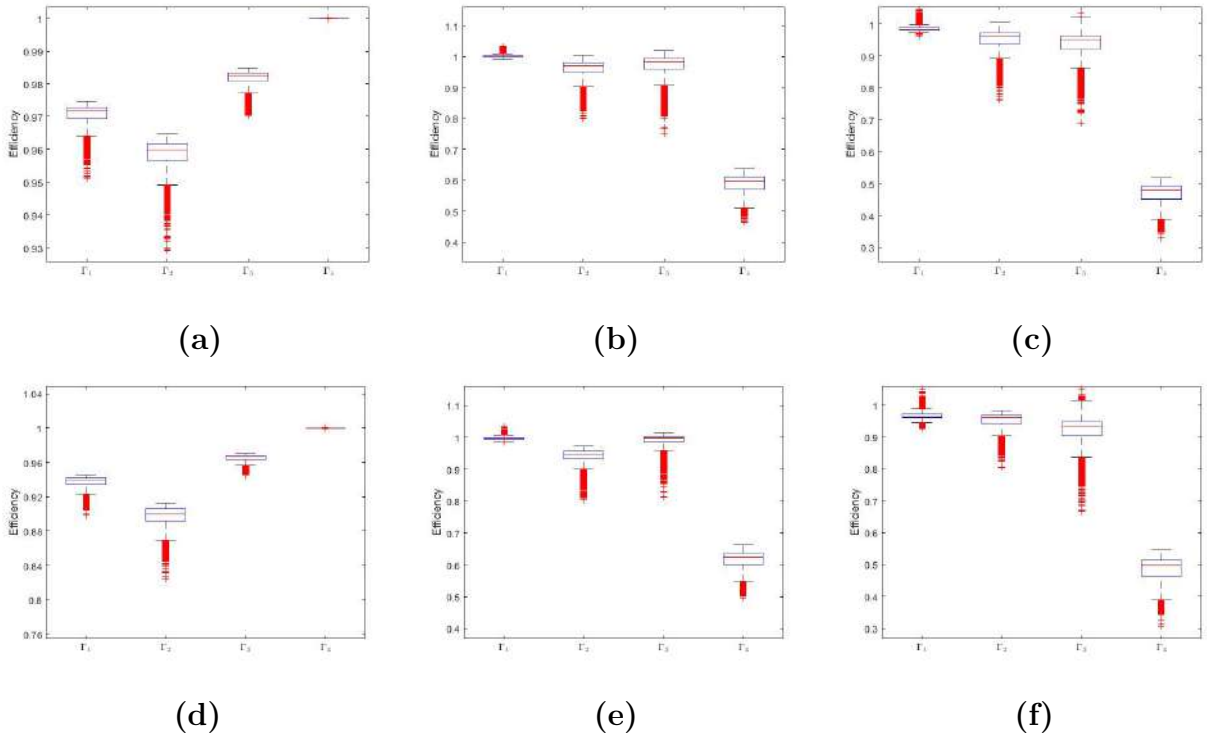


Figure 8: Binary 4×2 crossover trials with AR structure. The efficiencies of design $\zeta^* = D^{(4)}$ when compared to $D_1^{(4)}$, $D_2^{(4)}$, $D_3^{(4)}$ and $D_4^{(4)}$ are denoted by "Γ_1", "Γ_2", "Γ_3" and "Γ_4" respectively and given as boxplots. The red line indicates the median and the red dots the outliers. (a) Scenario AR1 with $\alpha = 0.1$ (b) Scenario AR2 with $\alpha = 0.1$ (c) Scenario AR3 with $\alpha = 0.1$ (d) Scenario AR1 with $\alpha = 0.4$ (e) Scenario AR2 with $\alpha = 0.4$ (f) Scenario AR3 with $\alpha = 0.4$

Table 5: Optimal proportion ($\mathbf{p}^*$) for 4×2 crossover trials

				$\alpha = 0.1$					
Sequence	ID1	ID2	ID3	CS1	CS2	CS3	AR1	AR2	AR3
$AABB$	0.0730	0.2251	0.0219	0.0081	0.3530	0.0614	0	0.2096	0.2015
$BBAA$	0.0698	0.1363	0	0	0.3415	0	0	0.2427	0.2007
$ABBA$	0.0710	0.0152	0	0.4771	0.0884	0	0	0.1235	0
$BAAB$	0.0717	0	0.4503	0.4786	0	0.4338	0	0	0
$ABAB$	0.0704	0.0168	0	0.0153	0.0930	0	0.4977	0	0
$BABA$	0.0723	0	0	0.0189	0.1240	0	0.5023	0.0021	0
$ABBB$	0.0739	0.0153	0.4153	0	0	0.4302	0	0	0
$BABB$	0.0754	0	0.0310	0	0	0	0	0.2427	0.2772
$BBAB$	0.0726	0.2490	0	0	0	0	0	0	0
$BBBA$	0.0732	0.1027	0.0139	0	0	0	0	0	0
$BAAA$	0.0690	0	0	0	0	0.0744	0	0	0.0290
$ABAA$	0.0679	0.0158	0	0	0	0	0	0.1771	0.2917
$AABA$	0.0702	0.2235	0	0	0	0	0	0.0021	0
$AAAB$	0.0696	0	0.0670	0	0	0	0	0	0
$AAAA$	0	0	0	0	0	0	0	0	0
$BBBB$	0	0	0	0	0	0	0	0	0

	$\alpha = 0.4$					
	CS1	CS2	CS3	AR1	AR2	AR3
$AABB$	0.0072	0.3538	0.0592	0	0.1253	0.1999
$BBAA$	0.0012	0.3423	0	0	0.1119	0.2240
$ABBA$	0.4842	0.0889	0	0	0.3732	0
$BAAB$	0.4857	0	0.3696	0	0.3644	0
$ABAB$	0.0091	0.0918	0	0.4973	0	0
$BABA$	0.0127	0.1232	0	0.5027	0	0
$ABBB$	0	0	0.4311	0	0	0
$BABB$	0	0	0	0	0	0.2838
$BBAB$	0	0	0	0	0.0251	0
$BBBA$	0	0	0	0	0	0
$BAAA$	0	0	0.1400	0	0	0
$ABAA$	0	0	0	0	0	0.2923
$AABA$	0	0	0	0	0	0
$AAAB$	0	0	0	0	0	0
$AAAA$	0	0	0	0	0	0
$BBBB$	0	0	0	0	0	0

3.5. An equivalence theorem

For the linear regression models equivalence theorems were developed by Whittle (1973) and Kiefer (1974). For the GLMs, equivalence theorem for the Bayesian setup was discussed in Chaloner and Larntz (1989). Optimality of min–max crossover designs for the binary response model was verified by an equivalence theorem discussed in Singh *et al.* (2020). Singh and Mukhopadhyay (2016) provided an equivalence theorem to confirm the optimality of numerically obtained crossover designs for the GLMs. In this section we provide an equivalence theorem which can be utilized to verify the optimality of the LODs obtained in this article.

Let the design space be defined as a unit simplex $\boldsymbol{\Xi} = \{\mathbf{p}' = (p_{\omega_1}, \ldots, p_{\omega_s}) : \sum_{i=1}^{s} p_{\omega_i} = 1, \text{ and } 0 \leq p_{\omega_i} \leq 1\}$. Note that $Var_{\zeta}(\widehat{\boldsymbol{\theta}})$ given in (5) depends on $\mathbf{p}$ via ζ. Therefore, $Var_{\zeta}(\widehat{\boldsymbol{\theta}})$ can be represented as $Var_{\zeta(\mathbf{p})}(\widehat{\boldsymbol{\tau}})$. Suppose the interest is in estimating a estimable linear function of the parameters say $\lambda = L'\boldsymbol{\theta}$, where L is a $m \times s$ matrix, m is the length of the vector $\boldsymbol{\theta}$, and $s \leq m$. The information matrix of λ is given by $C = (L'Var_{\zeta(\mathbf{p})}(\widehat{\boldsymbol{\theta}})L)^{-1}$.

Theorem 1: A locally optimal design $\zeta^* \equiv \zeta(\mathbf{p}^*)$ at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ obtained as

$$\zeta^* = \min_{\mathbf{p} \in \boldsymbol{\Xi}} \log(Det(L'Var_{\zeta(\boldsymbol{p})}(\widehat{\boldsymbol{\theta}})L))$$

satisfies the condition

$$trace\{Var_{\zeta^*}(\widehat{\boldsymbol{\theta}})LCL'Var_{\zeta^*}(\widehat{\boldsymbol{\theta}})(Var_{\zeta_\omega}(\widehat{\boldsymbol{\theta}}))^{-1}\} \leq s, \quad \text{for all } \omega \in \Omega, \tag{9}$$

where $Var_{\zeta_\omega}(\widehat{\boldsymbol{\theta}})$ is the variance with respect to the design ζ_ω having unit mass at single treatment sequence ω. Equality holds in (9) if the treatment sequence ω is included in ζ^* with positive probability.

The proof of Theorem 1 follows from Theorem 1 of Pettersson, H. (2005) and Theorem 2.1 of Müller, C. H. and Pázman, A. (1998).

4. Summary

Crossover designs for two treatments and binary responses are determined for $p = 2, 3, 4$. Since these designs depend on the model parameters, various intervals estimated from the data sets based on the historical studies are considered and LODs are found in each case. Within subject correlation is modelled using working correlation matrix assuming: independent, compound symmetric and auto–regressive structures. LODs are compared with designs optimal for normal responses in each case.

In Table 6 we list designs optimal for normal responses which are as efficient as locally optimal designs obtained in this article under various scenarios.

Table 6: Efficient designs

Periods	ID1	ID2	ID3	CS1	CS2	CS3	AR1	AR2	AR3
2	$\{D_i^{(2)} : i = 1, 2\}$	$D_2^{(2)}$	$D_2^{(2)}$	$D_1^{(2)}$	$D_2^{(2)}$	$D_2^{(2)}$			
3	$\{D_i^{(3)} : i = 1, 2, 3\}$	$D_2^{(3)}$	$D_2^{(3)}$	$\{D_i^{(3)} : i = 1, 2\}$	$D_2^{(3)}$	$D_1^{(3)}$	$D_3^{(3)}$	$D_2^{(3)}$	$D_1^{(3)}$
4	$\{D_i^{(4)} : i = 1, \ldots, 4\}$	$D_1^{(4)}$	$D_1^{(4)}$	$\{D_i^{(4)} : i = 1, 2, 3, 4\}$	$D_1^{(4)}$	$D_1^{(4)}$	$D_4^{(4)}$	$D_4^{(4)}$	$D_1^{(4)}$

In conclusion it clear from the numerical studies that the results in the logistic regression case are quite similar to the available results in the continuous case in most of the scenarios.

Acknowledgements

We would like to acknowledge Late Professor Aloke Dey for his insightful guidance and support in formulating the optimal crossover design problem and developing the methodology. This work could not have been completed without the many fruitful discussions we had with him.

The work of Siuli Mukhopadhyay is supported by the DST grant EMR/2016/005142 and gratefully acknowledged. The work of Satya Prakash Singh is supported by the SEED grant Indian Institute of Technology Hyderabad, India and gratefully acknowledged.

References

Bose, M. and Dey, A. (2009). *Optimal Crossover Designs.* World Scientific, Singapore.

Bose, M. and Dey, A. (2015). Crossover Designs. In *Handbook of Design and Analysis of Experiments* (A. M. Dean, M. Morris, J. Stufken, D. Bingham, Eds.), Chapman and Hall/CRC Press, London, pp. 159–195.

Carriere, K. C. (1994). Crossover designs for clinical trials. *Statistics in Medicine* **13**, 1063–1069.

Carriere, K. C. and Huang, R. (2000). Crossover designs for two-treatment clinical trials. *Journal of Statistical Planning and Inference*, **87**, 125–134.

Carriere, K. C. and Reinsel, G. (1992). Investigation of dual-balanced crossover designs for two treatments. *Biometrics*, **48**, 1157–1164.

Chaloner, K. and Larntz, K. (1989). Optimal Bayesian design applied to logistic regression experiments. *Journal of Statistical Planning and Inference*, **21**, 191-208.

Gondaliya, J. and Divecha, J. (2015). Estimation of treatment and carryover effects in optimal cross-over designs for clinical trials. *Statistics in Biopharmaceutical Research*, **7**, 95-105.

Gondaliya, J. and Divecha, J. (2018). Optimal and/or efficient two treatment crossover designs for five carryover models. *The International Journal of Biostatistics*, **24**, 1-11.

Grizzle, J.E., (1965). The two-period changeover design and its use in clinical trials. *Biometrics*, **21**, 467-480.

Ebbutt, A. F. (1984). Three-period crossover designs for two treatments. *Biometrics*, **40**, 219–224.

Jankar, J., Mandal, A. and Yang, J. (2020).Optimal crossover designs for generalized linear models. *Journal of Statistical Theory and Practice*, **14**, https://doi.org/10.1007/s42519-020-00089-5.

Jones, B. and Kenward, M. G. (2014). *Design and Analysis of Crossover Trials*, 3rd ed. CRC Press, London.

Kershner, R. P. and Federer, W. T. (1981). Two-treatment crossover designs for estimating a variety of effects. *Journal of the American Statistical Association*, **76**, 612–619.

Kershner, R. P. (1986). Optimal 3-period 2-treatment crossover designs with and without baseline measurements. *Proceedings of the Biopharmaceutical Section of the American Statistical Association*, 152–156.

Kiefer, J. C. (1971). *The Role of Symmetry and Approximation in Exact Design Optimality.* In S. S. Gupta and J. Yackel (ed.), Statistical decision theory and related topics, Proceedings of a Symposium, Purdue University, New York, pages 109–118.

Kiefer, J. (1974). General equivalence theory for optimum designs (approximate theory). *The Annals of Statistics*, **2**, 849-879.

Kushner, H. B. (1997). Optimal repeated measurements designs: the linear optimality equations. *Annals of Statistics*, **25**, 2328–2344.

Kushner, H. B. (1998). Optimal and efficient repeated measurements designs for uncorrelated observations. *Journal of the American Statistical Association*, **93**, 1176–1187.

Laska, E. M. and Meisner, M. (1985). A variational approach to optimal two-treatment crossover designs: application to carryover-effect models. *Journal of the American Statistical Association*, **80**, 704–710.

Laska, E. M., Meisner, M. and Kushner, H. B. (1983). Optimal crossover designs in the presence of carryover effects. *Biometrics*, **39**, 1087–1091.

Liang, K. Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, **73**, 13–22.

Matthews, J. N. S. (1987). Optimal crossover designs for the comparison of two treatments in the presence of carryover effects and autocorrelated errors. *Biometrika*, **74**, 311–320.

McKnight, B., and Van Den Eeden, S.K. (1993). A conditional analysis for two-treatment multiple period crossover designs with binomial or Poisson outcomes and subjects who drop out. *Statistics in Medicine*, **12**, 825-834.

Morrey, G. H. (1989). Binary responses and the three period crossover. *Biometrical Journal*, **31**, 589–598.

Müller, C. H. and Pázman, A. (1998). Applications of necessary and sufficient conditions for maximin efficient designs. *Metrika*, **48**, 1-19.

Pettersson, H. (2005). Optimal design in average for inference in generalized linear models. *Statistical Papers*, **46**, 79-100.

Pukelsheim, F. and Rieder, S. (1992). Efficient rounding of approximate designs. *Biometrika*, **79**, 763–770.

Robinson, K. S. and Khuri, A. I. (2003). Quantile dispersion graphs for evaluating and comparing designs for logistic regression models. *Computational Statistics and Data Analysis*, **43**, 47–62.

Senn, S. (2003). *Cross-over Trials in Clinical Research*, 2nd ed. Wiley, Chichester, England.

Singh, S. P. and Mukhopadhyay, S. (2016). Bayesian Crossover Designs for Generalized Linear Models. *Computational Statistics and Data Analysis*, **104**, 35–50.

Singh, S. P., Mukhopadhyay, S. and Raj, H. (2020). Min-max crossover designs for generalized linear models. *Submitted.*

Waterhouse, T. H., Eccleston, J. A. and Dufull, S. B. (2006). Optimal crossover designs for logistic regression models in pharmacodynamics. *Journal of Biopharmaceutical Statistics*, **16**, 881–894.

Whittle, P. (1973). Some general points in the theory of optimal experimental design. *Journal of the Royal Statistical Society, B*, **35**, 123-130.

Zeger, S. L., Liang, K. Y. and Albert, P. S. (1988). Models for longitudinal data: a generalized estimating equation approach. *Biometrics*, **44**, 1049–1060.

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 247–255

Multiway Blocking of Designs of Experiments

Nam-Ky Nguyen[1], **Tung-Dinh Pham**[2], **Mai Phuong Vuong**[3]
[1] *Vietnam Institute for Advanced Study in Mathematics, Hanoi, Vietnam*
[2] *VNU University of Science, Vietnam National University, Hanoi*
[3] *Hanoi University of Science & Technology, Vietnam*

Received: 07 January 2021; Revised: 05 February 2021; Accepted: 09 February 2021

Abstract

Fisher's three R's or three principles of designs of experiments are (i) Randomisation; (ii) Replications; and (iii) Local control or blocking (also called noise reduction). Of the three, blocking is the most difficult. Works on blocked 3-level designs are very limited. In addition, there might be more than one extraneous variations or blocking factors. As such, there is a need for a general method to do multiway blocking of experimental designs. This paper extends the idea of orthogonal blocking of Box and Hunter (1957) from one blocking factor to several blocking factors. It then presents a blocking algorithm which can impose several blocking/noise factors on popular experimental designs. Particular attention will be given to 2-level, 3-level and mixed-level screening designs such as those introduced by Jones and Nachtsheim (2013) and Nguyen *et al.* (2020).

Key words: Box-Behnken designs; Central composite designs; Definitive screening designs; Fractional factorial designs; Orthogonal blocking.

1. Introduction

An experiment is conducted to determine the relationship between input factors affecting a process and the output of that process. There are controllable input factors to be studied as well as nuisance (uncontrollable) ones to be eliminated. While the former can be modified to optimise the output, this is not the case of the latter. Examples of uncontrollable factors are: (i) different batches of raw material; (ii) different machine; (iii) different operators; (iv) different locations; (v) different times, *etc.* A well-designed experiment minimises the effects of these uncontrollable factors by partitioning the set of experimental runs into more homogeneous subsets. This noise reduction exercise is called local control or blocking. It makes experiments more sensitive in detecting significant effects and hence less experimentation may be required. Examples of the scenarios when designs of more than one blocking factors can be used are given below.

Example 1: A 2^5 factorial experiment to identify interaction effects for different additives in linear low-density polyethylene film (Hoang *et al.*, 2004, Mee, 2009, p. 79). The factors and levels are: (**A**) Antioxidant A (ppm), 0 and 400; (**B**) Antioxidant B (ppm), 0 and 400; (**C**) Acid Scavenger (ppm), 0 and 1000; (**D**) Anti-block agent (ppm), 0 and 2000; (**E**) Slip

Corresponding Author: Nam-Ky Nguyen
Email: nknam@viasm.edu.vn

additive (ppm), 0 and 800. As the full 2^5 factorial would take at least three days to complete, it was divided into four blocks. Let's assume the experimenter wishes to add to the model an additional block factor, *i.e.* times of the day (8AM and 2PM) and find a suitable design which can accommodate the new blocking factor.

Example 2: A 9-factor DSD in 21 runs was conducted to investigate the oxidation reactions in homogeneous Co^{2+}/PMS system (Zhang *et al.*, 2018). The objective is to evaluate the suitability of the DSD approach in optimising the operating parameters of Co^{2+}/PMS system and to identify the significant effects involved in the reaction system. See Jones and Nachtsheim (2011) for the use of DSD as a screening design. The nine factors in this experiment are: (**1**) NaCl, (**2**) NaH_2PO_4, (**3**) $NaHCO_3$, (**4**) $NaNO_3$, (**5**) Na_2SO_4, (**6**) HA, (**7**) PMS, (**8**) AO II, and (**9**) Co^{2+}. The first five factors were set at 0, 10, and 20mM, HA at 0, 20, 40mg dm^{-3}, PMS at 2, 6, 10mM. AO II at 50, 75, 100 mg dm^{-3} and Co^{2+} at 0, 0.68, and 1.36mM. Zhang *et al.* (2018) stated that they could not use response surface designs (RSDs), such as the Box-Behnken designs (BBD), the Doehlert design and the central composite design, as they could not afford the enormous number of runs required by these designs. At the same time, the popular Plackett-Burman design is unable to capture the quadratic and interaction effects. Let's assume the experiment was performed in two different reactors and three different days and the experimenters wish to add these two blocking factors to the model.

2. Conditions for Orthogonal Blocks

Consider the following model for an n-run design with m factors $x_1, \ldots, x_m$, out of which m_3 factors are at 3-level and the rest are at 2-level, arranged in b blocks:

$$\mathbf{y} = \mathbf{Z}\boldsymbol{\delta} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{1}$$

where $\mathbf{y}$ is an $n \times 1$ column vector of response, $\mathbf{Z}$ an $n \times b$ matrix containing b dummy variables, $\boldsymbol{\delta}$ a $b \times 1$ column vector representing block effects, $\mathbf{X}$ is an $n \times p$ expanded design matrix, $\boldsymbol{\beta}$ a $p \times 1$ column vector of parameters to be estimated, and $\boldsymbol{\epsilon}$ an $n \times 1$ column vector of random errors assumed to have zero mean and variance σ^2. $\mathbf{X}$ includes a column of 1's, representing the intercept; m columns representing the main-effects (MEs); and depending on the model might also include $\binom{m}{2}$ columns representing the 2-factor interactions (2FIs) and m_3 columns representing the second-order effects (SOEs).

Following Box and Hunter 1957, Section 8 and Khuri and Cornell, 1996, Chapter 8, we scale the columns of $\mathbf{Z}$ by subtracting the value of each column from the column mean. Equation (1) can then be written as:

$$\mathbf{y} = \tilde{\mathbf{Z}}\boldsymbol{\delta} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{2}$$

The condition for orthogonal blocks can be written as:

$$\tilde{\mathbf{Z}}'\mathbf{X} = \mathbf{0}. \tag{3}$$

As the row sum of $\tilde{\mathbf{Z}}$ in (2) is zero, which is an example of perfect multicollinearity, to avoid the singular data matrix, we drop the last column of $\tilde{\mathbf{Z}}$ and the last element of $\boldsymbol{\delta}$ in

(2). The least square solution for the unknown parameters $\boldsymbol{\delta}$ and $\boldsymbol{\beta}$ in (2) is the solution of the following equation:

$$\begin{pmatrix} \tilde{\mathbf{Z}}' \\ \mathbf{X}' \end{pmatrix} \mathbf{y} = \begin{pmatrix} \tilde{\mathbf{Z}}'\tilde{\mathbf{Z}} & \tilde{\mathbf{Z}}'\mathbf{X} \\ \mathbf{X}'\tilde{\mathbf{Z}} & \mathbf{X}'\mathbf{X} \end{pmatrix} \begin{pmatrix} \hat{\boldsymbol{\delta}} \\ \hat{\boldsymbol{\beta}} \end{pmatrix} \tag{4}$$

When the orthogonal block condition in Equation (3) is satisfied, it can be seen that the solution for $\boldsymbol{\beta}$ from (4) will be the same as the one from the equation $\mathbf{X}'\mathbf{y} = \mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}$, *i.e.* the equation for a model without blocking.

Remarks

1. When there are r blocking factors, the matrix $\tilde{\mathbf{Z}}$ in (2) can be partitioned as $(\tilde{\mathbf{Z}}_1 \dots \tilde{\mathbf{Z}}_r)$, where $\tilde{\mathbf{Z}}_l$ $(l = 1, \dots, r)$ is matrix of size $n \times (b_l - 1)$ and b_l the settings of the blocking factor l.

2. Let $\mathbf{x}_i'$ and $\mathbf{x}_u'$ be two rows of $\mathbf{X}$. Let $\tilde{\mathbf{z}}_i'$ and $\tilde{\mathbf{z}}_u'$ be the corresponding row vectors of $\tilde{\mathbf{Z}}$. Swapping the ith and uth row of $\mathbf{X}$ is the same as adding the following matrix to $\tilde{\mathbf{Z}}'\mathbf{X}$:

$$-(\tilde{\mathbf{z}}_i - \tilde{\mathbf{z}}_u)(\mathbf{x}_i - \mathbf{x}_u)'. \tag{5}$$

We use this matrix result to develop BLOCK, an algorithm for blocking various types of experimental designs, including DSDs and DSD-based mixed-level designs.

3. Two Steps of The BLOCK Algorithm

Here are two steps of BLOCK with r blocking factors using the results in Equation (5):

1. Allocate the n runs of the unblocked design to the blocking factors randomly. Calculate f, the sum of squares of the elements of $\tilde{\mathbf{Z}}'\mathbf{X}$.

2. Repeat searching for a pair of runs in different blocks such that the swap of the run positions results in the biggest reduction in f. If the search is successful, swap their positions, update f and $\tilde{\mathbf{Z}}'\mathbf{X}$. This step is repeated until f=0 or until f cannot be reduced further.

Each computer try has these two steps. Several tries are required for each design and the one with the smallest f will be chosen. For designs with the same f, the one with the smallest block factor (BF) will be chosen where BF is calculated as:

$$BF = (|X'X|/(|\tilde{\mathbf{Z}}'\tilde{\mathbf{Z}}||\mathbf{X}'\mathbf{X}|))^{1/(p-v)} \tag{6}$$

Here $X = (\tilde{\mathbf{Z}}\ \mathbf{X})$, p is the number of parameters to be estimated and $v = \sum_{l=1}^{r}(b_l - 1)$ is the degree of freedom associated with the r blocking factors or columns of $\tilde{\mathbf{Z}}$. Clearly, BF equals 1 means the design is orthogonally blocked.

Remarks

1. For a factorial or fractional factorial design (FFD), the orthogonality between block variables and MEs is considered more important than the one between the former and 2FIs. For a 3-level or a mixed-level screening design, the orthogonality between block variables and MEs is considered more important than the one between the former and SOEs. In these situations, partition $\mathbf{X}$ as $(\mathbf{X}_1\ \mathbf{X}_2)$ where $\mathbf{X}_1$ is associated with the more important effects and partition $\tilde{\mathbf{Z}}'\mathbf{X}$ as $(\tilde{\mathbf{Z}}'\mathbf{X}_1\ \tilde{\mathbf{Z}}'\mathbf{X}_2)$. Let g be the sum of squares of the elements of $\tilde{\mathbf{Z}}'\mathbf{X}_1$ and f the sum of squares of the elements of $\tilde{\mathbf{Z}}'\mathbf{X}$ as defined previously. g and f will then be used as the primary and secondary objective functions respectively.

2. BLOCK can be implemented sequentially, *i.e.* at each step, a blocking factor is added to the design. We use this sequential approach to block very large designs with several blocking factors.

3. In a sense, our blocking algorithm is an extension of one of Nguyen (2001), which only works with one blocking factors, to several blocking factors. Ours is more general than the one of Gilmour and Trinca (2003), which only work with two blocking factors. Clearly, ours does not require matrix inversions and therefore is considered faster (and less prone to get trapped in the local optima) than the ones by other authors (See *e.g.* Cook and Nachtsheim, 1989; Gilmour and Trinca, 2003 and Jones and Nachtsheim, 2016).

4. Discussion

In the followings, we will show the solutions for the two designs problems mentioned in the Introduction.

Example 1: Table 1 shows how a 2^5 factorial can be blocked using two blocking factors (day and time of the day) and the interaction model which includes the MEs and 2fi's terms. Our constructed design is an orthogonally blocked one, meaning the factors **A**, **B**, **C**, **D** and **E** are orthogonal to both days and times of day (8AM and 2PM).

Example 2: Table 2 (a) shows the layout of a 9-factor DSD with two blocking factors, *i.e.* reactor and day, using the pure quadratic model which includes the SOE terms and MEs terms. Three centre runs have been added to the original 21-run DSD. For this design, $g = 0,\ f = 29$ and $BF = 0.807$. Now let us assume that the unblocked design for the above experiment is a 24-run mixed-level design with the first five factors at 3-level and the rest are at 2-level constructed from a Hadamard matrix of order 12 (see Nguyen *et al.*, 2020). Table 2 (b) shows the layout the blocked design. For this design, $g = 0,\ f = 7.11$ and $BF = 0.928$.

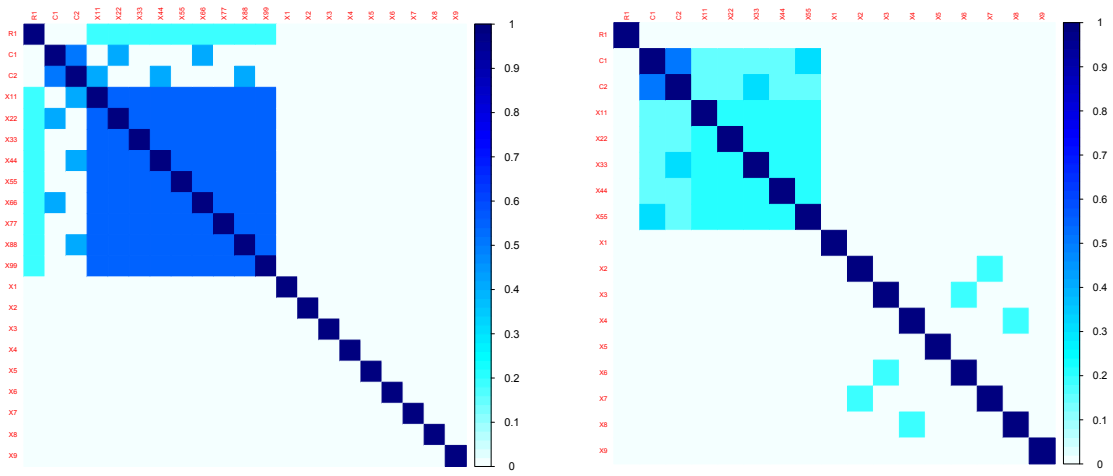
To visualise the confounding patterns of blocked designs we use the correlation cell plots (CCPs). These CCPs, proposed by Jones and Nachtsheim (2011), display the magnitude of the correlation between the blocking factors and the MEs, SOEs (of 3-level factors) and 2FIs of the designs under study. The colour of each cell in these plots ranges from white (no correlation) to dark (correlation of 1 or close to 1).

The two CCPs in Figures 1 (a) and 1 (b) show the confounding patterns of the two designs in Tables 2 (a) and 2 (b). It can be seen the first blocking factor (reactor) is orthogonal to both MEs and SOEs in Figure 1 (b) but is only orthogonal to MEs in Figure

Table 1: A 2^5 factorial with two blocking factors: day and time of the day. The low and high levels are coded as -1 and 1

Day	Time	A	B	C	D	E	Day	Time	A	B	C	D	E
1	**1**	1	1	-1	1	-1	**3**	**1**	1	-1	-1	-1	-1
1	**1**	-1	1	1	-1	-1	**3**	**1**	-1	-1	1	1	-1
1	**1**	-1	-1	-1	-1	1	**3**	**1**	1	1	1	-1	1
1	**1**	1	-1	1	1	1	**3**	**1**	-1	1	-1	1	1
1	**2**	1	-1	1	-1	-1	**3**	**2**	1	1	1	1	-1
1	**2**	-1	1	1	1	1	**3**	**2**	-1	1	-1	-1	-1
1	**2**	1	1	-1	-1	1	**3**	**2**	-1	-1	1	-1	1
1	**2**	-1	-1	-1	1	-1	**3**	**2**	1	-1	-1	1	1
2	**1**	-1	-1	-1	-1	-1	**4**	**1**	-1	1	1	-1	1
2	**1**	-1	1	1	1	-1	**4**	**1**	1	1	-1	-1	-1
2	**1**	1	1	-1	1	1	**4**	**1**	-1	-1	-1	1	1
2	**1**	1	-1	1	-1	1	**4**	**1**	1	-1	1	1	-1
2	**2**	1	1	1	-1	-1	**4**	**2**	-1	-1	1	-1	-1
2	**2**	1	-1	-1	1	-1	**4**	**2**	1	-1	-1	-1	1
2	**2**	-1	1	-1	-1	1	**4**	**2**	-1	1	-1	1	-1
2	**2**	-1	-1	1	1	1	**4**	**2**	1	1	1	1	1

1 (a). The second blocking factor (day) is orthogonal to MEs in both Figures 1 (a) and 1 (b) but is only near-orthogonal to the SOEs in these figures. We say that for both designs, the MEs of both designs are clear of block effects but the SOEs are partially confounded (slightly correlated) with block effects.



(a) (b)

Figure 1: CCPs showing the confounding patterns of: (a) a blocked 9-factor DSD, (b) a blocked mixed-level screening designs in Table 2

Table 2: Two blocked 24-run designs for Example 2: (a) a blocked 9-factor DSD; (b) a blocked mixed-level screening designs with five 3-level factors and four 2-level factors. The low, mid- and high levels are coded as $-1, 0$ and 1

Reactor	Day	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	**1**	1	1	1	-1	-1	1	-1	1	0
1	**1**	0	0	0	0	0	0	0	0	0
1	**1**	0	0	0	0	0	0	0	0	0
1	**1**	-1	-1	1	-1	1	-1	0	1	1
1	**2**	0	-1	-1	-1	-1	-1	-1	-1	-1
1	**2**	-1	1	0	1	-1	-1	1	1	-1
1	**2**	1	-1	-1	0	1	1	1	1	-1
1	**2**	0	0	0	0	0	0	0	0	0
1	**3**	1	1	-1	1	1	-1	-1	0	1
1	**3**	1	-1	1	1	-1	0	1	-1	1
1	**3**	-1	1	-1	-1	0	1	1	-1	1
1	**3**	-1	0	1	1	1	1	-1	-1	-1
2	**1**	-1	-1	-1	1	-1	1	-1	1	1
2	**1**	-1	-1	-1	1	1	-1	1	-1	0
2	**1**	1	1	-1	1	-1	1	0	-1	-1
2	**1**	1	1	1	-1	1	-1	1	-1	-1
2	**2**	0	0	0	0	0	0	0	0	0
2	**2**	-1	1	1	0	-1	-1	-1	-1	1
2	**2**	1	-1	0	-1	1	1	-1	-1	1
2	**2**	0	1	1	1	1	1	1	1	1
2	**3**	-1	-1	1	-1	-1	1	1	0	-1
2	**3**	1	-1	1	1	0	-1	-1	1	-1
2	**3**	-1	1	-1	-1	1	0	-1	1	-1
2	**3**	1	0	-1	-1	-1	-1	1	1	1

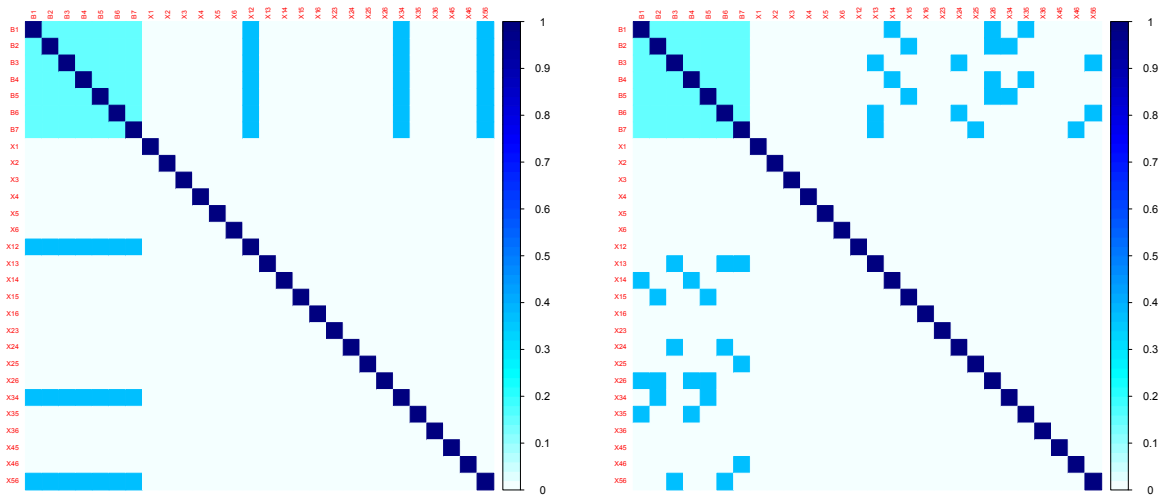
(a)

Reactor	Day	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	**1**	-1	1	-1	0	-1	-1	-1	1	1
1	**1**	1	-1	1	0	1	1	1	-1	-1
1	**1**	0	1	-1	1	1	-1	-1	-1	-1
1	**1**	0	-1	1	-1	-1	1	1	1	1
1	**2**	-1	-1	1	1	0	1	-1	1	-1
1	**2**	1	0	-1	1	1	-1	1	1	-1
1	**2**	-1	0	1	-1	-1	1	-1	-1	1
1	**2**	1	1	-1	-1	0	-1	1	-1	1
1	**3**	1	-1	0	1	-1	-1	-1	1	1
1	**3**	1	1	1	1	-1	1	-1	-1	-1
1	**3**	-1	1	0	-1	1	1	1	-1	-1
1	**3**	-1	-1	-1	-1	1	-1	1	1	1
2	**1**	-1	-1	0	1	-1	-1	1	-1	-1
2	**1**	-1	0	1	-1	1	-1	-1	1	-1
2	**1**	1	0	-1	1	-1	1	1	-1	1
2	**1**	1	1	0	-1	1	1	-1	1	1
2	**2**	0	1	1	1	1	1	1	1	1
2	**2**	0	-1	-1	-1	-1	-1	-1	-1	-1
2	**2**	-1	1	-1	0	-1	1	1	1	-1
2	**2**	1	-1	1	0	1	-1	-1	-1	1
2	**3**	1	1	1	-1	-1	-1	1	1	-1
2	**3**	-1	1	1	1	0	-1	1	-1	1
2	**3**	-1	-1	-1	1	1	1	-1	-1	1
2	**3**	1	-1	-1	-1	0	1	-1	1	-1

(b)

Table 3: A resolution V 2^{6-1} FFD arranged in eight blocks by using the BLOCK algorithm

Block	A	B	C	D	E	F	Block	A	B	C	D	E	F
1	-1	1	-1	-1	1	-1	**5**	-1	1	1	1	1	-1
1	1	1	1	1	-1	-1	**5**	-1	1	-1	-1	-1	1
1	1	-1	-1	-1	-1	1	**5**	1	-1	1	-1	1	1
1	-1	-1	1	1	1	1	**5**	1	-1	-1	1	-1	-1
2	1	1	-1	1	1	-1	**6**	-1	-1	1	-1	1	-1
2	-1	1	1	1	-1	1	**6**	-1	-1	-1	1	-1	1
2	-1	-1	-1	-1	1	1	**6**	1	1	1	1	1	1
2	1	-1	1	-1	-1	-1	**6**	1	1	-1	-1	-1	-1
3	1	-1	1	1	-1	1	**7**	1	-1	-1	1	1	1
3	-1	1	1	-1	-1	-1	**7**	-1	1	-1	1	-1	-1
3	-1	-1	-1	1	1	-1	**7**	1	1	1	-1	1	-1
3	1	1	-1	-1	1	1	**7**	-1	-1	1	-1	-1	1
4	1	1	1	-1	-1	1	**8**	1	-1	-1	-1	1	-1
4	-1	-1	-1	-1	-1	-1	**8**	-1	1	1	-1	1	1
4	1	-1	1	1	1	-1	**8**	1	1	-1	1	-1	1
4	-1	1	-1	1	1	1	**8**	-1	-1	1	1	-1	-1



(a) (b)

Figure 2: CCPs showing the confounding patterns of a resolution V 2^{6-1} FFD arranged in eight blocks by: (a) using the block generators ACE, BCE and ADE, (b) using the BLOCK algorithm

It is reasonable to compare a block design available in a catalog and that constructed by BLOCK. Let's block a resolution V 2^{6-1} FFD generated by the design generator **F=ABCDE** in eight blocks. To construct this block design, we can use the block generators **ACE**, **BCE** and **ADE** (See Table 5B.3 of Wu and Hamada, 2009). The resulting design has clear MEs and 2FIs, except **AB**, **BC** and **CD**. In other words, these 2fi's are fully confounded with blocks and cannot be estimated. In contrast, our block design in Table 3 has clear MEs and

some clear 2FIs. Most 2FIs are however, partially confounded with blocks but can still be estimated.

Figures 2 (a) and 2 (b) show the confounding patterns of the blocked resolution V 2^{6-1} FFDs in Tables 5B.3 of Wu and Hamada, 2009, and in Table 3, respectively. It can be seen from the two CCPs that the main effects of both designs are clear of block effects. However, for the CCP in Figure 2 (a), the 2fi's **AB**, **BC** and **CD** are confounded with block effects. At the same time, for the CCP in Figure 2 (b) the 2fi's are either clear or partially confounded with the block effects.

The runtime for BLOCK is minimal. BLOCK requires a couple of seconds to find a design solution for each design problem in this paper on a Mac mini M1 (each design parameter requires 1,000 tries).

5. Conclusion

Most block designs in the literature are cataloged designs and as such they are not flexible enough. The 2-level factorials and FFDs are only available in 2^q blocks, but not available in five, six or seven blocks. The 4-factor BBD, for example is available in three blocks but not in two blocks. Besides, catalogs do not offer designs having more than one blocking factor. BLOCK was developed with the philosophy "Design for experiment, not experiment for the design" in mind. We hope it could offer efficient alternatives to the existing catalog of block designs. In the pre-computer age, the constructed block design strived for simplicity in the analysis. Nowadays, it is not simplicity in the analysis but the design efficiency and the saving of experimental resources that counts.

The link to the supplemental material contains the Java implementation of the BLOCK algorithm in Section 3 is `https://drive.google.com/drive/folders/1lju4hWxI0tY4zA_sxT5qVtgtYCydH4LV`. It also contains additional examples of blocked factorial and FFDs, blocked RSDs, blocked mixture designs, and blocked DSDs and DSD-based mixed-level designs (Jones and Nachtsheim, 2013) and Hadamard design-based mixed-level designs (Nguyen *et al.*, 2020).

References

Box, G. E. P. and Hunter, J. S. (1957). Multi-factor experimental designs for exploring response surfaces. *The Annals of Mathematical Statistics*, **28**, 159-241.

Cook, R. D. and Nachtsheim, C. J. (1989). Computer-aided blocking of factorial and response surface designs. *Technometrics*, **31**, 339-346.

Gilmour, S. G. and Trinca L. A. (2003). Row-column response surface designs. *Journal of Quality Technology*, **35**, 184-193.

Hoang, E. M., Liauw, C. M., Allen, N. S., Fontan, E. and Lafuente, P. (2004). Effects of additive interactions on the thermo-oxidative stabilization of a film grade metallocene LLDPE. *Journal of Vinyl and Additive Technology*, **10**, 148-156.

Jones, B. and Nachtsheim, C. J. (2011). A class of three levels designs for definitive screening in the presence of second-order effects. *Journal of Quality Technology*, **43**, 1-15.

Jones, B. and Nachtsheim, C. J. (2013). Definitive screening designs with added two-level categorical factors. *Journal of Quality Technology*, **45**, 121-129.

Jones, B. and Nachtsheim, C. J. (2016). Blocking schemes for definitive screening designs, *Technometrics* **58**, 74-83.

Khuri, A. I. and Cornell, J. A. (1996). Response Surfaces, Designs and Analyses. 2nd ed. New York, Marcel Dekker.

Mee, R. W. (2009).*A Comprehensive Guide to Factorial Two-level Experimentation.* New York, NY: Springer.

Nguyen, N-K (2001). Cutting experimental designs into blocks. *Australia and New Zealand Journal of Statistics*, **43**, 367-374.

Nguyen, N-K., Pham, D-T. and Vuong, M. P. (2020). Constructing D-efficient mixed-level foldover designs using Hadamard matrices. *Technometrics*, **62**, 48-56,

Wu, C. F. J and Hamada, M. (2009). Experiments: Planning, Analysis and Parameter Design Optimisation. New York: John Wiley and Sons, Inc.

Zhang, C, Wei Chen, W, Xiana, J. and Degang Fub, D. (2018). Application of a novel definitive screening design to in situ chemical oxidation of acid orange-II dye by a Co^{2+}/PMS system. *RSC Advances*, **8**, 3934-3940.

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 257–275

Optimal Row-Column Designs with Three Rows

J. P. Morgan[1] and Siriluck Jermjitpornchai[2]
[1]*Department of Statistics*
Virginia Tech, Blacksburg, Virginia, United States
[2]*Department of Mathematics*
Mahasarakham University, Mahasarakham, Thailand

Received: 31 December 2020; Revised: 7 February 2021; Accepted: 10 February 2021

Abstract

E and MV optimality results are established for several series of 3-rowed row-column designs. All of these designs are generalized binary in rows, and their column component designs have completely symmetric information matrices. Included are optimal designs which are nonbinary in columns, and which are superior to any competitor that is binary in columns. The optimality of designs with BIBD column components is extended beyond that of regular Youden designs.

Key words: Block design; Completely symmetric design; E optimality; MV optimality.

AMS Subject Classifications: 62K05, 05B05

1. Introduction

Block designs are useful for experiments where it is important to eliminate an identified source of heterogeneity in experimental units. In many experimental situations, the position that a unit occupies within a block can also affect observed responses. In such cases, row-column designs can often be used to additionally eliminate heterogeneity in this second, orthogonal direction. Applications of row-column designs range from agriculture to psychology to industry and beyond, and an abundance of examples can be found in design textbooks, for example Bailey (2008) or Dean *et al.* (2017).

Consider then bk experimental units arranged in a $k \times b$ array to which v treatments are to be assigned. Determining optimal row-column designs, aside from situations where treatments can either be equally replicated in each row or equally replicated in each column (which includes such well-known designs as Latin squares and regular Youden designs), has proven to be a challenging task. The limited results up to about 1989 are summarized in the monograph Shah and Sinha (1989), and from there to 2015 in Morgan (2015). There has been little progress since.

We proceed with the standard, additive linear model with uncorrelated, equi-variable errors, and in which the expected response arising from the unit in row l, column j is the sum of the effect of row l, the effect of column j, and the effect of the treatment employed in

Corresponding Author: J. P. Morgan
Email: jpmorgan@vt.edu

that cell. With this model, the $v \times v$ treatment information matrix C_d for a $k \times b$ row-column design d can be written (see, for instance, Shah and Sinha, 1989, p. 66) as

$$C_d = diag(r_{d1}, \ldots, r_{dv}) - \frac{1}{k} N_d N_d' - \frac{1}{b} M_d M_d' + \frac{1}{bk} r_d r_d' \tag{1}$$

where

- N_d denotes the $v \times b$ treatment-column incidence matrix whose entries n_{dij} are nonnegative integers indicating the number of times treatment i occurs in column j,
- M_d denotes the $v \times k$ treatment-row incidence matrix whose entries m_{dil} are nonnegative integers indicating the number of times treatment i occurs in row l,
- r_d denotes the $v \times 1$ vector with entries r_{di}, where r_{di} is the replication of treatment i in design d (*i.e.* the number of experimental units to which treatment i is assigned).

The ith row sum of both N_d and M_d is r_{di}. The matrices $N_d N_d'$ and $M_d M_d'$ are respectively the column-concurrence and row-concurrence matrices for d, with entries denoted by $\lambda_{dii'}$ and $\mu_{dii'}$ respectively. The individual entries of C_d can be displayed thusly:

$$c_{dii'} = r_{di}\delta_{ii'} - \frac{1}{k}\lambda_{dii'} - \frac{1}{b}\mu_{dii'} + \frac{1}{bk} r_{di} r_{di'} \tag{2}$$

where $\delta_{ii'}$ is the Kronecker delta.

C_d is known to be symmetric and nonnegative definite with zero row and column sums. Hence the rank of C_d satisfies $r(C_d) \leq v - 1$. Here only designs with $r(C_d) = v - 1$ are considered. These are exactly the designs for which every contrast $\sum_i l_i \tau_i$ ($\sum_i l_i = 0$) is estimable, commonly termed the connected designs. For given v, b, and k, $\mathcal{D}(v, b, k)$ will denote the class of all connected $k \times b$ row-column designs for v treatments.

For a given class $\mathcal{D}(v, b, k)$, define the *replication target* r by $r = \lfloor \frac{bk}{v} \rfloor$, and so $bk = vr + p$ where $p \in \{0, 1, \ldots, v-1\}$ is the *plot excess*, that is, the number of experimental units available beyond that needed for equal replication. Also define $\lambda = \lfloor \frac{r(k-1)}{(v-1)} \rfloor$ as the *concurrence target* for the column component design. Then $\lambda(v-1) = r(k-1) + q$ where $q \in \{0, 1, \ldots, v-1\}$ is the *(column) concurrence excess* for a treatment replicated r times. In this paper we study settings $\mathcal{D}(v, b, k)$ for which $q = 0$, which is henceforth assumed, allowing frequent use of the equality $r(k-1) = \lambda(v-1)$.

Each row-column design $d \in \mathcal{D}(v, b, k)$ can be associated with two block designs: the column component design d_N and the row component design d_M, having respective information matrices $C_{d_N} = diag(r_{d1}, \ldots, r_{dv}) - \frac{1}{k} N_d N_d'$ and $C_{d_M} = diag(r_{d1}, \ldots, r_{dv}) - \frac{1}{b} M_d M_d'$. The assignment of treatment i in design d is said to be generalized binary in columns (rows) if $n_{dij} \in \{\lfloor \frac{k}{v} \rfloor, \lfloor \frac{k}{v} \rfloor + 1\}$ for all $j = 1, \ldots b$ ($m_{dil} \in \{\lfloor \frac{b}{v} \rfloor, \lfloor \frac{b}{v} \rfloor + 1\}$ for all $l = 1, \ldots k$). The component block design d_N (d_M) is called generalized binary if the assignment of every treatment to columns (rows) is generalized binary. A generalized binary assignment in which the two counts are 0 and 1 is said to be binary.

Following Kiefer (1975), the treatment information matrix for a design is said to be completely symmetric if its elements are constant on the diagonal and constant off the

diagonal. The condition $q = 0$ allows the possibility that C_{d_N} is completely symmetric, which will play an important role in the optimality proofs in this article. When C_{d_N} is completely symmetric, d_N will be termed a completely symmetric design, or CSD for short.

From (1), C_d can be written as

$$C_d = C_{d_N} - \frac{1}{b} M_d (I_k - \frac{1}{k} J_k) M_d' \tag{3}$$

$$\leq C_{d_N} \tag{4}$$

where "$\leq$" in (4) is with respect to the Loewner ordering, by virtue of the fact that $M_d(I_k - \frac{1}{k}J_k)M_d'$ is nonnegative definite. These relationships will be useful in what follows. It can be immediately noted that for any optimality criterion respecting the Loewner ordering, a design d whose column component design is optimal, and for which $M_d M_d' - \frac{1}{k} rr' = M_d(I_k - \frac{1}{k}J_k)M_d' = 0$, is optimal over $\mathcal{D}(v, b, k)$. Now $M_d(I_k - \frac{1}{k}J_k)M_d' = 0$ if, and only if, m_{dil} is constant in l for each i, so $m_{dil} = m_{di}$ (say) for $i = 1 \ldots, v$; row-column designs fulfilling this condition are said to be of *Youden type* (Das and Dey, 1990). If further m_{di} is constant in i, the design is said to be *row-regular*. This is the basis for proving optimality of regular Youden designs, which are row-regular row-column designs with optimal, completely symmetric, column component designs (see pp. 66-7 of Shah and Sinha, 1989). The settings explored in this article preclude row-regularity for designs that are equireplicate or nearly so.

The determination of optimal row-column designs is pursued here in terms of two distinct optimality criteria. Let $z_{d0} = 0 < z_{d1} \leq z_{d2} \leq \ldots \leq z_{dv-1}$ denote the eigenvalues of the information matrix C_d associated with $d \in \mathcal{D}(v, b, k)$. The E optimality criterion aims to maximize the smallest positive eigenvalue z_{d1} of C_d. In terms of contrast variances, an E-optimal design minimizes, over $d \in \mathcal{D}$, the maximum of $Var_d(\widehat{\sum_{i=1}^{v} l_i \tau_i})$ over all possible choices of normalized ($\sum_i l_i^2 = 1$) contrast coefficients. The MV criterion requires minimizing, over $d \in \mathcal{D}$, the maximum variance over all elementary treatment contrasts $\widehat{\tau_i - \tau_{i'}}$. That is, an MV-optimal design minimizes over $d \in \mathcal{D}$ the quantity

$$\Upsilon_d = \max_{i \neq i'} \frac{Var_d(\widehat{\tau_i - \tau_{i'}})}{\sigma^2}.$$

The E and MV criteria are both minimax criteria, and both criteria respect the Loewner ordering.

This article examines row-column designs with $k = 3$ from the perspectives of the E and MV optimality criteria. Section 2 develops a series of bounds that are useful in constructing the optimality arguments. The main results are in Section 3. Concluding discussion comprises Section 4.

2. Preliminaries

Useful lemmas, to be employed in the optimality proofs in Section 3, are established here. At the heart of these are two results, stated first, due to M. Jacroux. For information matrix $C_d = (c_{dii'})$ define the quantity $\theta_{dii'}$ for each $i \neq i'$ by $\theta_{dii'} = c_{dii} + c_{di'i'} - 2c_{dii'}$.

From (2), $\theta_{dii'}$ may be written as

$$\theta_{dii'} = r_{di} + r_{di'} - \frac{1}{bk}[b\sum_{j=1}^{b}(n_{dij} - n_{di'j})^2 + k\sum_{l=1}^{k}(m_{dil} - m_{di'l})^2] + \frac{(r_{di} - r_{di'})^2}{bk} \qquad (5)$$

Lemma 1: (Jacroux, 1982) Let $d \in \mathcal{D}(v, b, k)$ have information matrix $C_d = (c_{dii'})$.

(i) If M is a subset of $\{1, 2, \ldots, v\}$ of size m, $1 \le m \le v - 1$, then
$z_{d1} \le (v/m(v-m))(\sum_{i\in M} c_{dii} + \sum_{i\in M}\sum_{i'(\neq i)\in M} c_{dii'})$;

(ii) $z_{d1} \le \theta_{dii'}/2$ for $i, i' = 1, \ldots, v$ $(i \neq i')$.

Moreover, in light of (4), d in the right-hand side of each of these two inequalities can be replaced by d_N, providing two additional (and possibly less sharp) upper bounds for z_{d1}.

Lemma 2: (Jacroux, 1983) Let $d \in \mathcal{D}(v, b, k)$ have information matrix $C_d = (c_{dii'})$. For any $i \neq i'$,

$$(1/\sigma^2)Var_d(\widehat{\tau_i - \tau_{i'}}) \ge \frac{4}{\theta_{dii'}} \ge \frac{4}{\theta_{d_N ii'}}.$$

Proofs for Lemmas 1 and 2 may be found in the cited papers. Proofs for the lemmas that follow appear in the appendix.

Lemma 3: A row-column design $d \in \mathcal{D}(v, b, 3)$ for which the column design is a CSD with $c_{d_N^* ii} = \frac{r(k-1)}{k}$ for all $i = 1, \ldots, v$ satisfies

$$z_{d1} = \frac{\lambda v}{3} - \frac{1}{2b}\left(a_{1d} + \sqrt{2a_{2d} - a_{1d}^2}\right)$$

where $a_{1d} = tr(B_d)$ and $a_{2d} = tr(B_d^2)$ for the matrix $B_d = M_d M_d' - \frac{1}{k} r_d r_d'$.

Lemma 4: A row-column design $d \in \mathcal{D}(v, b, 3)$ for which $r_{di} \le r - 1$ for some i satisfies

$$\theta_{dii'} \le \frac{2\lambda v}{3} + \frac{2[v(q-1) + p]}{3(v-1)}.$$

for some $i' \neq i$.

Lemma 5: A row-column design $d \in \mathcal{D}(v, b, k)$ for which some treatment with replication r is nonbinary in columns satisfies

$$\theta_{dii'} \le \frac{2vr(k-1)}{k(v-1)} - \frac{2}{k} + \frac{p(k-1) - 4}{k(v-1)}$$

for some $i \neq i'$.

Lemma 6: For the row-column setting $\mathcal{D}(v, b, k)$ with $p \le v - 2$, let $d \in \mathcal{D}(v, b, k)$ have $r_{di} \ge r$ for all i. If two treatments $i \neq i'$ with $r_{di} = r = r_{di'}$ have $\lambda_{dii'} \le \lambda - 1$, then

$$\theta_{dii'} \le \frac{2(\lambda v + q - 1)}{k}.$$

Lemma 7: Let $d \in \mathcal{D}(v, b, 3)$ satisfy

(i) $r_{d1} = \ldots = r_{d,v-p} = r$,

(ii) $r_{d,v-p+1} = \ldots = r_{dv} = r + 1$, and

(iii) the treatments with replication $r + 1$ are generalized binary in rows.

If some treatment with replication r is not generalized binary in rows, where $r \equiv 1 \pmod 3$ and $p < (v + 3)/2$, or where $r \equiv 2 \pmod 3$, then

$$\theta_{dii'} \leq 2r - \frac{1}{3b}[b\sum_{j=1}^{b}(n_{dij} - n_{di'j})^2 + 18]$$

for some $i \neq i'$ with $i, i' \leq v - p$.

Lemma 8: Let $d \in \mathcal{D}(v, b, 3)$ satisfy

(i) $r_{d1} = \ldots = r_{d,v-p} = r$, and

(ii) $r_{d,v-p+1} = \ldots = r_{dv} = r + 1$.

If some treatment with replication $r+1$ is not generalized binary in rows, where $r \equiv 1 \pmod 3$ and $p < (v + 3)/2$, or where $r \equiv 2 \pmod 3$ and $p < (v + 6)/4$, then

$$\theta_{dii'} \leq 2r + 1 - \frac{1}{3b}[b\sum_{j=1}^{b}(n_{dij} - n_{di'j})^2 + 14]$$

for some $i \leq v - p$ and some $i' \geq v - p + 1$.

3. Main Results

As stated in Section 1, design classes $\mathcal{D}(v, b, k)$ for which the concurrence excess is $q = 0$ are the focus in the theorems to follow. Optimality results for the E and MV criteria will now be derived for the plot excess p taking the values $p = 1$ and $p = 0$. The same techniques can be successfully applied for some larger p, but for reasons of space are not pursued here.

3.1. Designs with $p = 1$

Theorem 1: Let $\mathcal{D}(v, b, 3)$ be the class of $3 \times b$ row-column designs for which the plot excess is $p = 1$, the concurrence excess is $q = 0$, and $b > v + 2$. If the target replication r satisfies $r \equiv 1 \pmod 3$ and $d^* \in \mathcal{D}(v, b, 3)$ satisfies

(i) the column design of d^* is a CSD with $c_{d_N^* ii} = \frac{r(k-1)}{k}$ for all $i = 1, \ldots, v$, and

(ii) the row design of d^* is generalized binary,

then d^* is E-optimal in $\mathcal{D}(v,b,3)$, and any design failing either (i) or (ii) is E-inferior to d^*. In particular, any design which is binary in columns is E-inferior to d^*.

Proof: Condition (i) implies that $r_{d^*1} = \ldots = r_{d^*v-1} = r$ and thus $r_{d^*v} = r+1$. It then further implies that treatment v is nonbinary in columns of d^*.

For each treatment i, define the vector of row counts m_{di} for design d as $m_{di} = (m_{di1}, m_{di2}, m_{di3})$, the i^{th} row of M_d. With appropriate labeling of treatments and ordering of rows, condition (ii) implies that $m_{d^*v} = (\frac{r+2}{3}, \frac{r+2}{3}, \frac{r-1}{3})$ and so $m_{d^*i} = (\frac{r+2}{3}, \frac{r-1}{3}, \frac{r-1}{3})$ for $i = 1, \ldots, \frac{v-2}{3}$; $m_{d^*i} = (\frac{r-1}{3}, \frac{r+2}{3}, \frac{r-1}{3})$ for $i = \frac{v+1}{3}, \ldots, \frac{2(v-2)}{3}$; and $m_{d^*i} = (\frac{r-1}{3}, \frac{r-1}{3}, \frac{r+2}{3})$ for $i = \frac{2v-1}{3}, \ldots, v-1$. With M_{d^*} so determined, write $v_1 = v_2 = (v-2)/3$, $v_3 = (v+1)/3$ and $v_4 = 1$ for the frequencies of its four distinct rows. It follows that the symmetric matrix $B_{d^*} = M_{d^*}(I_3 - \frac{1}{3}J_3)M'_{d^*}$ is the partitioned block matrix with diagonal blocks $\frac{2}{3}J_{v_g v_g}$ for $g = 1, \ldots, 4$, and blocks $-\frac{1}{3}J_{v_g v_h}$ for $g < h \leq 3$, $\frac{1}{3}J_{v_g v_4}$ for $g = 1, 2$, and $-\frac{2}{3}J_{v_3 v_4}$ above the diagonal. It is now simple to calculate $a_{1d^*} = tr(B_{d^*}) = \frac{2v}{3}$ and $a_{2d^*} = tr(B^2_{d^*}) = \frac{2(v^2+4)}{9}$, so that by Lemma 3,

$$z_{d^*1} = \frac{\lambda v}{3} - \frac{(v+2)}{3b}.$$

Let d be any other design with $r_{d1} \leq \ldots \leq r_{dv}$. Suppose d has $r_{di} \leq r-1$ for some $i = 1, \ldots, v$. Then d is E-inferior to d^*, as by Lemma 1, and Lemma 4 with $p = 1$,

$$z_{d1} \leq \frac{\lambda v}{3} + \frac{v(q-1)+p}{3(v-1)} = \frac{\lambda v}{3} - \frac{1}{3} < \frac{\lambda v}{3} - \frac{(v+2)}{3b} = z_{d^*1}, \tag{6}$$

the last inequality due to the condition $b > v+2$.

So assume d has $r_{di} \geq r$ for all i; with appropriate labeling of treatments, $r_{di} = r_{d^*i}$ for all i. If d has some treatment with replication r nonbinary in columns, Lemmas 1 and 5 give

$$z_{d1} \leq \frac{vr(k-1)}{k(v-1)} - \frac{1}{k} + \frac{p(k-1)-4}{2k(v-1)} = \frac{\lambda v}{3} - \frac{1}{3} - \frac{1}{3(v-1)} < \frac{\lambda v}{3} - \frac{(v+2)}{3b} = z_{d^*1} \tag{7}$$

and again d is E-inferior to d^*.

So now suppose d has treatments $1, \ldots, v-1$ each binary in columns. If $\lambda_{dii'} \leq \lambda - 1$ for some $i, i' \leq v-1$ $(i \neq i')$, then by Lemmas 1 and 6,

$$z_{d1} \leq \frac{\lambda v}{3} - \frac{1}{3} < \frac{\lambda v}{3} - \frac{(v+2)}{3b} = z_{d^*1} \tag{8}$$

and again d is E-inferior to d^*.

Thus for d to be E-admissible, it must have $r_{d1} = \ldots = r_{dv-1} = r$, and all treatments replicated r times are binary in columns with $\lambda_{dii'} = \lambda$ for all $i \neq i'$ where $i, i' = 1, \ldots, v-1$. If in addition treatment v is binary in columns then $\sum_{i=1}^{v-1} \lambda_{dvi} = 2(r+1)$, implying

$$\sum_{i=1}^{v-2} \sum_{i'=i+i}^{v-1} \lambda_{dii'} \;=\; \sum_{i=1}^{v-1} \sum_{i'=i+i}^{v} \lambda_{dii'} - \sum_{i=1}^{v-1} \lambda_{dvi} \;=\; 3b - 2(r+1) \;=\; (v-2)r - 1,$$

the last equality because $bk = 3b = vr + 1$. Thus the average concurrence among treatments replicated r times is

$$\frac{(v-2)r-1}{(v-1)(v-2)/2} = \frac{(v-2)r-1}{(v-2)r/\lambda} = \lambda - \frac{1}{(v-2)r}$$

so that $\lambda_{dii'} < \lambda$ for some $i \neq i'$, $i, i' < v$. As shown above (see (8)), this implies d is E-inferior to d^*, so d cannot have treatment v binary in columns.

Now $c_{dvv} \leq c_{d_N vv} = r + 1 - (\sum_{j=1}^{b} n_{dvj}^2)$ and nonbinarity of treatment v in columns implies $\sum_{j=1}^{b} n_{dvj}^2 \in \{r+3, r+5, r+7 \ldots\}$. If $\sum_{j=1}^{b} n_{dvj}^2 \geq r+5$ then $c_{dvv} \leq (r+1)-(r+5)/3 = 2(r-1)/3$ and by Lemma 1(i),

$$z_{d1} \leq \left(\frac{v}{v-1}\right) c_{dvv} = \frac{2v(r-1)}{3(v-1)} = \frac{\lambda v(r-1)}{3r}$$

$$\Rightarrow 3r(z_{d^*1} - z_{d1}) \geq \lambda vr - \frac{(v+2)r}{b} - \lambda v(r-1) = \lambda v - \frac{(v+2)r}{b} = \lambda + 2r - \frac{(v+2)r}{b} > 0,$$

again invoking $b > v + 2$. So d must have $\sum_{j=1}^{b} n_{dvj}^2 = r + 3 \Rightarrow c_{d_N vv} = 2r/3 = c_{d_N ii}$ for $i = 1, \ldots, v-1$. The same argument as in (8) implies $\lambda_{dvi} \geq \lambda$ for $i < v$ and hence $\lambda_{dii'} \geq \lambda$ for all $i \neq i' \Rightarrow \lambda_{dii'} = \lambda$ for all $i \neq i'$. This establishes that E-admissibility of d requires $C_{d_N} = C_{d_N^*}$, that is, d and d^* have identical information matrices for their column component designs. If C_d differs from C_{d^*}, it can do so only in the term $M_d M_d'$.

If the row design of d is generalized binary in rows then d fulfills the conditions of the theorem, that is, d is a version of d^*. So suppose d is not generalized binary in rows. Then $a_{1d} = \sum_i (\mu_{dii} - \frac{r_{di}^2}{k}) \geq (2 + \sum_i \mu_{d^*ii}) - (\sum_i \frac{r_{d^*i}^2}{k}) = a_{1d^*} + 2 = \frac{2(v+3)}{3}$. By Lemma 3,

$$z_{d1} = \frac{\lambda v}{3} - \frac{1}{2b}(a_{1d} + \sqrt{2a_{2d} - a_{1d}^2}) \leq \frac{\lambda v}{3} - \frac{1}{2b} a_{1d} \leq \frac{\lambda v}{3} - \frac{(v+3)}{3b} < \frac{\lambda v}{3} - \frac{(v+2)}{3b} = z_{d^*1},$$

completing the proof. □

The strict inequality $b > v+2$, employed in equations (6) to (8), serves only to guarantee that any design failing either condition (i) or (ii) will be E-inferior to d^*. With that note, E optimality of d^* also holds for $b = v + 2$. Following Theorem 2, it will be shown that when $b = v + 2$ it is possible to find other E-optimal designs not satisfying (i) and (ii). This issue does not arise in the MV optimality proof of Theorem 2, which otherwise covers the same settings $\mathcal{D}(v, b, k)$ as in Theorem 1.

Theorem 2: Let $\mathcal{D}(v, b, 3)$ be the class of $3 \times b$ row-column designs for which the plot excess is $p = 1$, the concurrence excess is $q = 0$, and $b \geq v + 2$. If the target replication r satisfies $r \equiv 1 \pmod 3$ and $d^* \in \mathcal{D}(v, b, 3)$ satisfies

(i) the column design of d^* is a CSD with $c_{d_N^* ii} = \frac{r(k-1)}{k}$ for all $i = 1, \ldots, v$, and

(ii) the row design of d^* is generalized binary,

then d^* is MV-optimal in $\mathcal{D}(v,b,3)$, and any design failing either (i) or (ii) is MV-inferior to d^*. In particular, any design which is binary in columns is MV-inferior to d^*.

Proof: C_{d^*}, as determined in the proof of Theorem 1, has generalized group divisible structure (see *e.g.* Srivastav and Morgan, 1998) with four groups, call them V_g for $g = 1,2,3,4$ with $|V_g| = v_g$. Accordingly, $\widehat{Var_{d^*}(\tau_i - \tau_{i'})}$ depends only on group membership for i and i'. Writing $T_{d^*} = C_{d^*} + (\frac{\lambda}{3} - \frac{1}{3b})J_{vv}$, a generalized inverse of C_{d^*} is $T_{d^*}^{-1}$, from which the pairwise variances arising from d^*, displayed in Table 1, are easily found. The reader may check that these five variances satisfy $var_1 < var_4 < var_2 < var_3 < var_5$. For the purposes of this proof $var_5 = \Upsilon_{d^*}$ is rewritten as

$$\Upsilon_{d^*} \;=\; \frac{6}{\lambda v}\left[1 + \frac{4}{\lambda vb - (v+2)}\right] \;=\; \frac{6}{\lambda v}\left[\frac{\lambda vb - (v-2)}{\lambda vb - (v+2)}\right] \;=\; \frac{2}{\frac{\lambda v}{3} - \frac{4}{3}\frac{\lambda v}{\lambda vb - (v-2)}}.$$

Let d be any other design with $r_{d1} \le \ldots \le r_{dv}$. First suppose d has $r_{di} \le r-1$ for some $i = 1,\ldots,v$. Assume $r_{d1} \le r-1$. From Lemmas 2 and 4,

$$\Upsilon_d \;\ge\; \frac{2}{\frac{\lambda v}{3} - \frac{1}{3}} \;\ge\; \frac{2}{\frac{\lambda v}{3} - \frac{(v+2)}{3b}} \;>\; \frac{2}{\frac{\lambda v}{3} - \frac{4}{3}\frac{\lambda v}{\lambda vb - \lambda v}} \;>\; \frac{2}{\frac{\lambda v}{3} - \frac{4}{3}\frac{\lambda v}{\lambda vb - (v-2)}} \;=\; \Upsilon_{d^*} \qquad (9)$$

and therefore d^* is MV-better than d.

Next, suppose d has $r_{di} \ge r$ for $i = 1,\ldots,v$. Since $p = 1$, there are $v-1$ treatments replicated r times and one treatment replicated $r+1$ times. Assume $r_{d1} = \ldots = r_{dv-1} = r$ and $r_{dv} = r+1$.

Suppose d has some treatment with replication r nonbinary in columns. Then by the same calculation as in (7), there is some $\theta_{dii'} \le \frac{2\lambda v}{3} - \frac{2(v+2)}{3b}$. Applying Lemma 2 as shown in (9) gives $\Upsilon_d > \Upsilon_{d^*}$. The same θ bound, and thus the same result for Υ_d, is found (see (8)) if two treatments $i < i' < v$ are binary in columns with $\lambda_{dii'} \le \lambda - 1$. Thus the column component design for d must be binary in treatments $1,\ldots,v-1$, and it must have $\lambda_{dii'} \ge \lambda$ for any two of these $v-1$ treatments.

If treatment v is binary in columns of d, then just as shown in the proof of Theorem 2, $\lambda_{dii'} \le \lambda - 1$ for some $i < i' < v$ and so, as established in the preceding paragraph, $\Upsilon_d > \Upsilon_{d^*}$. If treatment v is nonbinary in columns and $\sum_j n_{dvj}^2 \ge r+5$, then as established in the proof

Table 1: Pairwise variances for d^* of Theorem 2

$i \ne i'$	$\frac{1}{\sigma^2}\widehat{Var_{d^*}(\tau_i - \tau_{i'})}$
$i,i' \in V_g,\ g = 1,2,3$	$var_1 = \frac{6}{\lambda v}$
$i \in V_1,\ i' \in V_2$	$var_2 = \frac{6}{\lambda v}[1 + \frac{3}{\lambda vb - (v-2)}]$
$i \in V_1$ or $V_2,\ i' \in V_3$	$var_3 = \frac{6}{\lambda v}[1 + \frac{3(\lambda vb - (v-1))}{(\lambda vb - (v+2))(\lambda vb - (v-2))}]$
$i \in V_1$ or $V_2,\ i' \in V_4$	$var_4 = \frac{6}{\lambda v}[1 + \frac{2(\lambda vb - 3(v-1))}{(\lambda vb - (v+2))(\lambda vb - (v-2))}]$
$i \in V_3,\ i' \in V_4$	$var_5 = \frac{6}{\lambda v}[1 + \frac{4}{\lambda vb - (v+2)}]$

of Theorem 1, $c_{dvv} \le 2(r-1)/3$. This yields

$$\frac{\sum_{i=1}^{v-1}\lambda_{dvi}}{v-1} = \frac{3c_{dvv}}{v-1} \le \frac{2(r-1)}{v-1} = \frac{\lambda(v-1)-2}{v-1}$$

implying that for some $i < v$, say $i = 1$, $\lambda_{dv1} \le \lambda - 1$. Then θ_{dv1} is

$$\theta_{dv1} = \frac{2r}{3} + \frac{2(r-1)}{3} + \frac{2\lambda_{dv1}}{3} \le \frac{4r-1+2(\lambda-1)}{3} = \frac{2\lambda v-2}{3}$$

which upon applying Lemma 2 as shown in (9) again gives $\Upsilon_d > \Upsilon_{d^*}$. Thus $\sum_j n_{dvj}^2 = r+3$, $c_{d_N vv} = 2r/3$, and just as for E-admissibility in the proof of Theorem 2, it follows that that MV-admissibility of d requires $C_{d_N} = C_{d_N^*}$.

It remains to consider the row component design for d. If the row design of d is generalized binary in rows then d fulfills the conditions of the theorem, that is, d is a version of d^*. So suppose d is not generalized binary in rows.

First consider if treatment i, for some $i \le v-1$, is not generalized binary in rows, but treatment v is. Note that $\sum_{j=1}^{b}(n_{dij} - n_{di'j})^2 = 2(r-\lambda)$ for any $i < i' < v$. Lemma 7 now says there is $\theta_{dii'}$ for which

$$\theta_{dii'} \;\le\; 2r - \frac{1}{3b}[2b(r-\lambda)+18] \;=\; \frac{2(\lambda vb-9)}{3b}.$$

Employing this inequality in Lemma 2,

$$\Upsilon_d \;\ge\; \frac{4}{\theta_{dii'}} \;\ge\; \frac{6}{\lambda v}\left(\frac{\lambda vb}{\lambda vb-9}\right) \;=\; \frac{6}{\lambda v}\left(1+\frac{9}{\lambda vb-9}\right) \;>\; \frac{6}{\lambda v}\left(1+\frac{4}{\lambda vb-(v+2)}\right) \;=\; \Upsilon_{d^*}.$$

Suppose treatment v is not generalized binary in rows. It has been established above that $\sum_{j=1}^{b} n_{dvj}^2 = r+3$, so that treatment v occurs twice in one column, and once in each of $r-1$ columns. Let treatment 1 be the treatment that appears in the block with $n_{dvj} = 2$. Then $\sum_{j=1}^{b}(n_{dvj} - n_{d1j})^2 = 2(r-\lambda)+3$. Lemma 8 now says

$$\theta_{d1v} \;\le\; 2r+1-\frac{1}{3b}[b\sum_{j=1}^{b}(n_{dij}-n_{di'j})^2+14] \;=\; \lambda(v-1)+1-\frac{[2(r-\lambda)+3]}{3}-\frac{14}{3b} \;=\; \frac{2(\lambda vb-7)}{3b}.$$

By Lemma 2,

$$\Upsilon_d \;\ge\; \frac{4}{\theta_{div}} \;\ge\; \frac{6}{\lambda v}\left(\frac{\lambda vb}{\lambda vb-7}\right) \;=\; \frac{6}{\lambda v}\left(1+\frac{7}{\lambda vb-7}\right) > \frac{6}{\lambda v}\left(1+\frac{4}{\lambda vb-(v+2)}\right) \;=\; \Upsilon_{d^*}.$$

□

Example 1: Consider these designs $d_1, d_2 \in \mathcal{D}(5,7,3)$:

$$d_1: \begin{array}{ccccccc} 5 & 1 & 3 & 2 & 4 & 4 & 3 \\ 5 & 2 & 1 & 5 & 2 & 3 & 4 \\ 4 & 5 & 5 & 3 & 1 & 1 & 2 \end{array} \qquad d_2: \begin{array}{ccccccc} 4 & 1 & 1 & 2 & 5 & 4 & 3 \\ 1 & 2 & 5 & 3 & 2 & 3 & 4 \\ 5 & 5 & 3 & 1 & 4 & 5 & 2 \end{array}$$

Design d_1 satisfies the conditions of Theorems 1 and 2, and so is E-optimal and MV-optimal in $\mathcal{D}(5, 7, 3)$. As discussed following the proof of Theorem 1, since $b = v + 2$ the E optimality may not be unique. Design d_2 fails both conditions (i) and (ii) of Theorem 1, but has the same E value as d_1. Thus d_2 is also E-optimal, even while being MV-inferior to d_1.

Example 2: Consider this design d^* for $v = 8, b = 19$, and $k = 3$:

1	2	3	4	5	6	7	1	3	4	5	8	8	7	2	6	1	2	7
2	3	4	5	6	7	1	8	2	8	4	5	6	1	7	3	3	4	8
4	5	6	7	1	2	3	2	8	3	8	6	1	4	5	7	5	6	8

All conditions of Theorem 1 and Theorem 2 are satisfied, so d^* is E- and MV-optimal in $\mathcal{D}(8, 19, 3)$. Any design that is binary in columns, or which is row-regular, is inferior to d^* on both of these criteria.

The column component designs for the design in Example 2, and for d_1 in Example 1, are two instances of an infinite series of completely symmetric block designs with $k = 3$ constructed by Morgan and Uddin (1995). Those designs have, for each $t \geq 1$, parameters

$$v = 3t + 2,\ b = 3t^2 + 3t + 1,\ k = 3,\ r = 3t + 1,\ \lambda = 2,\ p = 1,\ q = 0. \tag{10}$$

Also, each of these block designs has $c_{dii} = r(k-1)/k = 2r/3$ for all i, so if the blocks are arranged as columns of a $3 \times b$ row-column design, condition (i) of the theorems will be satisfied. To satisfy condition (ii), treatments must then be ordered in each column so that each treatment appears in each row either t or $t + 1$ times. Examples 1 and 2 demonstrate that this can be done for $t = 1, 2$. That it can be done for all t can be proven using systems of distinct representatives, illustrating application of a result due to Das and Dey (1989).

Lemma 9: (Das and Dey, 1989) If a block design with v treatments in b blocks of size k has treatment replication numbers $r_i = km_i$ for integer values m_i and $i = 1, \ldots v$, then the blocks can be arranged as columns of a $k \times b$ row-column design so that, for $i = 1, \ldots v$, treatment i occurs in each row m_i times.

The row-column arrangement guaranteed by Lemma 9 is a Youden type design.

The Morgan and Uddin (1995) construction of CSDs with parameters (10) is divided into four cases. One of those (their Case 2(a)) will be employed here to show how condition (ii) can be achieved; the other three cases are handled similarly. Designs in this case comprise the subseries of (10) with $t = 4w + 2$ having $v = 12w + 8$ and $r = 12w + 7$. The treatment symbols in that construction are ∞_1, ∞_2, and the integers (mod $12w + 6$). To apply Lemma 9, partition the blocks of this subseries into four subdesigns as shown in Table 2. Counting replications in each of the subdesigns S_1, S_2, and S_3 shows that, by Lemma 9, each can can be arranged as a Youden type design on the treatments involved. Thus taken together, these three subdesigns form a $3 \times (b - (v + 1)/3)$ row-column design on all v treatments that is row-regular, all treatments being replicated $r - 1$ times. Adding the $(v + 1)/3$ columns of S_4 as displayed in Table 2, for which each treatment except ∞_1 appears exactly once, gives the required design satisfying condition (ii) of the theorems.

Example 2 contains the Case 2(a) design with $w = 0$. In that example, the first seven columns (making a Youden design) are subdesign S_1, the next nine (making a Youden type

design) are subdesign S_3, and the last three are subdesign S_4. Since $w = 0$, subdesign S_2 is empty.

For other CSDs having block size $k = 3$ with plot excess $p \geq 1$, to which the methods of this article can be applied, see Morgan and Srivastav (2002).

3.2. Designs with $p = 0$

Theorem 3: Let $\mathcal{D}(v, b, 3)$ be the class of $3 \times b$ row-column designs for which the plot excess p and the concurrence excess q are both zero. If $d^* \in \mathcal{D}(v, b, 3)$ satisfies

(i) the column design of d^* is a BIBD, and

(ii) the row design of d^* is generalized binary,

then d^* is E-optimal and MV-optimal in $\mathcal{D}(v, b, 3)$.

Proof: If $r = bk/v$ is a multiple of 3 then d^* is a regular GYD, for which the result is already known. So only $r = \lambda(v-1)/2$ which is not a multiple of 3 need be considered. Condition (i) then implies that v is a multiple of 3, see (15) and (16) below. Write $r_1 = \lfloor r/3 \rfloor$ and $r_2 = r_1 + 1$.

Let $m_{di} = (m_{di1}, m_{di2}, m_{di3})$. For any equireplicated d which satisfies (ii), $m_{dil} \in \{r_1, r_2\}$ for every i and l, that is, m_{di} is some permutation of (r_1, r_1, r_2) or (r_1, r_2, r_2) as $r \equiv 1 \pmod 3$ or $r \equiv 2 \pmod 3$. In either case, there are exactly $\frac{v}{3}$ treatments corresponding to each of the three permutations, so that generalized binarity in rows induces a grouping of the treatments. By "treatments in group g" is meant those i for which m_{dig} is the distinct member of m_{di}, $g = 1, 2, 3$. For generalized binary d, $\mu_{dii'} = \mu_1 = \frac{r^2-1}{3}$ if i, i' are in different groups, and $\mu_{dii'} = \mu_1 + 1$ otherwise. Since the column design for d^* is a BIBD and d^* is generalized binary in rows, C_{d^*} has group divisble structure. In addition, B_{d^*} may be written in partitioned block form with matrices $\frac{2}{3}J_{\frac{v}{3},\frac{v}{3}}$ along the main diagonal and matrices $-\frac{1}{3}J_{\frac{v}{3},\frac{v}{3}}$ along the off-diagonal. It follows that $tr(B_{d^*}) = \frac{2v}{3}$ and $tr(B_{d^*}^2) = \frac{2v^2}{9}$. By Lemma 3, and the fact that

Table 2: Partition of blocks for CSDs with $t = 4w+2$ in (10). Blocks are displayed as rows for compactness. All integers are reduced mod $(12w + 6)$.

Subdesign	Blocks
S_1	blocks of a BIBD($k = 3$, $\lambda = 1$) on the $v - 1$ treatments ∞_2 and $1, 2, \ldots, 12w + 6$
S_2	$(j,\ 5w{+}3{-}i{+}j,\ 5w{+}2{+}i{+}j)$ and $(j,\ 3w{+}1{-}i{+}j,\ 3w{+}1{+}i{+}j)$ for $i = 1, 2, \ldots, w$ and $j = 1, 2, \ldots, 12w + 6$
S_3	$(\infty_1,\ i,\ 3w{+}1{+}i)$ for $i = 1, 2, \ldots, 12w + 6$, $(\infty_2,\ j,\ 6w{+}3{+}j)$ for $i = 1, 2, \ldots, 6w + 3$.
S_4	$(j,\ 4w{+}2{+}j,\ 8w{+}4{+}j)$ for $j = 1, 2, \ldots, 4w + 2$, and $(\infty_1, \infty_1, \infty_2)$

$r > 3$ (see (15) and (16) below),

$$z_{d^*1} \;=\; \frac{\lambda v}{3} - \frac{2}{\lambda(v-1)} \;=\; \frac{\lambda v}{3} - \frac{1}{r} \;>\; \frac{\lambda v - 1}{3}. \tag{11}$$

Calculation of pairwise variances $\frac{1}{\sigma^2} Var_{d^*}(\widehat{\tau_i - \tau_{i'}})$ with d^* follows easily from the generalized group divisible form of B_{d^*} (and hence of C_{d^*}). They take only two values: $var_1 = \frac{6}{\lambda v}$, which is the "same group" variance, and $var_2 = \frac{6}{\lambda v}[1 + \frac{3}{\lambda v b - v}]$, which is the "different group" variance. This gives

$$\Upsilon_{d^*} \;=\; \frac{6}{\lambda v}\left(1 + \frac{3}{\lambda v b - v}\right) \;=\; \frac{2}{\frac{\lambda v}{3} - \frac{\lambda v}{\lambda v b - (v-3)}} \;<\; \frac{2}{\frac{\lambda v - 1}{3}}. \tag{12}$$

Now consider any $d \in \mathcal{D}(v, b, k)$ that does not satisfy (i). If d has $r_{di} \leq r - 1$ for some i, then putting $p = 0$ in Lemma 4, there is $i' \neq i$ such that

$$\theta_{dii'} \;\leq\; \frac{2\lambda v}{3} - \frac{2v}{3(v-1)} \;<\; \frac{2(\lambda v - 1)}{3}. \tag{13}$$

With (13), it is immediate from Lemma 1(ii) and (11), and from Lemma 2 and (12), that d is E-inferior and MV-inferior to d^*.

Next, suppose d has $r_{di} = r$ for all i but is nonbinary in columns. By Lemma 5 with $p = 0$, there is $i \neq i'$ for which

$$\theta_{dii'} \;\leq\; \frac{2\lambda v}{3} - \frac{2}{3} - \frac{4}{3(v-1)} \;<\; \frac{2(\lambda v - 1)}{3}. \tag{14}$$

The bound (14) is the same as found in (13), so again Lemmas 1(ii) and 2 show d is E-inferior and MV-inferior to d^*.

So now suppose d is equireplicate and binary in columns, but has $\lambda_{dii'} < \lambda$ for some $i \neq i'$, say $\lambda_{d12} \leq \lambda - 1$. Now Lemma 6 with $p = 0$ says $\theta_{dii'} \;\leq\; \frac{2(\lambda v - 1)}{3}$ and yet again Lemmas 1(ii) and 2 show immediately that d is E-inferior and MV-inferior to d^*.

Now consider any d which satisfies (i) but not (ii). Since d is not generalized binary in rows,

$$a_{1d} = \sum_i (\mu_{dii} - \frac{r_{di}^2}{k}) \geq (2 + \sum_i \mu_{d^*ii}) - (\sum_i \frac{r_{d^*i}^2}{k}) = a_{1d^*} + 2 = \frac{2v}{3} + 2 = \frac{2(v+3)}{3}.$$

By Lemma 3,

$$z_{d1} = \frac{\lambda v}{3} - \frac{1}{2b}(a_{1d} + \sqrt{2a_{2d} - a_{1d}^2}) \leq \frac{\lambda v}{3} - \frac{1}{2b} a_{1d} \leq \frac{\lambda v}{3} - \frac{(v+3)}{3b} < \frac{\lambda v}{3} - \frac{1}{r} = z_{d^*1}.$$

Thus d^* is E-better than d.

To complete the proof, note that by virtue of (i), $\sum_{j=1}^{b} (n_{dij} - n_{di'j})^2 = 2(r - \lambda)$ for every $i \neq i'$. Lemma 7 with $p = 0$ says

$$\theta_{dii'} \;\leq\; 2r - \frac{1}{3b}[2b(r-\lambda)b + 18] \;=\; \frac{2b(2r+\lambda) - 18}{3b} \;= \frac{2(\lambda v b - 9)}{3b}$$

for some $i \neq i'$. By Lemma 2,

$$\Upsilon_d \;\geq\; \frac{6}{\lambda v}\left(\frac{\lambda v b}{\lambda v b - 9}\right) \;=\; \frac{6}{\lambda v}\left(1 + \frac{9}{\lambda v b - 9}\right) \;>\; \frac{6}{\lambda v}\left(1 + \frac{4}{\lambda v b - (v+2)}\right) \;=\; \Upsilon_{d^*}.$$

and therefore d^* is MV-better than d. □

The BIBDs with $k = 3$ for which r is not a multiple of 3 are those with parameters

$$v = 6t + 3,\ b = \frac{\lambda v(v-1)}{6},\ k = 3,\ r = \frac{\lambda(v-1)}{2},\ \lambda \equiv 1 \text{ or } 2 \pmod 3, \tag{15}$$

and those with parameters

$$v = 6t,\ b = \frac{\lambda v(v-1)}{6},\ k = 3,\ r = \frac{\lambda(v-1)}{2},\ \lambda \equiv 2 \text{ or } 4 \pmod 6, \tag{16}$$

and any $t \geq 1$. All of these designs are known (Hanani, 1961). Invoking Theorem 3.2 of Agrawal (1966) (also see Chai, 1998), they can all be arranged as row-column designs satifying condition (ii) of Theorem 3. Thus all produce row-column designs that are E-optimal and MV-optimal. Moreover, as can be seen in the proof, these designs are superior, with respect to both of these optimality criteria, to any design failing either condition (i) or (ii) of Theorem 3.

Example 3: The two designs shown here:

$$\begin{array}{cccccccccc} 1 & 1 & 3 & 3 & 4 & 2 & 2 & 5 & 5 & 6 \\ 4 & 2 & 1 & 1 & 6 & 3 & 3 & 6 & 4 & 5 \\ 2 & 5 & 5 & 6 & 1 & 4 & 6 & 2 & 3 & 4 \end{array} \quad \text{and} \quad \begin{array}{cccccccccccc} 1 & 1 & 5 & 2 & 9 & 6 & 8 & 3 & 3 & 2 & 4 & 7 \\ 9 & 4 & 1 & 6 & 7 & 4 & 2 & 8 & 5 & 3 & 5 & 6 \\ 3 & 7 & 8 & 1 & 5 & 9 & 9 & 4 & 6 & 7 & 2 & 8 \end{array}$$

are E-optimal and MV-optimal in $\mathcal{D}(6, 10, 3)$ and $\mathcal{D}(9, 12, 3)$, respectively.

4. Discussion

An often employed property in determining optimal designs is maximal trace of the information matrix over the class of competing designs. The designs shown here to be E-optimal and MV-optimal do not have this property. The upper bound for $tr(C_d)$, achieved by Youden type designs with binary column components, is $2b$. Designs d^* satisfying Theorems 1 and 2 have $tr(C_{d^*}) = 2b - 2(v+b)/3b$, and those for Theorem 3 have $tr(C_{d^*}) = 2b - (2v/3b)$. All designs satisfying any of Theorems 1 to 3 are superior to any competing, maximal trace design. Indeed, they are superior to any competing design having larger trace, be that maximal or not. This is a property, see Kiefer (1975), shared with the nonregular generalized Youden designs, which also enjoy optimality properties without being of Youden type. The Youden designs are, however, binary or generalized binary in both rows and columns, while the designs satisfying Theorems 1 and 2 have nonbinarity in columns.

Acknowledgements

The research contributions of Professor Aloke Dey have been deep and impactful. This article is offered in honor of the memory of Professor Dey, whose excellence as a researcher in experimental design was exceeded only by his generosity of spirit.

References

Agrawal, H. (1966). Some generalizations of distinct representatives with applications to statistical designs. *Annals of Mathematical Statistics*, **37**, 525–528.

Bailey, R. A. (2008). *Design of Comparative Experiments*, volume 25 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge.

Chai, F.-S. (1998). A note on generalization of distinct representatives. *Statistics & Probability Letters*, **39**(**2**), 173–177.

Das, A. and Dey, A. (1989). A generalization of distinct representatives and its applications. *Calcutta Statistical Association Bulletin*, **38**(**149-150**), 57–63.

Das, A. and Dey, A. (1990). Optimality of row-column designs. *Calcutta Statistical Association Bulletin*, **39**(**153-154**), 63–72.

Dean, A., Voss, D. and Draguljić, D. (2017). *Design and Analysis of Experiments,* Second Edition. Springer Texts in Statistics. Springer, Cham.

Hanani, H. (1961). The existence and construction of balanced incomplete block designs. *Annals of Mathematical Statistics*, **32**, 361–386.

Jacroux, M. (1982). Some E-optimal designs for the one-way and two-way elimination of heterogeneity. *Journal of the Royal Statistical Society, Series B*, **44**(**2**), 253–261.

Jacroux, M. (1983). Some minimum variance block designs for estimating treatment differences. *Journal of the Royal Statistical Society, Series B*, **45**(**1**), 70–76.

Kiefer, J. (1975). Construction and optimality of generalized Youden designs. In *A Survey of Statistical Design and Linear Models* (pp. 333–353). North-Holland, Amsterdam.

Morgan, J. P. (2015). Blocking with independent responses. In *Handbook of Design and Analysis of Experiments* (pp. 99–157). CRC Press, Boca Raton.

Morgan, J. P. and Srivastav, S. K. (2002). The completely symmetric designs with block size three. *Journal of Statistical Planning and Inference*, **106**(**1-2**), 21–30.

Morgan, J. P. and Uddin, N. (1995). Optimal, non-binary, variance balanced designs. *Statistica Sinica*, **5**(**2**), 535–546.

Shah, K. R. and Sinha, B. K. (1989). *Theory of Optimal Designs,* Lecture Notes in Statisitics, **54**. Springer-Verlag, New York.

Srivastav, S. K. and Morgan, J. P. (1998). Optimality of designs with generalized group divisible structure. *Journal of Statistical Planning and Inference*, **71**(**1-2**), 313–330.

APPENDIX

A. Proof of Lemma 3

Proof: From (3), the information matrix C_d for design d is

$$C_d = C_{d_N} - \frac{1}{b}B_d$$

where $B_d = M_d(I_3 - \frac{1}{3}J_3)M_d'$ is nonnegative definite. By assumption, C_{d_N} has $v-1$ eigenvalues equal to $\frac{\lambda v}{3}$. Since C_{d_N} is completely symmetric, and since B_d and C_{d_N} each have zero row sums, it follows that $C_d\mathbf{u} = \frac{\lambda v}{3}\mathbf{u} - \frac{1}{b}B_d\mathbf{u}$, where $\mathbf{u}$ is any eigenvector of B_d satisfying $\mathbf{u'1} = 0$. Thus the nonzero eigenvalues of C_d are $\frac{\lambda v}{3} - \frac{1}{b}\xi_i$ where $\xi_{d1} \geq \xi_{d2} \geq \ldots \geq \xi_{d,v-1}$ are the eigenvalues of B_d corresponding to eigenvectors $\mathbf{u}$ as identified above.

Next, since B_d is nonnegative definite, and since $r(B_d) \leq r(I_3 - \frac{1}{3}J_3) = 2$, it follows that $\xi_{di} = 0$ for $i \geq 3$.

Write $a_{1d} = tr(B_d) = \xi_{d1} + \xi_{d2}$ and $a_{2d} = tr(B_d^2) = \xi_{d1}^2 + \xi_{d2}^2$. Solving these two equations for the eigenvalues in terms of the trace terms gives

$$\xi_{d1} = \frac{1}{2}[a_{1d} + \sqrt{2a_{2d} - a_{1d}^2}] \quad \text{and} \quad \xi_{d2} = \frac{1}{2}[a_{1d} - \sqrt{2a_{2d} - a_{1d}^2}].$$

so that $z_{d1} = \frac{\lambda v}{3} - \frac{\xi_{d1}}{b}$ as claimed. □

B. Proof of Lemma 4

Proof: For any $d \in \mathcal{D}(v, b, k)$ (for any $k \geq 2$),

$$\begin{aligned} \theta_{dii'} \leq \theta_{d_N ii'} &= c_{d_N ii} + c_{d_N i'i'} - 2c_{d_N ii'} \\ &= r_{di} - \frac{\lambda_{dii}}{k} + r_{di'} - \frac{\lambda_{di'i'}}{k} + \frac{2\lambda_{dii'}}{k} \qquad (17) \\ &\leq \frac{k-1}{k}(r_{di} + r_{di'} + \frac{2\lambda_{dii'}}{k-1}), \qquad (18) \end{aligned}$$

the last inequality because $\lambda_{dii} \geq r_{di}$ and $\lambda_{di'i'} \geq r_{di'}$. Now write $\bar{\theta}_{di} = \sum_{i' \neq i} \theta_{dii'}/(v-1)$. Then under the conditions of the lemma, setting $k = 3$,

$$\begin{aligned} \bar{\theta}_{di} &\leq \frac{2}{3(v-1)} \sum_{i' \neq i} (r_{di} + r_{di'} + \lambda_{dii'}) \leq \frac{2}{3(v-1)}[(v-1)r_{di} + (vr + p - r_{di}) + 2r_{di}] \\ &\leq \frac{2}{3(v-1)}[v(r-1) + (vr+p)] = \frac{2}{3(v-1)}[\lambda v(v-1) + v(q-1) + p)]. \end{aligned}$$

Since for some $i' \neq i$, $\theta_{dii'} \leq \bar{\theta}_{di}$, the result is proven. □

C. Proof of Lemma 5

Proof: Suppose treatment i having $r_{di} = r$ is nonbinary in columns. So $\lambda_{dii} \geq r + 2$ and

$\sum_{i' \neq i} \lambda_{dii'} \leq r(k-1) - 2$. Then from (17),

$$\begin{aligned}
\sum_{i' \neq i}^{v} \theta_{dii'} &\leq \sum_{i' \neq i}^{v} [r_{di} - \frac{\lambda_{dii}}{k} + r_{di'} - \frac{\lambda_{di'i'}}{k} + \frac{2\lambda_{dii'}}{k}] \\
&\leq (v-1)r - \frac{(r+2)}{k} + (\frac{k-1}{k})[(v-1)r + p] + \frac{2[r(k-1)-2]}{k} \\
&= \frac{1}{k}\left\{2vr(k-1) - 2(v-1) + [p(k-1)-4]\right\}.
\end{aligned}$$

Since for some i', $\theta_{dii'} \leq \bar{\theta}_{di.} = \sum_{i' \neq i}^{v} \theta_{dii'}/(v-1)$, the result follows. □

D. Proof of Lemma 6

Proof: From (18),

$$\theta_{dii'} \leq \frac{k-1}{k}(r_{di} + r_{di'} + \frac{2\lambda_{dii'}}{k-1}) \leq \frac{2}{k}[r(k-1) + \lambda - 1].$$

□

E. Proof of Lemma 7

Proof: For any two treatments $i \neq i'$ having replication r, equation (5) says

$$\theta_{dii'} = 2r - \frac{1}{3b}[b\sum_{j=1}^{b}(n_{dij} - n_{di'j})^2 + 3\sum_{l=1}^{3}(m_{dil} - m_{di'l})^2].$$

It will be shown that for some such $i \neq i'$, $\sum_{l=1}^{3}(m_{dil} - m_{di'l})^2 \geq 6$, establishing the result. This is done in two cases, depending on the $r \pmod 3$ value.

Case 1: $r \equiv 1 \pmod 3$ and $p < (v+3)/2$.

Suppose treatment 1 is not generalized binary in rows. Then without loss of generality (WLOG), $m_{d11} \leq (r-4)/3$. There are two subcases to consider.

Case 1(a): $m_{di1} \geq (r+2)/3$ for some $i \leq v - p$. So take $m_{d21} \geq (r+2)/3$ and write $\overline{m}_{i(1)} = \sum_{l=2}^{3} m_{dil}/2$. We have

$$\overline{m}_{1(1)} \geq \frac{r - \frac{r-4}{3}}{2} = \frac{r+2}{3} \quad \text{and} \quad \overline{m}_{2(1)} \leq \frac{r - \frac{r+2}{3}}{2} = \frac{r-1}{3}$$

so that

$$\begin{aligned}
\sum_{l=1}^{3}(m_{d1l} - m_{d2l})^2 &\geq (m_{d11} - m_{d21})^2 + 2(\overline{m}_{1(1)} - \overline{m}_{2(1)})^2 \\
&\geq \left(\frac{(r-4)}{3} - \frac{(r+2)}{3}\right)^2 + 2\left(\frac{(r+2)}{3} - \frac{(r-1)}{3}\right)^2 = 6.
\end{aligned}$$

Case 1(b): $m_{di1} \le (r-1)/3$ for $i = 2, \ldots, v-p$. Then

$$\begin{aligned} \sum_{i=1}^{v} m_{di1} = b = \frac{vr+p}{3} &\le \frac{r-4}{3} + (v-p-1)\frac{r-1}{3} + \sum_{i=v-p+1}^{v} m_{di1} \\ \Rightarrow \sum_{i=v-p+1}^{v} m_{di1} &\ge \frac{v+pr+3}{3} \end{aligned}$$

Also, since treatments $i > v-p$ are generalized binary in rows, $\sum_{i=v-p+1}^{v} m_{di1} \le p(r+2)/3$. Hence we must have

$$p(r+2)/3 \ \ge \ \frac{v+pr+3}{3} \quad \Leftrightarrow \quad p \ \ge \ \frac{v+3}{2}$$

contradicting the requirement on p and thus showing that Case 1(b) cannot occur.

Case 2: $r \equiv 2 \pmod 3$

Suppose treatment 1 is not generalized binary in rows. Then WLOG $m_{d11} \ge (r+4)/3$. Again there are two subcases to consider.

Case 2(a): $m_{di1} \le (r-2)/3$ for some $i \le v-p$. So take $m_{d21} \le (r-2)/3$ and with the notation employed in Case 1(a),

$$\overline{m}_{1(1)} \ \le \ \frac{r - \frac{r+4}{3}}{2} \ = \ \frac{r-2}{3} \quad \text{and} \quad \overline{m}_{2(1)} \ \ge \ \frac{r - \frac{r-2}{3}}{2} \ = \ \frac{r+1}{3}$$

so that

$$\begin{aligned} \sum_{l=1}^{3} (m_{d1l} - m_{d2l})^2 &\ge (m_{d11} - m_{d21})^2 + 2(\overline{m}_{1(1)} - \overline{m}_{2(1)})^2 \\ &\ge \left(\frac{(r+4)}{3} - \frac{(r-2)}{3}\right)^2 + 2\left(\frac{(r-2)}{3} - \frac{(r+1)}{3}\right)^2 \ = \ 6. \end{aligned}$$

Case 2(b): $m_{di1} \ge (r+1)/3$ for $i = 2, \ldots, v-p$. Since treatments $i > v-p$ are generalized binary in rows, $m_{di1} = (r+1)/3$ for $i = v-p+1, \ldots, v$. Thus

$$\begin{aligned} 3\sum_{i=1}^{v} m_{di1} = bk = vr+p &\ge 3\left[\frac{r+4}{3} + (v-p-1)\frac{r+1}{3} + p\frac{(r+1)}{3}\right] \\ &= vr+p+3, \end{aligned}$$

a contradiction, showing that Case 2(b) cannot occur. □

F. Proof of Lemma 8

Proof: For any two treatments $i \ne i'$ having replication r and $r+1$ respectively, equation (5) says

$$\theta_{dii'} = 2r + 1 - \frac{1}{3b}[b\sum_{j=1}^{b}(n_{dij} - n_{di'j})^2 + 3\sum_{l=1}^{3}(m_{dil} - m_{di'l})^2 - 1].$$

It will be shown that for some such $i \neq i'$, $\sum_{l=1}^{3}(m_{dil} - m_{di'l})^2 \geq 5$, establishing the result. This is done in two cases, depending on the $r \pmod 3$ value.

Case 1: $r \equiv 1 \pmod 3$ and $p < (v+3)/2$.

Suppose treatment v is not generalized binary in rows. Then WLOG $m_{dv1} \geq (r+5)/3$. There are two subcases to consider.

Case 1(a): $m_{di1} \leq (r-1)/3$ for some $i \leq v-p$. So take $m_{d11} \leq (r-1)/3$ and write $\overline{m}_{i(1)} = \sum_{l=2}^{3} m_{dil}/2$. We have

$$\overline{m}_{1(1)} \;\geq\; \frac{r - \frac{r-1}{3}}{2} \;=\; \frac{2r+1}{6} \quad \text{and} \quad \overline{m}_{v(1)} \;\leq\; \frac{r - \frac{r+5}{3}}{2} \;=\; \frac{r-1}{3}$$

so that

$$\begin{aligned}\sum_{l=1}^{3}(m_{d1l} - m_{dvl})^2 \;&\geq\; (m_{d11} - m_{dv1})^2 + 2(\overline{m}_{1(1)} - \overline{m}_{v(1)})^2 \\ &\geq\; \left(\frac{(r-1)}{3} - \frac{(r+5)}{3}\right)^2 + 2\left(\frac{(r-1)}{3} - \frac{(2r+1)}{6}\right)^2 \;=\; 4\frac{1}{2},\end{aligned}$$

and hence the integer $\sum_{l=1}^{3}(m_{d1l} - m_{dvl})^2$ is at least 5.

Case 1(b): $m_{di1} \geq (r+2)/3$ for $i = 1, \ldots, v-p$. Here p must satisfy $p \geq 2$, for otherwise,

$$3\sum_{i=1}^{v} m_{di1} = bk = vr + p \;\;\geq\;\; 3\left[(v-1)\frac{r+2}{3} + \frac{(r+5)}{3}\right] \;=\; vr + 2v + 3,$$

a contradiction. It is now claimed that $m_{di1} \leq \frac{r-4}{3}$ for some $i \geq v-p+1$, for if not, employing the condition $p < (v+3)/2$,

$$\begin{aligned}0 \;=\; 3[(\sum_{i=1}^{v} m_{di1}) - b] \;&=\; 3\left[(v-p)\frac{r+2}{3} + (p-1)\frac{(r-1)}{3} + \frac{(r+5)}{3} - \frac{(vr+p)}{3}\right] \\ &=\; 2v - 4p + 6 \;>\; 2v - 4\frac{v+3}{2} + 6 \;=\; 0,\end{aligned}$$

another contradiction. Hence WLOG $m_{dv-1,1} \leq (r-4)/3$, so that with the same notation as in Case 1(a),

$$\overline{m}_{1(1)} \;\leq\; \frac{r - \frac{r+2}{3}}{2} \;=\; \frac{r-1}{3} \quad \text{and} \quad \overline{m}_{v-1(1)} \;\geq\; \frac{r+1 - \frac{r-4}{3}}{2} \;=\; \frac{2r+7}{6}.$$

Consequently,

$$\begin{aligned}\sum_{l=1}^{3}(m_{d1l} - m_{dv-1,l})^2 \;&\geq\; (m_{d11} - m_{dv-1,1})^2 + 2(\overline{m}_{1(1)} - \overline{m}_{v-1(1)})^2 \\ &\geq\; \left(\frac{(r+2)}{3} - \frac{(r-4)}{3}\right)^2 + 2\left(\frac{(r-1)}{3} - \frac{(2r+7)}{6}\right)^2 \;=\; 8\frac{1}{2}.\end{aligned}$$

Case 2: $r \equiv 2 \pmod 3$

Suppose treatment v is not generalized binary in rows. Then WLOG $m_{dv1} \geq (r+4)/3$. Again there are two subcases to consider.

Case 2(a): $m_{di1} \leq (r-2)/3$ for some $i \leq v-p$. So take $m_{d11} \leq (r-2)/3$ and with the notation employed in Case 1(a),

$$\overline{m}_{1(1)} \;\geq\; \frac{r-\frac{r-2}{3}}{2} \;=\; \frac{r+1}{3} \quad \text{and} \quad \overline{m}_{v(1)} \;\leq\; \frac{r+1-\frac{r+4}{3}}{2} \;=\; \frac{2r-1}{6}$$

so that

$$\begin{aligned}\sum_{l=1}^{3}(m_{d1l}-m_{dvl})^2 \;&\geq\; (m_{d11}-m_{dv1})^2+2(\overline{m}_{1(1)}-\overline{m}_{v(1)})^2\\ &\geq\; \left(\frac{(r-2)}{3}-\frac{(r+4)}{3}\right)^2+2\left(\frac{(r+1)}{3}-\frac{(2r-1)}{6}\right)^2 \;=\; 4\frac{1}{2}\end{aligned}$$

implying $\sum_{l=1}^{3}(m_{d1l}-m_{dvl})^2 \geq 5$.

Case 2(b): $m_{di1} \geq (r+1)/3$ for $i = 1, \ldots, v-p$. As in Case 1(b), p must satisfy $p \geq 2$, for otherwise

$$3\sum_{i=1}^{v} m_{di1} = bk = vr+p \;\;\geq\;\; 3\left[(v-1)\frac{r+1}{3}+\frac{(r+4)}{3}\right] \;=\; vr+v+3,$$

a contradiction. It is now claimed that $m_{di1} \leq \frac{r-5}{3}$ for some $i \in \{v-p+1, \ldots, v-1\}$, for if not, employing the condition $p < (v+6)/4$,

$$\begin{aligned}0 \;=\; 3[(\sum_{i=1}^{v} m_{di1})-b] \;&=\; 3\left[(v-p)\frac{r+1}{3}+(p-1)\frac{(r-2)}{3}+\frac{(r+4)}{3}-\frac{(vr+p)}{3}\right]\\ &=\; v-4p+6 \;>\; v-4\frac{v+6}{4}+6 \;=\; 0,\end{aligned}$$

another contradiction. Hence WLOG $m_{dv-1,1} \leq (r-5)/3$, so that with the same notation as in Case 1(a),

$$\overline{m}_{1(1)} \;\leq\; \frac{r-\frac{r+1}{3}}{2} \;=\; \frac{2r-1}{6} \quad \text{and} \quad \overline{m}_{v-1(1)} \;\geq\; \frac{r+1-\frac{r-5}{3}}{2} \;=\; \frac{r+4}{3}.$$

Consequently,

$$\begin{aligned}\sum_{l=1}^{3}(m_{d1l}-m_{dv-1,l})^2 \;&\geq\; (m_{d11}-m_{dv-1,1})^2+2(\overline{m}_{1(1)}-\overline{m}_{v-1(1)})^2\\ &\geq\; \left(\frac{(r+1)}{3}-\frac{(r-5)}{3}\right)^2+2\left(\frac{(2r-1)}{6}-\frac{(r+4)}{3}\right)^2 \;=\; 8\frac{1}{2}.\end{aligned}$$

□

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 277–286

An Algorithmic Approach to the Construction of Weighted A-optimal Balanced Treatment Incomplete Block Designs

Baidya Nath Mandal, Rajender Parsad and Sukanta Dash
ICAR-Indian Agricultural Statistics Research Institute
Library Avenue, New Delhi-110012

Received: 21 November 2020; Revised: 13 January 2021; Accepted: 23 February 2021

Abstract

The purpose of this paper is to present a heuristic algorithm for obtaining weighted A-optimal balanced treatment incomplete block (BTIB) designs for making test versus test and tests versus control comparisons. The proposed algorithm is implemented using R language. The proposed algorithm has been used to obtain weighted A-optimal BTIB designs in a restricted parametric range. A total of 369 weighted A-optimal BTIB designs are obtained in the restricted parametric range.

Key words: Algorithm; BTIB Designs; Linear Integer Programming; Test Treatment; Control; R package.

1. Introduction

There are many experimental situations where the experimenter is interested in comparing a set of new treatments, called tests, with a standard treatment, called control. In the presence of a single nuisance factor, block designs for tests versus control are recommended for conducting such experiments. A number of useful classes of designs such as balanced treatments incomplete block (BTIB) designs, group divisible treatment (GDT) designs, partially balanced treatment incomplete block (PBTIB) designs are available in literature and a lot of research efforts has been made on these designs. One can refer to Hedayat *et al.* (1988), Gupta and Parsad (2001) and Section 5.4 of Dey (2010) for a review on designs for tests versus control comparison.

Consider the experimental setting where v test treatments are to be compared with a control using a block design with b blocks each of size k. Let $D(v, b, k)$ denote the class of connected block designs in $v + 1$ treatments with b blocks with size k each. In the choice of A-optimal designs for tests versus controls comparisons, only comparison between test treatments and the control played the role in the choice of A-optimal designs. No consideration was made for pairwise comparison among test treatments. Since the designs in $D(v, b, k)$ are connected, they permit estimation of test versus test comparisons along with test versus control comparisons. Though, the pairwise comparisons among test versus test treatments would be required with lesser precision. To this end, Gupta *et al.* (1999) introduced weighted A-optimality of the block designs. Gupta *et al.* (1999) derived conditions

Corresponding Author: Baidya Nath Mandal
Email: bn.mandal@icar.gov.in

under which a design is weighted A-optimal for estimating these two sets of contrasts with unequal precisions. They provided a method of construction and a catalogue of weighted A-efficient BTIB designs. Parsad *et al.* (2009) proposed an algorithm based on interchange-exchange approach to obtain weighted A-efficient and weighted A-optimal designs for test versus test and tests versus control comparisons. They also obtained 15259 weighted A-efficient designs using the proposed algorithm. However, they reported only 43 weighted A-optimal BTIB designs in Table 3 of their article.

The purpose of this article is to present an algorithm for construction of weighted A-optimal BTIB designs and a list of 369 weighted A-optimal BTIB designs in a restricted parametric range. The article is organized as follows. Section 2 gives the concept of weighted A-optimality. An algorithm is proposed to obtain weighted A-optimal BTIB designs in section 3. The list of weighted A-optimal BTIB designs obtained using the proposed algorithm is presented in section 5. The article is concluded in section 6.

2. Preliminaries

Let the control treatment be indexed as 0 and the test treatments be denoted as $1, 2, ..., v$. Assume the two-way classified fixed effects homoscedastic model

$$y_{ijl} = \mu + \tau_i + \beta_j + \epsilon_{ijl} \tag{1}$$

where y_{ijl} denote the response from the lth experimental units in jth block receiving ith treatment, τ_i is the effect of ith treatment, β_j is the effect of jth block and ϵ_{ijl} are uncorrelated errors with mean zero and constant variance σ^2,$i = 0, 1, 2, ..., v$, $j = 1, 2, ..., b$ and $l = 1, 2, ..., k$. It may be mentioned here that Gupta *et al.* (1999) considered mixed effects model with random block effects. However, we shall restrict ourselves to fixed effects of blocks. A design $d \in D(v, b, k)$ is said to be weighted A-optimal if it minimizes

$$\beta \sum_{i=1}^{v} \text{var}\left(\hat{\tau}_{d0} - \hat{\tau}_{di}\right) + \alpha \sum_{i=1}^{v-1} \sum_{i'=i+1}^{v} \text{var}\left(\hat{\tau}_{di} - \hat{\tau}_{di'}\right)$$

with $\beta + \alpha = 1$ and $0 \leq \alpha, \beta \leq 1$. The expression above is the weighted sum of the variances of the estimates of test-control contrasts and test-test contrasts, respectively, with weights as β and α, respectively. Clearly, for $\alpha = 0$, the criterion reduces to A-optimality for tests vs controls and for $\alpha = \beta$, the criterion reduces to A-optimality for all pairwise comparisons. Since more precision is required for test-control comparisons than the test-test comparisons, β and α may be so chosen that $\beta > \alpha$.

Let $\mathbf{P}_c = [\mathbf{1}_v : -\mathbf{I}_v]$ and $\mathbf{P}_{T,Z} = [\mathbf{0}_Z : \mathbf{0}_{Z\times(v-Z-1)} : \mathbf{1}_Z : -\mathbf{I}_Z]$, $Z = 1, 2, ..., v-1$. Then

$$\mathbf{P} = \begin{pmatrix} \mathbf{P}_c \\ \mathbf{P}'_{T,1} \\ \mathbf{P}'_{T,2} \\ \vdots \\ \mathbf{P}'_{T,v-1} \end{pmatrix} \tag{2}$$

is the coefficient matrix of the contrasts for test-control and test-test comparisons.

To search for A-optimal block designs in $D(v, b, k)$, Gupta *et al.* (1999) focused their attention in the class of BTIB$(v, b, k; t, s)$ designs which was introduced by Stufken (1987). Gupta *et al.* (1999) presented the following result to characterize weighted A-optimality of BTIB designs.

Theorem 1: A BTIB$(v, b, k; t, s)$ design is A-optimal over $D(v, b, k)$ for fixed value of α if

$$g(t, s) = \min_{(x,z)\in\Delta} g(x, z)$$

where $\Delta = \{(x, z) : x = 0, 1, ..., int(k/2) - 1; z = 0, 1, ..., b \text{ with } z > 0 \text{ when } x = 0\}$ and $g(x, z) = \frac{(\beta+\alpha v)(v-1)^2}{A(x,z)} + \frac{\beta b}{B(x,z)}$,
$A(x, z) = k(v - 1)[b(k - x) - z] - [v\{b(k - x) - z\} - bk^2 - bx^2 - 2xz - z + 2k(bx + z)]$,
$B(x, z) = b[k(bx + z) - (bx^2 + 2xz + z)]$ and $\alpha/\beta \leq \frac{\{(2vk-2v-k+1)^2-(k-1)^2(v-1)^2\}}{v[(k-1)(v-1)]^2}$ when k is odd

and $\alpha/\beta \leq \frac{(2vk-2v-k)^2-k^2(v-1)^2}{v[k(v-1)]^2}$ when k is even.

Theorem 1 gives a sufficient condition to check weighted A-optimality of a given BTIB $(v, b, k; t, s)$ design and is useful to see whether a BTIB design is weighted A-optimal or not for given value of α. Gupta *et al.* (1999) used the result to check weighted A-optimality of designs from Parsad *et al.* (1995). The number of weighted A-optimal BTIB designs obtained by them are given below.

α	0	0.1	0.2	0.3	0.4
Number of designs	9	7	8	6	0

It is clear from above that more efforts are required to obtain weighted A-optimal BTIB designs. To this end, we present an algorithm to obtain weighted A-optimal BTIB designs.

3. The Algorithm

In this section, we present the algorithm to obtain weighted A-optimal BTIB designs in $D(v, b, k)$ for test-test and test-control comparisons. Given v, b, k, the algorithm computes the value of t and s which minimize $g(x, z)$ and then obtains other parameters through necessary parametric relations. Then it attempts to obtain the incidence matrix of a weighted A-optimal BTIB design with these parameters through linear integer programming approach.

The steps of the algorithm are detailed below.

Step 1: Given v, b, k and α, first check whether α satisfies the condition of Theorem 1. If α satisfies the condition of Theorem 1, obtain t and s which minimize $g(x, z)$.

Step 2: Compute $r_0 = s + bt, r = (bk - r_0)/v, \lambda_0 = (s(t + 1)(k - t - 1) + (b - s)t(k - t))/v$ and $\lambda_1 = (r(k - 1) - \lambda_0)/(v - 1)$. If all of r, r_0, λ and λ_0 are integers then proceed, else a weighted A-optimal BTIB does not exist.

Step 3: (i) Create the first row of the incidence matrix $\mathbf{N}$ by assigning $t+1$ in s randomly chosen columns and by assigning t in the remaining $b-s$ columns of the row. The first row of incidence matrix indicates the allocation of the control to b blocks.

(ii) Obtain the ith $(i = 2, 3, \ldots, v+1)$ row for allocation of $(i-1)$th test treatment to blocks by solving the following linear integer programming formulation with respect to *binary* decision variables $x_1, x_2, \ldots, x_b$:

$$\begin{aligned}
&\text{Maximize } \phi = \sum_{j=1}^{b} w_j x_j \\
&\text{subject to constraints} \\
&\sum_{j=1}^{b} x_j = r \\
&x_j \leq k - k_j \forall j = 1, 2, \ldots, b \\
&\sum_{j=1}^{b} n_{1j} x_j = \lambda_0 \\
&\sum_{j=1}^{b} n_{i'j} x_j = \lambda_1, \ \forall i' = 2, 3, \ldots, i-1
\end{aligned} \tag{3}$$

where $w_j = \frac{1}{k_j}$ if $k_j > 0$ and $w_j = 1$ if $k_j = 0$, with k_j being the size of the jth block up to $(i-1)$ row and $n_{i'j}$ is the element at the i'th row and the jth column of $\mathbf{N}$.

(iii) If there is no optimal solution of the formulation (3), delete a random row m between 2 to $(i-1)$th row of the incidence matrix, store the deleted row in a matrix $\mathbf{T}$, update k_j values and try to obtain a newer solution to mth row by solving the formulation (4):

$$\begin{aligned}
&\text{Maximize } \phi = \sum_{j=1}^{b} w_j x_j \\
&\text{subject to constraints} \\
&\sum_{j=1}^{b} x_j = r \\
&x_j \leq k - k_j \ \forall j = 1, 2, \ldots, b \\
&\sum_{j=1}^{b} n_{1j} x_j = \lambda_0 \\
&\sum_{j=1}^{b} n_{i'j} x_j = \lambda_1 \ \forall i' = 2, 3, \ldots, m-1, m+1, \ldots, i-1 \\
&\sum_{j=1}^{b} t_{uj} x_j < r \ \forall u = 1, 2, \ldots, p
\end{aligned} \tag{4}$$

where p is the number of rows of the matrix $\mathbf{T}$ and t_{uj} is the element at the

uth row and the jth column of $\mathbf{T}$ matrix. If there is a solution then update the incidence matrix. If there is no solution, repeat this step by drawing another random number m. Once the mth row is obtained, then go back to step ii) to obtain the ith row.

Step 4: If all the $v+1$ rows of the matrix $\mathbf{N}$ are obtained, then compute the A-efficiency by using the formula $A_e = \text{trace}(\mathbf{PC}_{d*}^{-}\mathbf{P}')/\text{trace}(\mathbf{PC}_{d}^{-}\mathbf{P}')$ to confirm weighted A-optimality of the design. Here d^* is a hypothetical A-optimal design in $D(v,b,k)$ for which $\text{trace}(\mathbf{PC}_{d*}^{-}\mathbf{P}')$ is minimum. If $A_e = 1$, then the design is weighted A-optimal.

The formulations (3) and (4) allocate a particular test treatment to r blocks out of the b blocks. While doing so, the objective function gives less weight to those blocks which already contains more number of treatments compared to other blocks. The first constraint ensures that the number of replications of the treatment is r. The second constraint is to ensure that a block does not contain more than k treatments. The third and fourth set of constraints ensure that for a given test treatment, the concurrences with the control and with the other test treatments are λ_0 and λ_1, respectively. The additional fifth constraint in formulation (4) prevents an already deleted solution for the mth row to recur.

Even if a weighted A-optimal BTIB design exists, sometimes the proposed algorithm may not be able to obtain a weighted A-optimal design. For example, the algorithm may get $u < v+1$ rows of incidence matrix $\mathbf{N}$ and it may not be able to proceed after uth row. This indicates that in these u rows, there may be some row(s) which do not allow the desired structure of the required design. Though step (iii) of Step 3 is there to eliminate such rows, however, it is not known which row(s) are actually the culprit and so step (iii) of Step 3 may not be 100% effective and this is the reason that the algorithm may not be able to get a solution even though a weighted optimal exists for the given parameters.

We have seen that the algorithm works best when $v \leq 30$ and $k \leq 10$. The efficiency of the algorithm to obtain weighted A-optimal design goes down with larger values of v. This is due to the fact that the chances of entering improper candidate rows in the incidence matrix increases with larger v. Further research efforts are required to obtain weighted A-optimal BTIB designs for larger values of v and k.

The integer programming formulations (3) and (4) were solved using lpSolve R package of Berkelaar and Others (2011) and the complete algorithm is implemented using R language. Further, an R package Aoptbdtvc (Mandal *et al.*, 2017a) has been built and published on CRAN. The package is available on cran.r-project.org/web/packages/Aoptbdtvc/index.html. A manual showing the usage of functions to implement the proposed algorithm is also available in the same web page.

4. Working of The Algorithm

In this Section, we illustrate the working of the algorithm with the help of an example.

Example 1: Consider construction of weighted A-optimal BTIB design for $v = 4, b = 4, k = 4, \alpha = 0.4$ The algorithm finds that $t = 0, s = 4$ in Step 1. From Step 2, algorithm gives $r_0 = 4, r = 3, \lambda_0 = 3, \lambda = 2$. Now in Step 3, the algorithm attempts to obtain an treatment-block incidence matrix of such a BTIB design with these parameters.

In the first step of Step 3, the algorithm obtains first row of the treatment-block incidence matrix as $\begin{pmatrix}1 & 1 & 1 & 1\end{pmatrix}$. To obtain the second row of the incidence matrix, following linear integer program is solved:

Maximize $\phi = x_1 + x_2 + x_3 + x_4$ subject to constraints

$$\begin{aligned}
&x_1 + x_2 + x_3 + x_4 = 3\\
&x_1 \leq 4 - 1\\
&x_2 \leq 4 - 1\\
&x_3 \leq 4 - 1\\
&x_4 \leq 4 - 1\\
&x_1 + x_2 + x_3 + x_4 = 3.
\end{aligned}$$

An optimal solution to the above linear program is $\begin{pmatrix}0 & 1 & 1 & 1\end{pmatrix}$. So after two steps, the incidence matrix obtained is $\begin{pmatrix}1 & 1 & 1 & 1\\ 0 & 1 & 1 & 1\end{pmatrix}$. To obtain the third row, the linear integer formulation is Maximize $\phi = x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 + \frac{1}{2}x_4$ subject to constraints

$$\begin{aligned}
&x_1 + x_2 + x_3 + x_4 = 3\\
&x_1 \leq 4 - 1\\
&x_2 \leq 4 - 2\\
&x_3 \leq 4 - 2\\
&x_4 \leq 4 - 2\\
&x_1 + x_2 + x_3 + x_4 = 3\\
&x_2 + x_3 + x_4 = 2.
\end{aligned}$$

An optimal solution to this formulation is $\begin{pmatrix}1 & 0 & 1 & 1\end{pmatrix}$ which gives incidence matrix up to third row as $\begin{pmatrix}1 & 1 & 1 & 1\\ 0 & 1 & 1 & 1\\ 1 & 0 & 1 & 1\end{pmatrix}$. For obtaining the fourth row, the formulation is Maximize $\phi = \frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{1}{3}x_3 + \frac{1}{3}x_4$ subject to constraints

$$\begin{aligned}
&x_1 + x_2 + x_3 + x_4 = 3\\
&x_1 \leq 4 - 2\\
&x_2 \leq 4 - 2\\
&x_3 \leq 4 - 3\\
&x_4 \leq 4 - 3\\
&x_1 + x_2 + x_3 + x_4 = 3\\
&x_2 + x_3 + x_4 = 2\\
&x_1 + x_3 + x_4 = 2.
\end{aligned}$$

The algorithm gives an optimal solution to this formulation as $\begin{pmatrix}1 & 1 & 0 & 1\end{pmatrix}$ and hence, the

incidence matrix obtained till fourth row is $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}$. For getting the last row of the incidence matrix, the formulation is as follows: Maximize $\phi = \frac{1}{3}x_1 + \frac{1}{3}x_2 + \frac{1}{4}x_3 + \frac{1}{4}x_4$ subject to constraints

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 3 \\ x_1 &\leq 4 - 3 \\ x_2 &\leq 4 - 3 \\ x_3 &\leq 4 - 3 \\ x_4 &\leq 4 - 4 \\ x_1 + x_2 + x_3 + x_4 &= 3 \\ x_2 + x_3 + x_4 &= 2 \\ x_1 + x_3 + x_4 &= 2 \\ x_1 + x_2 + x_4 &= 2. \end{aligned}$$

An optimal solution to this formulation is $\begin{pmatrix} 1 & 1 & 1 & 0 \end{pmatrix}$. As a result, the algorithm gives treatment-block incidence matrix as $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$ and the corresponding design is

$$\begin{array}{ll} \text{Block-1} & (0 \quad 2 \quad 3 \quad 4) \\ \text{Block-2} & (0 \quad 1 \quad 3 \quad 4) \\ \text{Block-3} & (0 \quad 1 \quad 2 \quad 4) \\ \text{Block-4} & (0 \quad 1 \quad 2 \quad 3) \end{array}$$

For confirmation of A-optimality, A_e is computed which is $A_e = \text{trace}(\mathbf{P}\mathbf{C}_{d*}^{-}\mathbf{P}')/\text{trace}(\mathbf{P}\mathbf{C}_{d}^{-}\mathbf{P}') = g(t,s)/\text{trace}(\mathbf{P}\mathbf{C}_{d}^{-}\mathbf{P}') = 3.2/3.2 = 1$. Thus, the design so obtained is weighted A-optimal BTIB design for $v = b = k = 4, \alpha = 0.4$.

Let us consider another example with $\alpha = 0.5$.

Example 2: Consider $v = 6, b = 7, k = 4, \alpha = 0.5$. In Step 1, it can be found that $t = 0, s = 4$ which gives $r = r_0 = 4, \lambda_0 = \lambda_1 = 2$. Step 3 gives us the following design.

$$\begin{array}{ll} \text{Block-1} & (2 \quad 3 \quad 4 \quad 6) \\ \text{Block-2} & (0 \quad 2 \quad 3 \quad 4) \\ \text{Block-3} & (0 \quad 2 \quad 5 \quad 6) \\ \text{Block-4} & (1 \quad 2 \quad 4 \quad 5) \\ \text{Block-5} & (0 \quad 1 \quad 4 \quad 6) \\ \text{Block-6} & (1 \quad 2 \quad 3 \quad 6) \\ \text{Block-7} & (0 \quad 1 \quad 3 \quad 5) \end{array}$$

Clearly, the design is a balanced incomplete block design and is A-optimal for all possible pair wise comparisons.

5. List of Weighted A-optimal BTIB Designs

The proposed algorithm in Section 3 can be used to construct weighted A-optimal BTIB designs for given parameters v, b, k, and α. We utilized the algorithm to obtain weighted A-optimal BTIB designs in a limited parametric range $2 \leq v \leq 30, v+1 \leq b \leq 50, 2 \leq k \leq min(10, v), \alpha = 0.2, 0.4, 0.6, 0.8$. We denote this parametric range as $\mathfrak{P}$ for further reference.

Within $\mathfrak{P}$, we obtained 369 A-optimal designs out of which 70 are R-type and 299 are S-type. The list of designs along with the layouts is available at `https://drs.icar.gov.in/WAoptBTIB/WAoptBTIB.htm`, (Mandal *et al.*, 2017b). The distribution of the designs according to various values of block size k and α is given in Table 1.

Table 1: Distribution of weighted A-optimal designs according to block size and α

	α				
Block size	0.2	0.4	0.6	0.8	Total Number of Designs
3	19	18	17	19	73
4	28	16	14	20	78
5	31	12	11	10	64
6	5	13	8	11	37
7	16	14	7	6	43
8	0	12	6	10	28
9	3	10	6	6	25
10	3	10	3	5	21
Total Number of Designs	105	105	72	87	369

We made a comparison of the weighted A-optimal designs obtained above with those of Gupta *et al.* (1999) and Parsad *et al.* (2009). Out of 15 distinct weighted A-optimal BTIB design given by Gupta *et al.* (1999), 7 fall in the parametric range $\mathfrak{P}$. Out of these 7 designs, we obtained six of them and are shown in Table 2. Out of the 43 designs reported by Parsad *et al.* (2009), only 14 designs fall in the parametric range $\mathfrak{P}$. We have obtained all these 14 designs and are given in Table 3.

Table 2: Weighted A-optimal BTIB designs in $\mathfrak{P}$ from Gupta *et al.* (1999)

v	b	k	t	s	α	Type
6	15	5	0	15	0.2	S
6	18	3	0	12	0.2	S
4	12	4	0	8	0.6	S
4	18	4	0	16	0.4	S
7	7	7	0	7	0.4	S
4	24	4	0	12	0.8	S
4	36	4	0	32	0.4	S

Table 3: Weighted A-optimal BTIB designs in $\mathfrak{P}$ from Parsad *et al.* (2009)

v	b	k	t	s	α	Type
3	3	3	0	3	0.2	S
3	4	3	0	3	0.4	S
6	7	3	0	3	0.4	S
9	18	3	0	9	0.2	S
4	4	4	0	4	0.2	S
4	5	4	0	4	0.4	S
5	5	5	0	5	0.2	S
5	5	5	0	5	0.4	S
7	7	5	0	7	0.2	S
7	7	5	0	7	0.2	S
9	12	7	0	12	0.2	S
8	8	8	0	8	0.4	S
9	9	9	0	9	0.4	S
10	10	10	0	10	0.4	S

An interesting observation is that among the 369 designs in the parametric range $\mathfrak{P}$, we found certain designs with same parametric combinations which are weighted A-optimal for more than one value of α. The list of those designs are depicted in Table 4.

Table 4: Weighted A-optimal designs for multiple values of α

Sr No.	v	b	k	t	s	α	Type
1	3	4	3	0	3	0.4, 0.6	S
2	4	4	4	0	4	0.2, 0.4	S
3	4	4	4	1	0	0.2, 0.4	R
4	4	5	4	0	4	0.4, 0.6	S
5	5	5	5	0	5	0.2, 0.4	S
6	5	5	5	1	0	0.2, 0.4	R
7	6	7	3	0	3	0.4, 0.6	S

6. Concluding Remarks

We have presented an algorithm to construct weighted A-optimal BTIB designs and also listed 369 weighted A-optimal designs. We believe most of the designs, particularly those with $\alpha > 0$ are new and has not been reported elsewhere. The proposed algorithm will be useful for experimenters and statisticians to obtain weighted A-optimal BTIB designs for various values of parameters including other values of weights given to the contrasts. Further efforts are required to devise algorithms which are able to construct weighted A-optimal BTIB designs for larger number of treatments. The proposed algorithm in this article has been restricted to construct weighted A-optimal BTIB designs where only one control is considered. The algorithm may be extended for weighted A-optimal block designs for more than one control treatment. Effort may also be made to obtain weighted A-optimal block designs beyond the class of BTIB designs.

Acknowledgements

Authors are thankful to the reviewer for careful pointing out of a mistake in an important expression in an earlier version of the manuscript. Authors duly acknowledge the support and guidance by the Chair Editor.

References

Berkelaar, M. and Others (2011). lpsolve: Interface to lp solve v. 5.5 to solve linear/integer programs. *R package version 5.6.6*.

Dey, A. (2010). *Incomplete Block Designs*, World Scientific.

Gupta, V. K. and Parsad, R. (2001). Block designs for comparing test treatments with control treatments-an overview. *Statistics and Applications*, **3(1 & 2)**, 133–146.

Gupta, V. K., Ramana, D. V. V. and Parsad, R. (1999). Weighted A-efficiency of block designs for making treatment-control and treatment-treatment comparisons. *Journal of Statistical Planning and Inference*, **77(2)**, 301–319.

Hedayat, A., Jacroux, M. and Majumdar, D. (1988). Optimal designs for comparing test treatments with controls. *Statistical Science*, **3(4)**, 462–476.

Mandal, B. N., Dash, S. and Parsad, R. (2017a). *Aoptbdtvc: A-Optimal Block Designs for Comparing Test Treatments with Controls*. R package version 0.0.2. https://CRAN.R-project.org/package=Aoptbdtvc.

Mandal, B. N., Parsad, R. and Dash, S. (2017b). Weighted A-optimal balanced treatment incomplete block designs : Design resources server. https://drs.icar.gov.in/WAoptBTIB/WAoptBTIB.htm.

Parsad, R., Gupta, V. K. and Prasad, N. S. G. (1995). On construction of A-efficient balanced test treatment incomplete block designs. *Utilitas Mathematica*, **47**, 185–190.

Parsad, R., Rathore, A. and Gupta, V. K. (2009). Computer aided construction of efficient designs for making treatment-treatment and treatment-control comparisons. *American Journal of Mathematical and Management Sciences*, **29(1-2)**, 201–228.

Stufken, J. (1987). A-optimal block designs for comparing test treatments with a control. *The Annals of Statistics*, **15(4)**, 1629–1638.

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 287–306

New Plans Orthogonal Through the Block Factor

Sunanda Bagchi*
1363, 10th cross, Kengeri Satellite Town, Bangalore 560060, India

Received: 16 November 2020 ; Revised: 16 February 2021; Accepted: 27 February 2021

Abstract

In the present paper we construct plans orthogonal through the block factor (POTBs). We describe procedures for adding blocks as well as factors to an initial plan and thus generate a bigger plan. Using these procedures we construct POTBs for symmetrical experiments with factors having three or more levels. We also construct plans inter-class orthogonal through the block factor for two-level as well as higher level factors.

Key words: Factorial experiments; Main effects; Orthogonal; Block factor.

1. Introduction

A situation in which a treatment factor is neither orthogonal nor confounded to a nuisance factor was first explored in Morgan and Uddin (1996) in the context of nested row-column designs. They derived a sufficient condition for a treatment factor, possibly non-orthogonal to the nuisance factors, to be orthogonal to another treatment factor. They also derived a sufficient condition for optimality and constructed several series of orthogonal main effect plans (OMEPs) satisfying optimality properties. Mukherjee, Dey and Chatterjee (2002) discussed and constructed main effect plans (MEPs) on small-sized blocks, not necessarily orthogonal to all treatment factors. Their plans also satisfy optimality properties. Optimal blocked MEPs of similar type are also constructed in Das and Dey (2004).

Bose and Bagchi (2007) provided plans satisfying properties similar to those of the plans of Mukherjee, Dey and Chatterjee (2002), but requiring fewer blocks. In Bagchi (2010) the concept of orthogonality through the block factor [see Definition 2.2] is introduced. In that paper it has been shown that a plan orthogonal through the block factor (POTB) may exist in a set up, where an OMEP can not exist. Making use of the Hadamard matrices in various way, Jacroux and his co-authors (2011, 2017) have come up with a number of such plans, mostly for two-level factors, many of them satisfying optimality properties. POTBs are also constructed in Saharay and Dutta (2016).

Wang (2004) constructed plans for two-level factors on blocks of size two, estimating interaction effects also. A general method of estimating interaction effects from plans on blocks of small size is provided in Bagchi (2020). The method is applied to plans for two and three-level factors estimating two-factor interactions.

Corresponding Author: Sunanda Bagchi
Email: sunandab355@gmail.com
*The author is retired from Indian Statistical Institute, Bangalore Center.

Preece (1966) constructed 'BIBDs for two sets of treatments'. Subsequently several authors constructed similar combinatorial objects. Among these, the ones relevant to the present paper are 'balanced Graeco-Latin block designs' of Seberry (1979), 'Graeco-Latin designs of type 1' of Street (1981) and 'Perfect Graeco-Latin balanced incomplete block designs (PERGOLAs)' of Rees and Preece (1999). We note that all these combinatorial designs are, in fact, two-factor POTBs satifying certain additional properties. We discuss these interesting combinatorial designs briefly in Section 3.

In the present paper our main objective is to provide plans in those set ups where no OMEP is available, accommodating as many factors as possible and deviating "as little as possible" from orthogonality. We construct a few series of POTBs for symmetrical experiment with factors having three or more levels. We also define plans inter-class orthogonal through the block factor (PIOTBs) [see Definition 6.1] and construct a series of such plans.

In Section 2 we present the definition of a POTB along with its attractive features. The later sections are devoted to construction. In Section 3 we obtain a few infinite series of POTBs for symmetric experiments with four or less factors, each with five or more levels [see Theorems 3.1, 3.2 and 3.3]. In Section 4 we describe methods of recursive construction. In Section 5 we use these methods and construct two series of POTBs for three-level factors on blocks of size four [see Theorems 5.1 and 5.3]. Finally, in Section 6 we construct two infinite series of PIOTBs with orthogonal classes of small size, one for factors with level ≥ 5 [see Theorem 6.1] and another for two-level factors [see Theorem 6.3]. Many of the plans constructed here are saturated. Moreover, many satisfy optimality properties.

2. Preliminaries

We shall consider main effect plans for a symmetrical experiment with m factors, laid out on blocks of constant size.

Notation 2.1: (0) m and n will denote integers. $M = \{1, 2, \cdots m\}$. Further, $\bar{i} = M \backslash \{i\},\ i \in M$.

(1) $\mathcal{P}$ will denote a main effect plan for a s^m experiment consisting of b blocks each of size k. n will denote the total number of runs. Thus, $n = bk$.

(2) s will denote an integer and S the set of integers modulo s. The set of levels for each factor is S unless stated otherwise. S^m will denote the following set of $m \times 1$ vectors. $S^m = \{(x_1, \cdots x_m)' : x_i \in S\}$.

(3) A_i denotes the ith factor, $i = 1, 2, \cdots m$. The vector $x = (x_1, x_2, \cdots x_m)' \in S^m$ represents a level combination or run, in which A_i is at level $x_i, i = 1, 2, \cdots m$.

(4) $\mathcal{B} = \{B_j,\ j = 1, \cdots b\}$ will denote the set of all blocks of $\mathcal{P}_0$, Thus, $B_j \subset S^m, |B_j| = k,\ 1 \leq j \leq b$. Sometimes we describe a plan in terms of its blocks.

(5) The replication vector of A_i is denoted by the $s \times 1$ vector r_i, the pth entry of which is the number of runs x of $\mathcal{P}$ such that $x_i = p, p \in S$. R_i denotes a diagonal matrix with diagonal entries same as those of r_i in the same order, $1 \leq i \leq m$.

(6) For $1 \leq i, j \leq m$, the A_i versus A_j incidence matrix is the $s_i \times s_j$ matrix N_{ij}.

The (p,q)th entry of this matrix is $N^{ij}(p,q)$, which is the number of runs x of $\mathcal{P}$ such that $x_i = p,\ x_j = q,\ p \in S_i,\ q \in S_j$. When $j = i, N_{ij} = R_i$.

(7) Fix $i \in M$.

(a) L_i will denote the A_i-versus block incidence matrix. Thus, the (p,j)th entry of the L_i is

$$L^i(p,j) = |x \in B_j : x_i = p|,\ p \in S, 1 \le j \le b.$$

(b) The $s \times 1$ vector α^i will denote the vector of unknown effects of A_i. The vector of raw totals for A_i will be denoted by T_i. B will denote the vector of block totals.

(c) We define the following vectors and matrices in terms of $\bar{i}$ [recall (0)]. The vector $\mathbf{U}_{\bar{i}}$ is of order $(m-1)s \times 1$, the matrices $\mathbf{V}_{\bar{i}}$ and $W_{\bar{i}}$ are of orders $s \times (m-1)s$ and $(m-1)s \times (m-1)s$ respectively

$$\mathbf{U}_{\bar{i}} = ((T_j))_{j \in \bar{i}}, \tag{1}$$
$$\mathbf{V}_{\bar{i}} = ((N_{ij}))_{j \in \bar{i}}, \tag{2}$$
$$W_{\bar{i}} = ((N_{pq}))_{p,q \in \bar{i}}. \tag{3}$$

Consider the normal equations for a plan $\mathcal{P}$ as described above. If we eliminate the general effects and the vector of block effects from this system of equations, we get the reduced normal equation for the vectors of all (unknown) effects of all the treatment factors. This is a system of ms equations, but it is convenient to view it as m systems of s equations each, the ith system equations is of the form

$$\sum_{j=1}^{m} C_{ij;B}\widehat{\alpha^j} = \mathbf{Q}_{i;B}. \tag{4}$$

Here $C_{ij;B},\ 1 \le j \le m$ are the coefficient matrices and $\mathbf{Q}_{i;B}$ is the vector of adjusted (for the blocks) totals for A_i. The expressions of them are as follows.

$$C_{ij;B} = N_{ij} - (1/k)L_iL_j' \text{ and } \mathbf{Q}_{i;B} = T_i - (1/k)L_iB. \tag{5}$$

For a fixed i, we can eliminate $\widehat{\alpha^j}, j \in \bar{i}$ from (4) and get

$$\text{the reduced normal equation for } \widehat{\alpha^i} \text{ as } C_{i;\bar{i}}\widehat{\alpha^i} = \mathbf{Q}_{i;\bar{i}}, \tag{6}$$

where $C_{i;\bar{i}} = R_i - \mathbf{V}_{\bar{i}}W_{\bar{i}}^{-}\mathbf{V}_{\bar{i}}'$ and $\mathbf{Q}_{i;\bar{i}} = T_i - \mathbf{V}_{\bar{i}}W_{\bar{i}}^{-}\mathbf{U}_{\bar{i}}'$, U_i, V_i and W_i are as in (1) and the equations next.

With this background we present a few definitions.

Definition 2.1: An m-factor MEP is said to be 'connected' if $Rank(C_{i;\bar{i}}) = s-1$, for every $i = 1, 2, \cdots m$.

Definition 2.2: [Bagchi (2010)] Fix $i \ne j\ 1 \le i, j \le m$. The factors A_i and A_j are said to be **orthogonal through the block factor (OTB)** if

$$kN_{ij} = L_i(L_j)'. \tag{7}$$

We denote this by $A_i \perp_{bl} A_j$.

A plan $\mathcal{P}$ is said to be a **plan orthogonal through the block factor (POTB)** if $A_i \perp_{bl} A_j$ for every pair $(i,j), i \neq j, i,j = 1, \cdots m$.

Remark 2.1: Condition (7) is equivalent to equation (7) of Morgan and Uddin (1996) in the context of nested row-column designs.

Let us try to see the implications of orthogonality through the block factor. Let $SS_{i;all}$ (respectively $SS_{i;B}$) denote sum of squares for A_i, adjusted for all other factors (respectively the block factor). The following results are known.

Theorem 2.1: Consider a plan $\mathcal{P}$. Fix $i \in \{1, \cdots m\}$.

(a) [Bagchi(2010)] If for $j \neq i$ $A_i \perp_{bl} A_j$, then

(i) $C_{ij;B} = 0$ and (ii) $Cov(l'\widehat{\alpha^i}, m'\widehat{\alpha^j}) = 0$, for $l'1_s = 0 = m'1_s$.

(b)[Bagchi (2020)] Further, $A_i \perp_{bl} A_j$, $\forall j \neq i$ is necessary and sufficient for the following.

(i) $C_{i;\bar{i}} = C_{ii;B}$ and (ii) $SS_{i;all} = SS_{i;B}$ with probability 1.

Discussion: Theorem 2.1 says the following about the inference on the factors of a connected main effect plan. The inference on a factor A_i depends only on the relationship between A_i and the block factor if and only if A_i is orthogonal to every other treatment factor through the block factor. Moreover, the data analysis of a POTB is very similar to the data analysis of a block design with s treatments.

It is well-known that the orthogonal MEP obtained from an orthogonal array is the best possible MEP in the sense that the estimates have the maximum precision among all MEPs in the same set up. The same cannot be said about an POTB since its performance also depends on the relationships of the treatment factors with the block factor. In the next theorem a guideline for the search for a 'good' POTB is provided. We omit the proof which can be obtained by going along the same lines as in the proofs of Lemma 1 and Theorem 1 of Mukherjee, Dey and Chatterjee (2002). [See Shah and Sinha (1989) for definitions, results and other details about standard optimality criteria]

Theorem 2.2: Suppose a connected POTB $\mathcal{P}^*$ satisfies the following condition. For a factor A_i and a non-increasing optimality criterion ϕ, L_i is the incidence matrix of a block design d which is ϕ -optimal in a certain class of connected block designs with s treatments and b blocks of size k each. Then, $\mathcal{P}^*$ is ϕ-optimal in a similar class of connected m-factor MEPs in the same set-up as $\mathcal{P}^*$ for the inference on A_i.

In particular, using the well-known optimality result of Kiefer (1975) we get the following result.

Corollary 2.1: Suppose $\mathcal{P}^*$ is a connected POTB. Fix $i \in \{1, \cdots m\}$.

If L_i is the incidence matrix of a BIBD, then, for the inference on A_i, $\mathcal{P}^*$ is universally optimal in the class of all m-factor connected MEP containing $\mathcal{P}^*$.

In view of the above result, we introduce the following term.

Definition 2.3: A connected POTB is said to be **balanced** if **each of its factors form a BIBD with the block factor**, that is L_i is the incidence matrix of a BIBD for each i, $1 \leq i \leq m$

We now present a small example of a balanced POTB on six blocks of size two each. It has two factors, each with four levels 0,1,2,3.

Example 1 (Bagchi and Bose, 2007) :

Blocks	$\rightarrow$	B_1		B_2		B_3		B_4		B_5		B_6	
Factors $\downarrow$	A_1	0	2	1	3	0	3	1	2	0	1	3	2
	A_2	1	3	0	2	2	1	3	0	3	2	0	1

3. Construction of Plans with a Small Number of Factors

We shall now proceed to construct POTBs for a symmetric experiment. Most of the constructions are of recursive type, in the sensethat from a given initial plan we generate a plan by adding blocks and/or factors. Before going to the actual constructions, we present a small POTB. This is not connected and therefore has no use in practice. However, its property will be used in proving the results below.

Lemma 3.1: Let $T = \{a, b, c, d\} \subset S$. Consider the following plan for a 4^2 experiment consisting of two blocks, in which the set of levels of A_1 is T, while that of A_2 is $-T$. Then, $A_1 \perp_{bl} A_2$.

Blocks	$\rightarrow$	B_1		B_2	
Factors $\downarrow$	A_1	a	c	b	d
	A_2	-d	-b	-a	-c

.

Proof is by direct verification of (7).

Definition 3.1: Consider an initial plan $\mathcal{P}_0$ for an s^m experiment as described in Notation 2.1. For $B \in \mathcal{B}$ and $v \in S^m$, $B+v$ will denote the following set of k runs. $B+v = \{x+v,\ x \in B\}$. Here $x + v = [x_i + v_i : 1 \leq i \leq m]'$, where the addition in each co-ordinate is modulo s.

By the plan **generated from $\mathcal{P}_0$ by adding** S we shall mean the plan (for the same experiment) having the set of blocks $\{B + u1_m : u \in S, B \in \mathcal{B}\}$. The new plan $\mathcal{P}$ will be denoted by $\mathcal{P}_0 \oplus S$.

We present an useful result. The simple proof is omitted.

Lemma 3.2: Consider a t^2 experiment, say E, such that the set of levels of A_i is T_i, which is a subset of S and $|T_i| = t, i = 1, 2$. If there is a plan $\mathcal{P}_0$ for E in which $A_1 \perp_{bl} A_2$, then $A_1 \perp_{bl} A_2$ in $\mathcal{P} = \mathcal{P}_0 \oplus S$ too ; although in $\mathcal{P}$ the set of levels of each factor is S. $\mathcal{P}$ is connected, even if $\mathcal{P}_0$ is not so.

We shall now proceed to construction. We begin with plans with a small set of factors. Let S^+ denote $S \cup \{\infty\}$. The following rule will define addition in S^+.

$$u + \infty = \infty = \infty + u,\ u \in S. \tag{8}$$

Theorem 3.1: Suppose s is an integer ≥ 5. Then connected POTBs with block size two exists for the following experiments.

(a) For an s^2 experiment a POTB $\mathcal{P}$ on $2s$ blocks exists. In the case $s = 5$, $\mathcal{P}$ is balanced.

(b) For an s^4 experiment a POTB on $4s$ blocks exists, whenever $s \geq 9$. If $s = 9$, $\mathcal{P}_2$ is balanced.

(c) A POTB $\mathcal{P}$ for a $(s+1)^4$ experiment with $6s$ blocks exists, whenever $n \geq 7$.

Proof: In each case, we present the blocks, of an initial plan $\mathcal{P}_0$ The required plan is $\mathcal{P}_0 \oplus S$ [see Definition 3.1]. Here a, b, c, d are distinct members of $S \setminus \{0\}$.

(a) The blocks of $\mathcal{P}_0$ are given below.

Blocks	$\rightarrow$	B_1		B_2	
Factors $\downarrow$	A_1	a	-a	b	-b
	A_2	b	-b	-a	a

.

We apply Lemma 3.1 with $c = -a, d = -b$ and see that $A_1 \perp_{bl} A_2$ in $\mathcal{P}_0$. Then we apply Lemma 3.2 and get the result.

If $s = 5$, taking $a = 1, b = 2$ we see that $L_1 = L_2$, which is the incidence matrix of the BIBD with parameters $v = 5, b = 10, r = 4, k = 2, \lambda = 1$. Hence the result.

(b) The blocks $B_l,\ l = 1, \cdots 4$ of $\mathcal{P}_0$ are as follows.

Blocks	$\rightarrow$	B_1		B_2		B_3		B_4	
Factors $\downarrow$	A_1	a	-a	b	-b	c	-c	-d	d
	A_2	b	-b	-a	a	-d	d	-c	c
	A_3	c	-c	d	-d	-a	a	b	-b
	A_4	d	-d	-c	c	b	-b	a	-a

To show that the generated plan is a POTB, we proceed as follows. We consider the following subplans.

Let $\mathcal{P}_{ij} = \{B_i, B_j\}\ i \neq j, i, j = 1, 2, 3, 4$. Further, let $\bar{\mathcal{P}}_{ij}$ denote the plan $\mathcal{P}_{kl}$, where $\{k, l\} = \{1, 2, 3, 4\} \setminus \{i, j\}$. [Thus, $\bar{\mathcal{P}}_{34} = \mathcal{P}_{12}$ and so on].

Applying Lemma 3.1 on these subplans we see that $A_i \perp_{bl} A_j$ in $\mathcal{P}_{ij}$ as well as in $\bar{\mathcal{P}}_{ij}$. Now the result follows from Lemma 3.2.

By taking $a = 1, b = 2, c = 3$ and $d = 4$ in the case $s = 9$, we get L_i is the incidence matrix of the BIBD with parameters $v = 9, b = 36, r = 8, k = 2, \lambda = 1$. Hence the result.

(c) The set of levels for each factor is S^+. The blocks $B_l,\ l = 1, \cdots 6$ of the initial plan are as follows.

Blocks	$\rightarrow$	B_1		B_2		B_3		B_4		B_5		B_6	
Factors $\downarrow$	A_1	0	∞	a	-a	b	-b	c	-c	a	-a	a	-a
	A_2	a	-a	0	∞	c	-c	-b	b	a	-a	-a	a
	A_3	b	-b	c	-c	0	∞	a	-a	-c	c	-c	c
	A_4	c	-c	-b	b	a	-a	0	∞	-c	c	c	-c

. □

By using arguments similar to that in Case (b), together with the relation (8) we get the result. □

We now list a few combinatorial structures in the literature which are actually balanced POTBs (for symmetrical or asymmetrical experiments).

(a) **Balanced Graco-Latin block design** defined and constructed in Seberry (1979) heve two factors.

(b) **Graco-Latin block design of type 1** of Street (1981) are also two-factor balanced POTBs satisfying $\mathbf{N}_{12} = J$.

(c) **Perfect Graeco-Latin balanced incomplete block designs (PERGOLAs)** defined and discussed extensively in Rees and Preece (1999) are two-factor balanced POTBs satisfying

$$\mathbf{N}_{12}\mathbf{N}'_{12} = \mathbf{N}'_{12}\mathbf{N}_{12} = fI_s + gJ_s, \quad \text{where f, g are integers.} \tag{9}$$

Here I_n is the identity matrix and J_n is the all-one matrix of order n.

(d) **Mutually orthogonal BIBDs** defined and constructed by Morgan and Uddin (1996) are multi-factor balanced POTBs.

Remark 3.1: The definition of neither balanced Graco-Latin block designs nor of mutually orthogonal BIBDs include condition (9). However, it is interesting to note that all these designs constructed so far do satisfy this condition. One would, therefore, suspect that this condition is implicit in the definition. We have, however, found a balanced POTB which does not satisfy this condition, as is shown below [see Theorem 3.2 (b)].

For the next construction we need a few notations. Recall that a multiset is a set in which the number of occurrences of an element is counted.

Notation 3.1: (i) $\bigsqcup$ denotes an union counting multiplicity.

(ii) For a set A and an integer n, nA denotes the multiset in which every member of A occurs n times.

(iii) For subsets A and B of a group $(G, +)$, A_B will denote the following multiset.

$$A - B = \{a - b : a \in A, b \in B\}.$$

Notation 3.2: Consider a plan $\mathcal{P}_0$ for an S^2 experiment as in Notation 2.1.

(a) Consider the jth block B_j of $\mathcal{P}_0$. $(x_{jt}, y_{jt})'$ will denote the tth run of B_j, $j = 1, \cdots b$. $X_j = \{x_{jt}, t = 1, \cdots k\}$ (respectively $Y_j = \{y_{jt}, t = 1, \cdots k\}$) is the set of levels of A_1 (respectively A_2) appearing in B_j.

(b) We define the following multisets.

$$D_j = \{y_{jt} - x_{jt}, t = 1, \cdots k\} \text{ and } D = \bigsqcup_{1 \le j \le b} D_j. \text{ Further, } \delta = \bigsqcup_{1 \le j \le b} (Y_j - X_j).$$

The following result can be proved by direct verification.

Lemma 3.3: (a) If an element u of S appears p times in D, then the following holds for $\mathcal{P}_0 \oplus S$.

$$N_{12}(i, j) = p \text{ if } j - i = u.$$

(b) A sufficient condition for $A_1 \perp_{bl} A_2$ in $\mathcal{P}_0 \oplus S$ is that $\delta = kD$ in $\mathcal{P}_0$.

Now we go to another construction.

Theorem 3.2: Let s be a positive integer ≥ 5. Then

(a) there exists a symmetric POTB $\mathcal{P}$ with three factors each having $s + 1$ levels on $b = 6s$ blocks of size two.

(b) In the case $s = 5$, we get a Balanced POTB. The restriction to any two of the factors reduces it to a PERGOLA, except that condition (9) is not satisfied.

Proof : Let S^+ be the set of levels for each factor. Consider an initial plan $\mathcal{P}_0$ with the set of factors $\{A_0, A_1, A_2\}$ and $\mathcal{B} = \{B_{ij},\ i = 1, 2,\ j = 0, 1, 2\}$, where B_{ij}' are as shown in the table below. The required plan $\mathcal{P} = \mathcal{P}_0 \oplus S$.

Blocks	$\rightarrow$	B_{10}		B_{11}		B_{11}		B_{20}		B_{21}		B_{22}	
Factors $\downarrow$	A_0	∞	0	-1	1	0	1	∞	0	1	2	0	2
	A_1	0	1	∞	0	-1	1	0	2	∞	0	1	2
	A_2	-1	1	0	1	∞	0	1	2	0	2	∞	0

.

(a) Consider a pair of factors of $\mathcal{P}_0$, say A_i, A_j. One can verify that the following are satisfied.

$$D = \{\infty^4, 0^2, 1^2, -1^2, 2, -2\} \text{ and } \delta = 2D \tag{10}$$

Here the superscript denote the multiplicity. In the computation we have used (8).

Now an application of Lemma 3.3 implies that $A_i \perp_{bl} A_j$ in $\mathcal{P}$. Since i, j are arbitrary, $\mathcal{P}$ is a POTB.

(b) From (10), in view of Lemma 3.3 (a) it follows that in $\mathcal{P}$

$$N_{pq} = \begin{bmatrix} 0 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 1 & 1 & 2 \\ 2 & 2 & 2 & 2 & 1 & 1 \\ 2 & 1 & 2 & 2 & 2 & 1 \\ 2 & 1 & 1 & 2 & 2 & 2 \\ 2 & 2 & 1 & 1 & 2 & 2 \end{bmatrix}, \; p \neq q, p, q = 0, 1, 2. \tag{11}$$

Moreover, by the construction,

$$L_i(L_i)' = 8I_6 + 2J_6, \; i = 0, 1, 2.$$

From the relation above, we see that each L_i is the incidence matrix of a BIBD with parameters $(v = 6, b = 30, r = 10, k = 2, \lambda = 2)$. Thus, by Definition 2.3 $\mathcal{P}$ is a balanced POTB. However, N_{pq} does not satisfy (9), $p \neq q, \; p, q = 0, 1, 2$. □

Next we construct a series of balanced POTBs using finite fields. We need the following notation.

Notation 3.3: (i) s is an odd prime power. $t = (s-1)/2$. F denotes the Galois field of order s. Further, $F^* = F \setminus \{0\}$ and $F^+ = F \cup \{\infty\}$.

(ii) α denotes a primitive element of F.

(iii) C_0 denotes the subgroup of order t of the multiplicative group of F and C_1 the coset of C_0. Thus, C_0 is the set of all non-zero squares of F, while C_1 is the set of all non-zero non-squares of F.

(iv) (i, j) denotes the number of ordered pairs of integers (k,l) such that the following equation is satisfied in F. [This notation is borrowed from the theory of cyclotomy]

$$1 + \alpha^k = \alpha^l, \; k \equiv i, l \equiv j \pmod 2.$$

We present the following well-known result for ready reference. [See equations (11.6.30), (11.6.40) and (11.6.43) of Hall (1986)].

Lemma 3.4: The difference between the cosets of F^* can be expressed in terms of the cyclotomy numbers as follows.

$$C_1 - C_0 = \bigcup_{k=0}^{1} (k, 1) C_k.$$

The following cyclotomy numbers are known.

Case 1: t odd. $(0, 0) = (1, 1) = (1, 0) = (t-1)/2, (0, 1) = (t+1)/2$.

Case 2: t even. $(0, 0) = t/2 - 1, (0, 1) = (1, 0) = (1, 1) = t/2$.

A series of two-factor balanced POTBs :

Theorem 3.3: Suppose s is an odd prime or a prime power. Then there exists a balanced POTB $\mathcal{P}^*$ for a $(s+1)^2$ experiment on $b = 2s$ blocks of size $(s+1)/2$.

Proof: The set of levels of each factor is F^+. We shall present the initial plan $\mathcal{P}_0$ consisting of a pair of blocks. The required POTB is $\mathcal{P}^* = \mathcal{P}_0 \oplus F$.

Let $\delta \in C_1$. Consider three $2 \times (t+1)$ arrays R^0, R^1 and R^2, the rows of which are indexed by $\{0,1\}$ and the columns by $C_0 \cup \{0\}$. The entries of the arrays are as given below.

$$R^0(1,0) = R^1(0,0) = R^2(0,0) = 0 \text{ and } R^0(0,0) = R^1(1,0) = R^2(1,0) = \infty. \tag{12}$$

$$\text{For } x = 0,1,\ y \in C_0,\ R^0(x,y) = \delta^x y, R^1(x,y) = \delta^{-x} y \text{ and } R^2(x,y) = \delta^{x-1} y. \tag{13}$$

For $i = 0,1,2$, let B_i be the block, the runs of which are the columns of R^i. When t is even, B_0 and B_1 constitute $\mathcal{P}_0$, while B_0 and B_2 constitute $\mathcal{P}_0$ when t is odd.

Clearly block size is $t+1 = (s+1)/2$. To show that $\mathcal{P}^*$ satisfies the required property, we have to show that

(a) $\mathcal{P}^*$ is a POTB and (b) each factor forms a BIBD with the block factor.

Condition (b) follows from the construction in view of Lemma 3.4. So, we prove (a). Let us write N for N_{12}. The rows and columns of N are indexed by F^+. From (12) and (3.9), we see that

$$N(ii) = 0,\ i \in F^+ \text{ and } N(\infty, i) = N(i, \infty) = 1,\ i \in F. \tag{14}$$

So, we assume $i \neq j \in F$. Let $u = j - i$. Then, $N(ij)$ is the number of times u appears in the multiset

$$\begin{cases} (\delta - 1)C_0 \bigsqcup (\delta^{-1} - 1)C_0 & \text{if t is even} \\ (\delta - 1)C_0 \bigsqcup (1 - \delta^{-1})C_0 & \text{if t is odd} \end{cases}$$

Since $-1 \in C_0$ if and only if t is even, $\delta^{-1} - 1$ is in the same coset as $\delta - 1$ if and only if t is odd. Therefore, the relations above together with (14) above imply that

$$N = J_{s+1} - I_{s+1}. \tag{15}$$

Now we take up $L_1 L_2' = H$ (say). From (12) and (3.9), we see that

$$H(ii) = 0,\ i \in F^+. \tag{16}$$

Further, for every $i \in F$, $H(\infty, i)$ is the replication number of i in the block design generated by the initial block $\{0\} \cup C_1$. Similarly, $H(i, \infty)$ is the replication number of i in the block design generated by the initial block $\{0\} \cup C_0$ if t is odd and $\{0\} \cup C_1$ otherwise. Thus,

$$H(\infty, i) = H(i, \infty) = t + 1,\ i \in F. \tag{17}$$

We, therefore, assume $i \neq j,\ i, j \in F$. Let $u = j - i$. Then, $H(ij)$ is the number of times u appears in the multiset

$$\begin{cases} ((\{0\} \cup C_1) - C_0) \bigsqcup (C_1 - (\{0\} \cup C_0)) & \text{if } t \text{ is even} \\ ((\{0\} \cup C_1) - C_0) \bigsqcup (C_0 - (\{0\} \cup C_1)) & \text{if } t \text{ is odd} \end{cases}$$

These relations, together with (16), (17) and Lemma 3.4 imply that $H = (t+1)(J_{s+1} - I_{s+1})$. Therefore, in view of (15), (7) follows and we are done. □

4. More on Recursive Construction

In this section we describe procedures for adding factors as well as blocks to an initial plan.

Notation 4.1: Consider a subset V of S^m. For every $i, 1 \leq i \leq m$, V_i will denote the following multiset of $|V|$ members of S. $V_i = \{v_i : v = (v_1, \cdots v_m)' \in V\}$.

Definition 4.1: Consider an initial plan $\mathcal{P}_0$ for an s^m experiment as described in Notation 2.1. Let V be as in Notation 4.1. By the plan $\mathcal{P}_0 + V$ **generated from** $\mathcal{P}_0$ **along** V we shall mean the plan (for the same experiment) having the set of blocks $\mathcal{B} + V = \{B + v : v \in V, B \in \mathcal{B}\}$, where $B + v$ is as in Definition 3.1. Usually, V will contain the 0-vector, so that the blocks of $\mathcal{P}_0$ will also be blocks of $\mathcal{P}_0 + V$.

The next lemma provides a few sufficient conditions on $\mathcal{P}_0$ and V so that a given pair of factors are orthogonal through the block factor in $\mathcal{P}_0 + V$. The proof is by direct verification.

Remark 4.1: In an initial plan, say $\mathcal{P}_0$, one or more levels of one or more factors may be absent. $\mathcal{P}_0$ may still be a POTB if (7) holds (with one or more row/column of N'_{ij}s being null vectors) for every unordered pair of (i, j). In such cases one has to choose V such that all levels of all factors do appear in $\mathcal{P}_0 + V$.

Lemma 4.1: Consider an initial plan $\mathcal{P}_0$ for an s^2 experiment and a subset V of S^2. The following conditions on $\mathcal{P}_0$ and V are sufficient for $\mathcal{P}_0 + V$ to be a POTB.

(a) In $\mathcal{P}_0$ all the levels of the first factor appear and $V = \{(0, i),\ i \in S\}$.

(b) $\mathcal{P}_0$ is arbitrary and $V = \{(i, j),\ i, j \in S\}$.

(c) $\mathcal{P}_0$ has a pair of blocks B_0, B_1 each of size 2, as described below. Let $i \neq j,\ k \neq l \in S$. Let $x_0 = (i, i)'$, $y_0 = (j, j)$, $x_1 = (k, l)'$ and $y_1 = (l, k)'$. B_i consists of runs x_i and y_i, $i = 0, 1$. $V = (u, u),\ u \in S$.

(d) $\mathcal{P}_0$ is a POTB in which with one or more levels of one or both factors may be absent. V is such that every member of S appears at least once in each $V_i, i = 1, 2$.

Our next procedure enlarges the set of factors of a given plan, while keeping the number of blocks fixed.

Definition 4.2: (a) Consider a plan $\mathcal{P}$ as in Notation 2.1. Suppose there is another plan $\mathcal{P}'$ having b blocks of size k each. We shall combine these two plans to get another one with a larger set of factors.

Let x_{ij} (respectively $\tilde{x}_{ij}$) denote the jth run in the ith block of $\mathcal{P}$ (respectively $\mathcal{P}'$), $1 \leq j \leq k,\ 1 \leq i \leq b$. Let $y_{ij} = \begin{bmatrix} x_{ij} & \tilde{x}_{ij} \end{bmatrix}'$, $1 \leq j \leq k,\ 1 \leq i \leq b$. Then, the plan on b blocks of size k with y_{ij} as the jth run in the ith block, $1 \leq j \leq k,\ 1 \leq i \leq b$ is said

to obtained by joining the factors of $\mathcal{P}$ and $\mathcal{P}'$ together. The new plan will be denoted by $\left[\begin{array}{cc} \mathcal{P} & \mathcal{P}' \end{array}\right]$.

(b) In case $\mathcal{P}'$ is a copy of $\mathcal{P}$ then $\left[\begin{array}{cc} \mathcal{P} & \mathcal{P}' \end{array}\right]$ is denoted by $\mathcal{P}^2$. For $t \geq 3$, the plan $\mathcal{P}^t$ is defined in the same way.

We find it convenient to name the factors of plan $\mathcal{P}$ in a manner different from Notation 2.1. The names the factors of $\mathcal{P}$ and its power $\mathcal{P}^t$ will be as in the notation below.

Notation 4.2: Consider a plan $\mathcal{P}$ having the set of factors $\mathcal{F}_0 = \{A, B, \cdots H\}$. The set of factors of $\mathcal{P}^t$ will be named as

$$\mathcal{F} = \bigcup_{i=1}^{t} \mathcal{F}_i, \text{where } \mathcal{F}_i = \{A_i, B_i, \cdots H_i\}.$$

Combining Definitions 4.1 and 4.2 we get a recursive construction described below.

Definition 4.3: Consider an initial plan $\mathcal{P}_0$ for an s^m experiment laid on b blocks of size k each. Consider a $p \times q$ array $H = ((h_{ij}))_{1 \leq i \leq p, 1 \leq j \leq q}$. We now obtain a plan for an s^{mq} experiment on bp blocks of size k using the array H as follows. We first obtain $\mathcal{P}_0^q$ following Definition 4.2.

Let $v_i = \left[\begin{array}{cccc} h_{i1}.1_t' & h_{i2}.1_t' & \cdots & h_{iq}.1_t' \end{array}\right]'$, $1 \leq i \leq p$ and $V_H = \{v_i,\ 1 \leq i \leq p\}$.

Our required plan $\mathcal{P}$ is $\mathcal{P}_0^q + V_H$ and it will be denoted by $H \diamond \mathcal{P}$. Symbolically,

$$\mathcal{P} = H \diamond \mathcal{P}_0 = \mathcal{P}_0^q + V_H. \tag{18}$$

Our task is to find a suitable array H so that the plan $H \diamond \mathcal{P}_0$ satisfies certain desirable properties. A natural choice for H is an orthogonal array of strength 2. We shall use a modification of such an orthogonal array so as to accommodate a few more factors.

Definition 4.4: (Rao, 1946) Let $m, N, t \geq 2$ be integers and s is an integer ≥ 2. Then an orthogonal array of strength t is an $m \times N$ array, with the entries from a set S of s symbols satisfying the following. All the s^t t-tuples with symbols from S appear equally often as columns in every $t \times N$ subarray. Such an array is denoted by $OA(N, m, s, t)$.

Notation 4.3: (a) The set of symbols of an $OA(N, m, s, 2)$ is assumed to be the set of integers modulo s.

(b) The array obtained by adding a column of all zeros (in the mth position, say) to an $OA(N, m-1, s, 2)$ will be denoted by $Q(N, m, s)$.

Exploring the properties of an orthogonal array of strength 2, we get the following result from the recursive construction described in Definition 4.3.

Theorem 4.1: Consider a plan $\mathcal{P}_0$ for an s^t experiment on b blocks of size k each. If an $OA(N, m-1, s, 2)$ exists, then $\exists$ a plan $\mathcal{P}$ with a set of s^{mt} factors on bN blocks of size k each with the following properties. Here the factors of $\mathcal{P}_0^q$ as well as $\mathcal{P}$ are named according to Notation 4.2. Let $P \neq Q, P, Q \in \mathcal{F}_0$.

(a) $P_i \perp_{bl} Q_j,\ i \neq j, 1 \leq i, j \leq m.$

(b) $P_i \perp_{bl} Q_i$ for every $i,\ 1 \leq i \leq m$, if and only if $P \perp_{bl} Q$ in $\mathcal{P}_0$.

Proof: By assumption $Q = Q(N, m, s)$ exists. The required plan $\mathcal{P}$ is $Q \diamond \mathcal{P}_0$. Property (a) follows from the construction while (b) follows from (b) of Lemma 4.1. □

Remark 4.2: Table 1 of Rees and Preece (1999) presents a number of examples of two-factor POTBs (which they call PERGOLAs). An Application of Theorem 4.1 on each of them would yield a balanced POTB for a larger set of factors.

Finally, we describe a procedure of modifying the sets of levels of factors. Specifically, given a pair of plans with the same number of factors and the same block size, we obtain a plan by merging the sets of levels of the corresponding factors of the given plans.

Definition 4.5: Consider a pair of plans $\mathcal{P}_1$ and $\mathcal{P}_2$ each having t factors and blocks of size k. Let S_i denote the set of levels of each factor of $\mathcal{P}_i,\ s_i = |S_i|,\ i = 1, 2$. We assume that $S_1 \neq S_2$. Let $U = S_1 \cup S_2$ and $u = |U|$. The plan consisting of all the blocks of $\mathcal{P}_1$ and $\mathcal{P}_2$ taken together will be viewed as a plan, say $\mathcal{P}_1 \cup \mathcal{P}_2$, for an u^t experiment in the following sense.

(a) Each factor of $\mathcal{P}_1 \cup \mathcal{P}_2$ will have U as the set of levels.

(b) Fix $p \in U$. Let $\mathcal{R}_p^{ij}$ denote the set of runs of $\mathcal{P}_j$, in which the level p of the ith factor appears, $j = 1, 2, 1 \leq i \leq t$. [Needless to mention that $\mathcal{R}_p^{ij} = \phi$ if p is not in S_j.] Then, the level p of the ith factor of $\mathcal{P}_1 \cup \mathcal{P}_2$ appears in exactly the runs in $\mathcal{R}_p^{i1} \sqcup \mathcal{R}_p^{i2}, 1 \leq i \leq t$.

Remark 4.3: From Definition 4.5 we see that for $p \in U$, the replication number of level p of the ith factor of $\mathcal{P}_1 \cup \mathcal{P}_2$ is $r^{i1}(p) + r^{i2}(p)$, where $r^{ij}(p)$ is the replication number of level p of the ith factor of $\mathcal{P}_j$.

For instance, in Theorem 5.1 below, Definition 4.5 is used to construct the plan $\mathcal{P}_h$ by merging the corresponding factors of $\mathcal{P}_{1h}$ and $\mathcal{P}_{2h}$. There, $S_1 = \{0, 1\}$, while $S_2 = \{0, 2\}$. Thus, while both $\mathcal{P}_{1h}$ and $\mathcal{P}_{2h}$ are equireplicate, the replication number of level 0 of each factor of $\mathcal{P}_h$ is double of that of the levels 1 and 2 of the same factor.

The following result is an immediate consequence of Definition 4.5 .

Lemma 4.2: Consider a pair of connected plans $\mathcal{P}_1$ and $\mathcal{P}_2$, as in Definition 4.5 (recall Definition 2.1). Then, we can say the following about the plan $\mathcal{P} = \mathcal{P}_1 \cup \mathcal{P}_2$.

(a) If both $\mathcal{P}_1$ and $\mathcal{P}_2$ are POTB, then so is $\mathcal{P}$.

(b) $\mathcal{P}$ is connected, if and only if $S_1 \cap S_2 \neq \phi$.

5. Construction of POTBs for Three-level Factors

In this section we make use of the tools described in Section 4 to generate plans for three-level factors. The factors of the initial and final plans are named in accordance with Notation 4.2.

Theorem 5.1: If h is the order of a Hadamard matrix, then there exists a connected and saturated POTB $\mathcal{P}_h$ for a 3^{3h} experiment in $2h$ blocks of size 4 each.

Proof: Let $O_4 = OA(4, 3, 2, 2)$ with $S = \{0, 1\}$. Let $\mathcal{P}_0$ be the plan consisting of a single block consisting of the four columns of O_4 as runs. Thus, $\mathcal{P}_0$ is an OMEP for a 2^3 experiment.

Using the relation between orthogonal arrays of strength two and Hadamard matrices, [see Theorem 7.5 in Hedayat, Sloane and Stuffken (1999), for instance], we see that a $Q(h, h, 2)$ exists, in view of the hypothesis.

Let $\mathcal{P}_{1h} = Q(h, h, 2)\diamond\mathcal{P}_0$ and $\mathcal{P}_{2h}$ be obtained from $\mathcal{P}_{1h}$ by replacing level 1 of every factor by the level 2. Next we construct our required plan $\mathcal{P}_h = \mathcal{P}_{1h} \cup \mathcal{P}_{2h}$ by using Definition 4.5. By construction $\mathcal{P}_h$ has $2h$ blocks of size 4 each.

We now show that $\mathcal{P}_h$ is a POTB. We note that by Theorem 4.1, each of $\mathcal{P}_{1h}$ and $\mathcal{P}_{2h}$ is a POTB for a 2^{3h} experiment on h blocks of size 4 each. The sets of levels of each factor of them are $\{0, 1\}$ and $\{0, 2\}$ respectively. It follows from Lemma 4.2 that $\mathcal{P}_h$ is a connected POTB for an experiment with $3h$ factors, the set of levels of each factor being $\{0, 1, 2\}$. Since the available degrees of freedom for the treatment factors is $6h$ which is the same as the required degrees of freedom, the plan is saturated. □

We now take $h = 2$ and present the plan $\mathcal{P}_2$ for a 3^6 experiment on four blocks of size four each.

Table 1 : The plan $\mathcal{P}_2$

Blocks	$\rightarrow$	B_{01}	B_{02}	B_{11}	B_{12}
Factors $\downarrow$	A_1	00 11	00 11	00 22	00 22
	B_1	01 01	01 01	02 02	02 02
	C_1	01 10	01 10	02 20	02 20
	A_2	00 11	11 00	00 22	22 00
	B_2	01 01	10 10	02 02	20 20
	C_2	01 10	10 01	02 20	20 02

For the next construction we need some more notations.

Notation 5.1: O_4 is as in the proof of Theorem 5.1. T_4 will denote the array obtained from O_4 by replacing each 1 by 2 and $\tilde{T}_4$ the array obtained from T_4 by interchanging 0 and 2.

Theorem 5.2: A POTB for a 3^3 experiment on two blocks of size four exists.

Proof: Let $B_{10} = O_4$, $B_{20} = T_4$ and $B_{02} = \tilde{T}_4$. The set of columns of each of these 3×4 array constitutes an OMEP for a 2^3 experiment, the set of levels of factors being $\{0, 1\}$ for B_{10}, while $\{0, 2\}$ for the other two.

Let ρ_1 (respectively ρ_2) denote the plan consisting of the pair of blocks B_{10}, B_{20} (respectively B_{10}, B_{02}). By Lemma 4.2, each of ρ_1 and ρ_2 is a POTB for a 3^3 experiment. □

Using the pair of plans constructed above, we generate a bigger plan.

Theorem 5.3: (a) If there exists an $OA(N, m, 3, 2)$, then there exists a connected POTB $\mathcal{P}_m$ for a $3^{3(2m+1)}$ experiment in $2N$ blocks of size 4 each.

In particular $\mathcal{P}_m$ is saturated whenever $N = 3^n$ and $m = (3^{n-1} - 1)/2$, for an integer $n \geq 2$.

(b) There exists a connected POTB $\mathcal{P}_1$ for a 3^9 experiment in 6 blocks of size 4 each.

Proof: Part (a): Let the factors of ρ_1 and ρ_2 be named as A, B, C and $\tilde{A}, \tilde{B}, \tilde{C}$ respectively. Let $O = OA(N, m, 3, 2)$ and $Q = Q(N, m, 3)$. We now use Definition 4.5 to generate bigger plans $\mathcal{P}_{1m}$ and $\mathcal{P}_{2m}$ as follows.

$$\mathcal{P}_{1m} = Q \diamond \rho_1 \text{ and } \mathcal{P}_{2m} = O \diamond \rho_2.$$

Clearly, $\mathcal{P}_{1m}$ and $\mathcal{P}_{2m}$ are plans for $3^{3(m+1)}$ and 3^{3m} experiments respectively, each on $2N$ blocks of size 4. Following Notation 4.2, we name of the factors of these plans as follows.

$$\begin{aligned} \text{The factors of } \mathcal{P}_{1m} \text{ are} \quad & A_0, B_0, C_0, A_1, B_1, C_1, \cdots A_m, B_m, C_m \\ \text{and the factors of } \mathcal{P}_{2m} \text{ are} \quad & \tilde{A}_1, \tilde{B}_1, \tilde{C}_1, \cdots \tilde{A}_m, \tilde{B}_m, \tilde{C}_m. \end{aligned}$$

Now we combine the factors of $\mathcal{P}_{1m}$ and $\mathcal{P}_{2m}$ following Definition 4.2 (a) and thus obtain our required plan $\mathcal{P}_m$. Symbolically,

$$\mathcal{P}_m = \left[\begin{array}{cc} \mathcal{P}_{1m} & \mathcal{P}_{2m} \end{array} \right].$$

By construction, $\mathcal{P}_m$ is a plan for $2m+1$ three-level factors on $2N$ blocks of size 4 each. We shall now show that it is a POTB.

Theorems 4.1 and 5.2 imply that each one of $\mathcal{P}_{1m}$ and $\mathcal{P}_{2m}$ is a POTB. Therefore, if we show the following relation, then we are done.

$$P_i \perp_{bl} \tilde{Q}_j, \ P, Q \in \{A, B, C\}, \ i \in I \cup \{0\}, \ j \in I, \text{ where } I = \{1, \cdots m\}. \tag{19}$$

To show this relation, we fix P_i and $\tilde{Q}_j$ as above.

Case 1. $i, j \in I$ **:** Since ρ_1 and ρ_2 are POTBs, (19) follows from Lemma 4.1 (d), whenever $Q \neq P$. Again, (c) of the same Lemma proves (19) for the case $Q = P$.

Case 2. $i = 0, \ j \in I$ **:** We take P_0 as the first and $\tilde{Q}_j$ as the second factor. Then applying Lemma 4.1 (a) we get (19).

Hence the proof of the first part is complete.

To prove the second part, we see that $\mathcal{P}_m$ is saturated when $N = 2m + 1$. Now Rao (1946) has shown that an $OA(s^n, (s^n - 1)/(s - 1), s, 2)$ exists whenever $n \geq 2$. (see Theorem 3.20 of Hedayat, Sloane and Stufken (1999) for instance). Putting $s = 3$, we get the result.

Proof of (b) : Let $O = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$. and $Q = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$. Now the construction for the plan, say $\mathcal{P}_1$, is just like that in Case (a). The verification is also exactly like the same in Case (a) with $I = \{1\}$. □

We now present the plan of Theorem 5.3 (b).

Table 2 : The plan $\mathcal{P}_1$

Blocks	$\rightarrow$	B_{10}	B_{20}	B_{11}	B_{21}	B_{12}	B_{22}
Factors $\downarrow$	A_0	00 11	00 22	00 11	00 22	00 11	00 22
	B_0	01 01	02 02	01 01	02 02	01 01	02 02
	C_0	01 10	02 20	01 10	02 20	01 10	02 20
	A_1	00 11	00 22	11 22	11 00	22 00	22 11
	B_1	01 01	02 02	12 12	10 10	20 20	21 21
	C_1	01 10	02 20	12 21	10 01	20 02	21 12
	$\tilde{A}_1$	00 11	22 00	11 22	00 11	22 00	11 22
	$\tilde{B}_1$	01 01	20 20	12 12	01 01	20 20	12 12
	$\tilde{C}_1$	01 10	20 02	12 21	01 10	20 02	12 21

6. Inter-class Orthogonal Plans

Inter-class orthogonal plans are defined in Bagchi (2019) in the context of plans without any blocking factor. Here we extend the definition to the present context - the orthogonality being through the block factor.

Definition 6.1: Let us consider a plan ρ. Suppose the set of all factors of ρ can be divided into several classes in such a way that if two factors belong to different classes, then they are orthogonal through the block factor. Such a plan ρ is called a **"Plan Inter-class Orthogonal through the Blocks (PIOTB)"** and the classes will be referred to as "orthogonal classes".

We shall now proceed towards the construction of PIOTBs.

Theorem 6.1: If s is an integer ≥ 5, then a PIOTBs with $4s$ blocks of size two exists for an s^4 experiment.

Proof: The blocks $B_l,\ l = 1, \cdots 4$ of an initial plan $\mathcal{P}_0$ are presented below. Here a, b are distinct members of $S \setminus \{0\}$. The required plan $\mathcal{P}$ is $\mathcal{P}_0 \oplus S$ [see Definition 3.1].

Blocks	$\rightarrow$	B_1		B_2		B_3		B_4	
Factors $\downarrow$	A_1	0	a	a	-a	0	b	-b	b
	A_2	a	-a	0	-a	-b	b	0	-b
	A_3	0	-b	b	-b	-a	0	a	-a
	A_4	b	-b	0	b	a	-a	a	0

We define at subplans $\mathcal{P}_{ij}, i, j = 1, 2, 3, 4$ in the same way as in the proof of Theorem 3.1 (b). Using Lemma 3.1 we see that in each of the subplans $\mathcal{P}_{12}$ and $\mathcal{P}_{34}$ $A_1 \perp_{bl} A_2$ and $A_3 \perp_{bl} A_4$. Similarly, in each of the subplans $\mathcal{P}_{13}$ and $\mathcal{P}_{24}$ $A_1 \perp_{bl} A_3$ and $A_2 \perp_{bl} A_4$. Therefore, by Lemma 3.2 $A_i \perp_{bl} A_j$ in $\mathcal{P}$, where $i \in \{1, 4\}$ and $j \in \{2, 3\}$. However, A_1 is not orthogonal to A_4 and A_2 is not orthogonal to A_3. It, therefore, follows that $\mathcal{P}$ is interclass orthogonal with classes $\{A_1, A_4\}$ and $A_2, A_3\}$. □

Next we present a PIOTB for three-level factors.

Theorem 6.2: A saturated PIOTB exists for a 3^6 experiment on four blocks of size four each.

Proof: Consider the following plan $\mathcal{P}$. It is easy to see that it is a PIOTB with non-orthgonal classes $\{P_1, P_2\},\ P = A, B, C$.

Table 1 : Plan $\mathcal{P}$

Blocks	$\rightarrow$	B_1	B_2	B_3	B_4
Factors $\downarrow$	A_1	00 12	00 21	00 12	00 21
	B_1	01 02	02 01	10 20	20 10
	C_1	01 20	02 10	02 10	01 20
	A_2	01 01	02 02	01 01	02 02
	B_2	01 10	02 20	10 01	20 02
	C_2	00 11	00 22	11 00	22 00

□

Remark 6.1: A POTB for a 4^4 experiment on 4 blocks of size 4 is well-known [can be obtained by treating a row of OA(16,5,4,2) as the block factor]. By collapsing two of the levels of each factor to one level one gets a POTB for a 3^4 experiment on the same set up. Allowing non-orthogonality we have been able to accommodate two more three-level factors, making it saturated.

Theorem 6.3: Suppose Hadamard matrices of orders m and n exist. Then, there exists a saturated PIOTB $\mathcal{P}_{(m,n)}$ for a 2^{mn} experiment on n blocks of size $m + 1$ each. There are n orthogonal classes of size m each.

Proof: By hypothesis $Q_m = Q(m, m, 2)$ exists. Let R be the $m \times m + 1$ array obtained by juxtaposing a column of all-ones to Q_m. Let $\mathcal{P}_0$ be the plan for a 2^m experiment on a single block consisting of $m + 1$ runs, which are the columns of R. Let us name the factors of $\mathcal{P}_0$ as $A, B, \cdots H\}$. Note that the column added to Q_m saves H from being confounded with the block. However, this column also destroys the relation of orthogonality enjoyed by the factors in Q_m (represented by its rows).

By hypothesis, $Q_n = Q(n, n, 2)$ exists. Let $\mathcal{P}_{(m,n)} = Q_n \diamond \mathcal{P}_0$. Clearly, $\mathcal{P}_{(m,n)}$ is an main effect plan for a 2^{mn} experiment with parameters as in the statement. By construction, no factor is confounded with the block factor. Using Theorem 4.1 and the property of $\mathcal{P}_0$, we

see that $\mathcal{P}_n$ is interclass orthogonal with orthogonal classes $\{A_i, B_i, \cdots H_i\},\ 1 \leq i \leq n$ (recall Notation 4.2). Hence the result. □

We now present the plan $\mathcal{P}_{(4,2)}$.

Table 6.2 : The plan $\mathcal{P}_{4,2}$

Blocks	$\rightarrow$	B_1	B_1
Factors $\downarrow$	A_1	00 11 1	11 00 0
	B_1	01 01 1	10 10 0
	C_1	01 10 1	10 01 0
	D_1	00 00 1	11 11 0
	A_1	00 11 1	00 11 1
	B_1	01 01 1	01 01 1
	C_1	01 10 1	01 10 1
	D_1	00 00 1	00 00 1

There are two orthogonal classes, which are $\{A_i, B_i, C_i, D_i\}, i = 1, 2$.

A situation when a PIOTB is better than a POTB.

Since in a POTB the factors are orthogonal to each other, one expects that its performance would be better than that of a PIOTB, in which orthogonality does not hold for one or more pairs of factors. However, we find an example, where a PIOTB is better than a POTB w.r.t. the commonly used optimality criteria.

Consider the following three plans for two-level factors on the set up with two blocks of size 5 each.

$\mathcal{P}_1$ is a plan for four factors constructed in Jacroux and Kealy-Dichone (2015). A POTB having seven factors is constructed in Bagchi and Bagchi (2020). Let $\mathcal{P}_2$ denote the POTB obtained by deleting the last factor of it. Let $\mathcal{P}_3$ denote the plan obtained from $\mathcal{P}_{4,2}$ by deleting the factors D_1 and D_2.

We compare these plans in terms of the C-matrix for the contrasts. For the definition of this C-matrix we refer to Jacroux and Kealy-Dichone (2015). We find the following.

Theorem 6.4: (a) $\mathcal{P}_3$ is A- and D-better than $\mathcal{P}_2$.

(b) Let $\tilde{\mathcal{P}}_2$ denote the restriction of $\mathcal{P}_2$ to four factors. Let $\tilde{\mathcal{P}}_3$ be the plan obtained from $\mathcal{P}_{4,2}$ consisting of the factors A_1, B_1, A_2, B_2. Then, $\tilde{\mathcal{P}}_3$ is A- and D-better than $\mathcal{P}_1$ as well as $\tilde{\mathcal{P}}_2$.

(c) All the three plans are E-optimal.

Proof: We compute the C-matrix for the contrasts of each plan and find the following.

$$C_{\mathcal{P}_1} = 8I_4 + (8/5)J_4 \quad \text{and} \quad C_{\mathcal{P}_2} = 8I_6.$$

$$C_{\mathcal{P}_3} = \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}, \quad \text{where} \quad A = 8I_3 + (8/5)J_3.$$

The spectrums of these matrices are as follows.

$$\begin{aligned} C_{\mathcal{P}_1}: &\quad 8^3(8+32/5)^1, \\ C_{\mathcal{P}_2}: &\quad 8^5, \\ C_{\mathcal{P}_3}: &\quad 8^4(8+24/5)^2 \\ \text{and } C_{\tilde{\mathcal{P}}_3}: &\quad 8^2(8+16/5)^2. \end{aligned}$$

The rest is by simple computation. □

Conjecture : An universally optimal plan can not exist in the above set up and $\mathcal{P}_3$ is optimal w.r.t. a large class of optimality criteria.

References

Bagchi, S. (2010). Main effect plans orthogonal through the block factor. *Technometrics*, **52**, 243–249.

Bagchi, S. (2019). Inter-class orthogonal main effect plans for asymmetrical experiments. *Sankhya B*, **81**, 93–122.

Bagchi, S. (2020). Nearly Resolution V Plans on Blocks of Small Size. *Journal of Statistical Theory and Practice*, **14(1)**, https://doi.org/10.1007/s42519-019-0074-3.

Bagchi, S. and Bagchi, B. (2020). Aspects of optimality of plans orthogonal through other factors. *Submitted*.

Bose, M. and Bagchi, B. (2007). Optimal main effect plans in blocks of small size. *Statistics and Probability Letters*, **77**, 142–147.

Chen, X. P., Lin, J. G., Yang, J. F. and Wang, H. X. (2015). Construction of main effects plans orthogonal through the block factor, *Statistics and Probability Letters*, **106**, 58–64.

Das, A. and Dey, A. (2004). Optimal main effect plans with nonorthogonal blocks. *Sankhya*, **66**, 378–384.

Hall, M. (1986). *Combinatorial Theory*. Wiley-interscince, New York.

Hedayat, A. S., Sloan, N. J. A. and Stufken, J. (1999). *Orthogonal Arrays, Theory and Applications*. Springer Series in Statistics.

Jacroux, M. (2011). On the D-optimality of orthogonal and nonorthogonal blocked main effects plans. *Statistics and Probability Letters*, **81**, 116–120.

Jacroux, M. (2011). On the D-optimality of nonorthogonal blocked main effects plans. *Sankhya B* , **73**, 62–69.

Jacroux, M. (2013). A note on the optimality of 2-level main effects plans in blocks of odd size. *Statistics and Probability Letters*, **83**, 1163–1166.

Jacroux, M. and Kealy-Dichone, B. (2014). On the E-optimality of blocked main effects plans when $n \equiv 3 \pmod 4$. *Statistics and Probability Letters*, **87**, 143–148.

Jacroux, M. and Kealy-Dichone, B. (2015). On the E-optimality of blocked main effects plans when $n \equiv 2 \pmod 4$. *Sankhya B*, **77**, 165–174.

Jacroux, M. and Jacroux, T. (2016). On the E-optimality of blocked main effects plans when $n \equiv 1 \pmod 4$. *Communications in Statistics - Theory and Methods*, **45**, 5584–5589.

Jacroux, M. and Kealy-Dichone, B. (2017). On the E-optimality of blocked main effects plans in blocks of different sizes. *Communications in Statistics - Theory and Methods*, **46**, 2132–2138.

Kiefer, J. (1975). Construction and optimality of generalized Youden designs. In: Srivastava, J.N. (ed) *A survey of statistical design and linear models*. North-Holland, Amsterdam, 333 – 353.

Morgan, J. P. and Uddin, N. (1996). Optimal blocked main effect plans with nested rows and columns and related designs. *Annals of Statistics*, **24**, 1185–1208.

Mukerjee, R., Dey, A. and Chatterjee, K. (2002). Optimal main effect plans with non-orthogonal blocking. *Biometrika*, **89**, 225–229.

Preece, D. A. (1966). Some balanced incomplete block designs for two sets of treatment. *Biometrika*, **53**, 497–506.

Rees, D. H. and Preece, D. A. (1999). Perfect Graeco-Latin balanced incomplete block designs. *Discrete Mathematics*, **197/198**, 691–712.

Rao, C. R. (1946). On Hypercubes of strength d and a system of confounding in factorial experiments. *Bulletin of Calcutta. Mathematical Society*, **38**, 67-78.

SahaRay, R. and Dutta, G. (2016). On the Optimality of Blocked Main Effects Plans. *International Scholarly and Scientific Research and Innovation*, **10**, 583–586.

Seberry, J. (1979). A note on orthogonal Graeco-Latin designs. *Ars Combinatoria*, **8**, 85–94.

Shah, K. R. and Sinha, B. K. (1989). *Theory of Optimal Designs*, Lecture notes in Statistics, 54, Springer-Verlag, Berlin.

Street, D. J. (1981). *Graeco-Latin and Nested Row and Column Designs*. Combinatorial Mathematics VIII, Lecture notes in Mathematics, 884, 304–313.

Takeuchi, K. (1961). *On The Optimality of Certain Type of PBIB Designs*. Reports of statistical application research, Union of Japanese Scientists and Engineers, 8, 140–145.

Wang, P. C. (2004). Designing two-level fractional factorial experiments in blocks of size two. *Sankhya*, **66**, 327–342.

Statistics and Applications {ISSN 2454-7395(online)}
Volume 19, No. 1, 2021 (New Series), pp 307–317

On Iterative Analysis of Orthogonal Saturated Factorial Designs

Daniel T. Voss
Department of Mathematics and Statistics
Wright State University, Dayton, Ohio, USA

Received: 25 November 2020; Revised: 11 February 2021; Accepted: 18 February 2021

Abstract

Orthogonal saturated factorial designs are useful for screening a few important factors from many. Independent effect estimates can be normalized to have common variance but, with no independent estimate of variability, tests are based on the comparison of larger estimates to smaller ones under an assumption of effect sparsity. Early methods of analysis were proposed by Daniel (1959) and Birnbaum (1959), and subsequent work by Zahn (1975ab) rekindled interest in the problem. They each suggested methods to be applied iteratively, but justifications are generally empirical.

Analytical results establishing control of error rates remain limited. Voss (1988, 1999), Holm, Mark, and Adolfsson (2005), and Voss and Wang (2006a) provided a class of closed step-down tests shown to be of family-wise size-α, utilizing the order statistics of the normalized estimates or the corresponding sums of squares. These are non-iterative tests, utilizing the effect estimates as k order statistics, comparing the ith largest estimate to a critical value based on the distribution of the ith largest of k estimates. However, the step-down tests would be more powerful if conducted iteratively—namely, testing the effect with the ith largest estimate using a critical value based on the largest of i estimates, rather than the ith largest of k estimates. Iterative tests also require the tabulation of fewer critical values.

In this paper, simulations are used to support the conjecture that certain iterative step-down tests for analysis of orthogonal saturated designs do strongly control the family-wise error rate. Some insight is also garnered to guide efforts for an analytical proof.

Key words: Closed test; Family-wise error rate; Iterative testing; Screening experiment.

AMS Subject Classifications: Primary 62K15, 62F07, 62F03; secondary 62F35, 62L10.

1. Introduction

Orthogonal saturated factorial designs are useful for screening a few important factors from many. Such designs yield independent effect estimates that can be normalized to have common variance. Such designs provide no independent estimate of variability, but larger

Correponding Author: Daniel T. Voss
Email: dan.voss@wright.edu

estimates can be compared to smaller estimates, so a standard premise for the analyses is an assumption of *effect sparsity*—namely, that only a few of the effects under study are substantial.

Daniel (1959) introduced the use of half-normal probability plots for the graphical analysis of the normalized effect estimates, and he proposed a corresponding testing procedure. Birnbaum (1959) also proposed tests for the analysis of such designs. They each suggested that their proposed methods be applied iteratively—namely, given k effects and corresponding estimates, if the effect with largest normalized estimate is asserted nonzero, the remaining $k-1$ effect estimates are then analyzed as if said largest estimate was never part of the data, and this process is iterated till the largest remaining estimate is not significantly nonzero. Zahn (1975ab) proposed and evaluated several variations on the iterative analyses of Daniel (1959), including revised test statistics and critical values, comparing methods empirically. His work renewed interest in the analysis of orthogonal saturated factorial designs, but establishing control of error rates remained an open problem.

Many methods of analysis of orthogonal saturated factorial designs have been proposed over the years, but most authors have relied on simulation studies to justify the methods. Progress on analytic justification has been slow, despite long interest in the problem. Voss (1988) proposed a family-wise size-α step-down test of the effects based on the order statistics of the normalized estimates. Subsequently, Voss (1999) provided a rigorous proof that said test strongly controls the family-wise size of the test, characterizing the procedure as a closed, step-down test (see Marcus, Peritz, and Gabriel, 1976), the proof utilizing an obscure stochastic ordering lemma of Alam and Rizvi (1966) and Mahamunulu (1967). Holm, Mark, and Adolfsson (2005) and Voss and Wang (2006a) also provided step-down tests strongly controlling error rates. For reviews of methods of analysis of orthogonal factorial designs, see Hamada and Balakrishnan (1998) and Voss and Wang (2006b).

To make our discussion concrete, consider here the step-down test procedure and statistics utilized by Voss (1988). The test statistics are

$$ss_{(i)}/qmse, \quad i = 1, \ldots, k,$$

for independent, normalized effect estimators $\hat{\theta}_i \sim N(\theta_i, \sigma)$, where $ss_{(i)} = \hat{\theta}^2_{(i)}$ are the corresponding order statistics of the sums of squares, and where $qmse = \sum_{i=1}^{\nu} ss_{(i)}/\nu$ is the *quasi mean squared error* obtained as the average of the ν smallest sums of squares, for specified ν.

For the *step-down test* proposed by Voss (1988), one asserts $\theta_h \neq 0$ if $\hat{\theta}_h = \hat{\theta}_{(j)}$ and $ss_{(i)}/qmse > c(\alpha, i, k)$ for all $i = j, ..., k$. In other words, if $ss_{(15)}/qmse > c(\alpha, 15, 15)$, then the effect corresponding to the largest order statistic $\hat{\theta}_{(15)}$ is asserted to be nonzero and one continues; else one stops. If $ss_{(15)}/qmse > c(\alpha, 15, 15)$ and $ss_{(14)}/qmse > c(\alpha, 14, 15)$, then one also asserts the effect corresponding to the second-largest order statistic to be nonzero and one continues; else one stops. The test procedure continues stepping down in this manner, testing each order statistic's effect in turn starting with the largest and stepping down, continuing as long as an assertion is made.

Voss (1999) showed that the above test procedure is a *closed step-down test* and strongly controls the family-wise error rate to be α if one uses critical value $c(\alpha, i, k)$ such that

$$P(SS_{(i)}/QMSE) > c(\alpha, i, k)) = \alpha, \tag{1}$$

for $SS_{(i)}/QMSE$ (i=1,...,k) the order statistics of the test statistics under the complete null distribution—namely, assuming all effects are zero. A test *strongly controls* the family-wise error rate to be at most α if the probability of any false assertions is at most α for all parameter configurations.

However, many authors have proposed that such step-down tests be conducted iteratively in the following sense. As the step-down test is conducted, if an effect is asserted to be nonzero, then the test proceeds as if that effect where never considered. For example, suppose $k = 15$ effects are initially analyzed. If the effect with largest absolute estimate (*i.e.* corresponding to $ss_{(15)}$) is asserted nonzero, then one proceeds to test the remaining 14 effects as if there never was a 15th effect, and one iteratively steps down in this way until one fails to make an assertion. Such is the case if the ith critical value $c(\alpha, i, k)$ is the upper α quantile of the distribution of $SS_{(i)}/QMSE$, the ith largest of i order statistics (rather than the ith largest of k), under the complete null distribution. Call this variation the *iterative step-down test*, rather than the closed step-down test. Such an iterative approach conjures up a sense of statistical magic since, when testing the effect of any estimate smaller than the largest, one would in essence and reality be ignoring the fact that effects with larger estimates have already been inferred to be nonzero.

There are two advantages to the iterative step-down test. First, as observed by Voss (1988), it is more powerful than the closed step-down test, having smaller critical values after the first. This is not obvious, since the test statistic numerator $SS_{(i)}$ and denominator $QMSE$ are each stochastically larger as a function of i estimators than as a function of k. However, this seems to be born out in practice. For example, Table 1 contains the critical values for the closed and iterative step-down tests for $k = 15$ effects, $\nu = 8$, and $\alpha = 0.10$, 0.05, 0.01, based on 999,999 simulated null samples. The critical values are by definition the same for the largest estimate, but the critical values for the iterative test are substantially smaller for testing all remaining estimates. The second advantage of the iterative test is that fewer critical values need be tabulated, since the iterative test only uses critical values corresponding to the largest order statistics, whereas the closed test requires critical values for each order statistic in a sample of size k.

The focus of this paper is the following conjecture, where *strong control of the error rate* means control under all effect parameter configurations, assuming the standard assumptions of normality and homogeneity of error variances.

Iterative step-down test conjecture. The iterative step-down test, using critical values satisfying equation (1), strongly controls family-wise error rate at the specified level.

A colleague and I have for years sought a rigorous proof of this conjecture, to establish its statistical magic as a happy reality. Alas, each time we 'found' a proof, we subsequently found a hole in it. While we hope someone will succeed where we have thus far failed, it is good to believe that what one hopes to prove is or may well be true. The simulation

Table 1: Critical values for the closed and iterative step-down tests for $k = 15$, $\nu = 8$, and $\alpha = 0.01$, 0.05, 0.10

α	Test Type	c15	c14	c13	c12	c11	c10	c9	c8
0.01	closed	151.2	123.4	98.13	76.31	56.78	39.79	24.66	6.360
	iterative	151.2	84.28	55.25	37.79	25.87	17.35	10.67	5.002
0.05	closed	81.75	67.22	53.93	42.03	31.60	22.42	14.11	5.434
	iterative	81.75	46.80	31.30	21.88	15.40	10.60	6.936	4.054
0.10	closed	60.23	49.51	39.93	31.30	23.65	16.88	10.78	4.907
	iterative	60.23	35.11	23.76	16.83	12.01	8.433	5.712	3.631

results presented in this paper strongly support our belief that the iterative step-down test conjecture is true. A secondary goal is to provide insight that may facilitate analytical proof of the conjecture.

In each of the simulations presented here, the step-down test statistics of Voss (1988) were utilized, but with the sharper critical values $c(\alpha, i, i)$ of the iterative test. Without loss of generality, fix $\sigma = 1$. Thus, pseudo random estimates were generated as $\hat{\theta}_i = \theta_i + \epsilon_i$, for ϵ_i pseudo random $N(0, 1)$, with each effect θ_i as specified, whether zero, a nonzero constant, or $N(0, \sigma = 5)$. In each case, the step-down test was stopped if the statistic $ss_{(\nu)}/qmse$ was significantly large, even though the closed step-down test if continued would still strongly control the error rate. All computations were done using the SAS software.

In the simulations, the following events were of interest. Let IA denote the event of an *incorrect assertion*—namely, that any effect with mean zero is asserted to be nonzero. Let MN denote the event that $ss_{(m)}/qmse > c(\alpha, m, k)$ for $ss_{(m)}$ the sum of squares corresponding to the *maximum nonactive* effect estimate. This condition is necessary but not sufficient for an incorrect assertion, since step-down testing may stop sooner, so $P(IA) \leq P(MN)$. Hence, in search of an analytic proof, showing $P(MN) \leq \alpha$ would establish the conjecture concerning iterative testing. Finally, let $P(A)$ denote the probability of an assertion (correct or not), and $P(CA)$ the probability of a *correct assertion*. Better understanding of the behavior of these probabilities may help someone prove the iterative step-down conjecture.

In Section 2, consider the common case of 15 estimates, corresponding for example to analysis of a 2^{15-11}_{III} fraction. For the setting, we ran an extensive simulation involving 100 distinct randomly chosen effect configurations for each number of active (nonzero) effects from one to seven, plus 100 replications of the null scenario, using family-wise significance level $\alpha = 0.01$. Section 3 contains the results of a similar simulation but using $\alpha = 0.10$. Subsequent sections present simulations with systematically chosen non-null parameter configurations. In Section 4, we consider a small simulation with only five effects and only one active effect. In Section 5, we revisit the common case of 15 estimates, with from zero to eight active effects, but systematically varying the values of the active effects over the values 0,3,6,9. The simulation in Section 6 likewise involves 15 estimates, but with from zero to three active effects with values varying over the values 0.0001, 2, 4, 6, 8. Conclusions are summarized in Section 7.

2. Simulation for 15 Effects, with Zero to Seven Random $N(0, 5)$ Active Effects, with $\alpha = 0.01$

In this section, we present the results of three iterations of an extensive three replicate simulation, including 700 different randomly selected non-null parameter configurations, providing substantial evidence of the conjecture that the iterative step-down test under considerations does strongly control family-wise error rates.

In particular, consider again the case of 15 estimates, here with from zero to seven active effects. For each number of active effects 1,2,...,7, we generated 100 random parameter configurations, where the active effects were independent $N(0, \sigma = 5)$, yielding 700 distinct non-null parameter configurations. Also included were 100 replications of the null parameter configuration, giving 800 cases in total. For each of these 800 cases, 10,000 sets of 15 estimates were generated by adding a $N(0, 1)$ error to each active or null effect, and the iterative step-down test was applied to each of the 10,000 sets of estimates set using $\alpha = 0.01$. In each of the 800 cases, the 10,000 tests were used to estimate the probability of an incorrect assertion, $P(IA)$. This same process was replicated three times, using the same 800 cases or parameter configurations but using different $N(0, 1)$ errors in each replicate, giving three estimates of $P(IA)$ for each of the 800 cases. The results are as follows.

For each of the 800 cases or parameter configurations, given an estimated value of $P(IA)$ from each of the three replicates, the minimum of the three values was saved. Only 11 of the 700 non-null parameter configurations yielded $\min P(IA) > 0.01$, compared to 20 false positives in the 100 null cases. Furthermore, of the 11 non-null cases so flagged, the largest estimate of $\min P(IA)$ was only 0.0110. This simulation strongly supports the truth of the conjecture that iterative application of the step-down test strongly controls the family-wise error rate.

While the above results seem convincing, one might want further evidence that the 11 (of 700) non-null parameter configurations yielding values of $\min P(IA)$ between 0.01 and 0.011 were indeed false positives. To this end, we repeated the above process two more times, using the same 700 non-null parameter configurations each time, but generating different random errors. In the second iteration of the simulation, only 13 of the 700 non-null parameter configurations yielded $\min P(IA) > 0.01$, compared to 17 false positives in the 100 null cases. In the third iteration of the simulation, only 13 of the 700 non-null parameter configurations yielded $\min P(IA) > 0.01$, compared to 15 false positives in the 100 null cases. More importantly, while the three iterations respectively flagged 11, 13 and 13 of 700 non-null parameter configurations as having $\min P(IA) > 0.01$, comparing the results of the three iterations of the three-replicate simulations, none of the 700 non-null parameter configurations where flagged in all three iterations. In other words, of the 700 randomly chosen non-null parameter configurations, there was no non-null parameter configuration for which $\min P(IA)$ exceeded $\alpha = 0.01$ for all three iterations of the simulation.

This simulation concerned the fairly common case of $k = 15$ effects, with $\nu = 8$ effects to form the denominator and a family-wise test size of $\alpha = 0.01$. In this case, the simulation

strongly supports the truth of the conjecture that iterative application of the step-down test strongly controls the family-wise error rate.

3. Simulation for 15 Effects, with from Zero to Seven Random $N(0, 5)$ Active Effects, with $\alpha = 0.10$

The prior section described three iterations of a simulation for $k = 15$ effects, using $\nu = 8$ effects to form the denominator, and with $\alpha = 0.01$. One may well prefer to use a larger value of α for a screening experiment and family-wise control of test size. In this section, we present the results of one iteration of the simulation presented in the prior section, but with $\alpha = 0.10$ rather than 0.01. As in the prior section, the simulation includes 700 different randomly selected non-null parameter configurations, plus 100 iterations for the null parameter configuration. The results for $\alpha = 0.10$ are as follows.

For each of the 800 cases or parameter configurations, the minimum estimated value of $P(IA)$ was computed over the three replications. Only 5 of the 700 non-null parameter configurations yielded $\min P(IA) > 0.10$, compared to 11 false positives in the 100 null cases. Furthermore, of the 5 non-null cases so flagged, the largest estimate of $\min P(IA)$ was only 0.1027, with the other estimates ranging from 0.1002 to 0.1005.

With more false positives (11 out of 100 null cases) than possible true positives (5 out of 700 non-null cases), and given the small estimates of $\min P(IA)$ in the prospective non-null cases, the results are again encouraging for $\alpha = 0.10$. In short, this simulation also strongly supports the truth of the conjecture that iterative application of the step-down test strongly controls the family-wise error rate.

4. A Small Simulation with Five Effects with One Active

In this Section, we consider a small simulation with only five effects ($k = 5$) and at most one active effect θ, forming $qmse$ from the $\nu = 3$ smallest sums of squares. The values considered for the 'active' effects are $\theta = 0, 1, \ldots, 10$, with $\theta = 0$ treated as active but negligible for estimating probabilities. The simulation included 10,000 runs for each value of θ, using $\alpha = 0.10$. Simulation results are provided in Table 2.

Table 2: Simulation for $k = 5$, $\nu = 3$, $\alpha = 0.1$, and one active effect θ

θ	$P(IA)$	$P(CA)$	$P(A)$	$P(MN)$
0	0.089	0.028	0.100	0.093
1	0.074	0.060	0.104	0.081
2	0.068	0.156	0.168	0.078
3	0.077	0.308	0.310	0.087
4	0.094	0.501	0.501	0.099
5	0.096	0.671	0.671	0.097
6	0.097	0.809	0.809	0.098
7	0.099	0.901	0.901	0.099
8	0.100	0.953	0.953	0.100
9	0.097	0.981	0.981	0.097
10	0.100	0.993	0.993	0.100

As noted above, the case with $\theta = 0$ is treated in the simulation as a non-zero but negligibly small effect. Since this is in essence the null case, any assertion is really an incorrect assertion, so it is not surprising that $P(A)$ essentially equals $\alpha = 0.10$. Not surprisingly, the results suggest that going from the null case to having one negligible active effect (approximated by and corresponding to $\theta = 0$) causes $P(IA)$ to drop discretely from 0.10 to 0.08909—the value listed for $P(IA)$ in Table 2. One might then anticipate $P(IA)$ being monotone increasing in θ for $\theta > 0$. Interestingly though, such is not the case. Instead, $P(IA)$ actually decreases as θ goes from 0 (*i.e.* negligible) to 1 to 2, then increases for $\theta > 2$; the slight exception when $\theta = 9$ is probably just simulation error. It is surprising that $P(IA)$ is not monotone in nonzero θ. This lack of monotonicity may be useful in proving the conjecture, if one can prove concavity, for example. It is perhaps also noteworthy that the greatest disparity between $P(IA)$ and $P(MN)$ is when θ is small, *i.e.* about 2 or 3.

Note that the probability of an incorrect assertion, $P(IA)$, is at most 0.10 for all nonzero values of θ, suggesting error rate control, so this simulation supports the iterative testing conjecture.

5. Simulation for 15 Effects, with Seven or Fewer Active Effects with Values 3, 6 or 9, with $\alpha = 0.01$

In this Section, we consider the common case of 15 estimates (*e.g.* corresponding to a 2^{15-11}_{III} fraction), with from zero to seven active effects, systematically varying the values of the last seven effects $\theta_9, \ldots, \theta_{15}$ to have nondecreasing values 0, 3, 6 or 9, yielding 120 distinct parameter configurations. Any effect with value zero is treated as inactive in estimating probabilities.

Simulation results are provided in Table 3 (after references). Only a few of the parameter configurations involving an effect of size three are shown, since $P(IA)$ is well below α in most such cases, including all cases not displayed. A few general observations are in order. The probability of making any assertions, $P(A)$, and the probability of making any correct assertions, $P(CA)$, are largest when there are a few large effects. The probability of an incorrect assertion, $P(IA)$, and the probability of the necessary condition for an incorrect assertion, $P(MN)$, are nearly equal when all active effects are very large, *i.e.* when all active effects are highly likely to be asserted to be nonzero. Most importantly for our purposes, note that the probability of an incorrect assertion, $P(IA)$, never exceeds $\alpha = 0.01$ by more than a negligible amount attributable to simulation error, supporting the conjecture.

6. Simulation for 15 Effects, with Three or Fewer Active Effects with Values 0.0001, 2, 4, 6, or 8, with $\alpha = 0.01$

In the prior section, it was seen that $P(IA)$ and $P(MN)$ are nearly equal when all active effects are very large, and that the test had more power when there were a few large effects. In this section, we examine whether the inclusion of a few small active effects among only a small number of active effects sheds any light on the relative behavior of $P(IA)$ and $P(MN)$, in case this helps in the quest for an analytic proof of the conjecture. In particular, we revisit the common case of 15 estimates, but with from zero to three active effects, systematically varying the values of the last three effects $\theta_{13}, \theta_{14}, \theta_{15}$ to have nondecreasing

values 0.0001, 2, 4, 6, or 8. This yields 56 distinct parameter configurations. Any effect with value zero is treated as inactive in estimating probabilities.

Simulation results are provided in Table 4 (after references). To save table space, the results are not shown for any parameter configurations with $\theta_{13} = 2$; these cases all yield $0.003 < P(IA) < 0.007$ so are not interesting.

A few general observations are in order. The probability of making any assertions, $P(A)$, and the probability of making any correct assertions, $P(CA)$, are largest when there are a few large effects. Conversely, consider the three parameter configurations when all active effects are small—namely, when the active effects consist of one, two or three effects of size 0.0001. These are close to the complete null case, when all 15 effects are zero, so it is not surprising that $P(A)$ is approximately α in each case. That said, incorrect assertions are more likely than correct assertions, since there are simply more null effects.

In examining the behavior of $P(IA)$ and $P(MN)$, perhaps the most interesting observations is that these can lack monotonicity in the effects. For example, in the cases with 14 null effects, these probabilities decrease as the lone active effect increases from 0.0001 to 2, but then they increase. It seems obvious that the conjecture should be true if all active effects are either very large or very small. In particular, any very large effects will almost surely correspond to the largest estimates and almost surely be asserted to be nonzero, after which the step-down test will proceed as if they were never in the picture; if so, then large effects shouldn't cause the size of the iterative step-down test to exceed α. Also, the estimates of any very small effects will behave like null effect estimates, except asserting any of them to be active would be correct assertions, so the existence of very small effects may cause $P(IA)$ to be less than α. In view of this, perhaps an analytic proof of the conjecture would follow if one could establish that either $P(IA)$ or $P(MN)$ is concave in each nonzero effect.

7. Concluding Remarks

In the analysis of orthogonal saturated designs, certain closed step-down tests are known to provide strong family-wise control of error rates. Many authors have proposed applying such step-down tests iteratively, but it remains an open problem to establish analytically the conjecture that iterative step-down tests strongly control family-wise error rates. The various simulations presented in this paper strongly support this conjecture. Then may also provide some insight that will be helpful in the search for an analytic proof of the conjecture. In particular, while the probability of making any incorrect assertions is apparently largest in the null case and when all active effects are very large, interestingly, this probability is apparently not monotone in the value of active effects. Simulations suggest that the behavior of the probability of incorrect assertions may be concave, but clearly it is not monotone.

Acknowledgements

It is my pleasure to contribute this paper for inclusion in the special volume to honor the late Professor Aloke Dey. It was my pleasure to have known him—a gentleman and a scholar.

References

Alam, K. and Rizvi, M. H. (1966). Selection from multivariate normal populations. *Annals of the Institute of Statistical Mathematics*, **18**, 307–318.

Birnbaum, A. (1959). On the analysis of factorial experiments without replication. *Technometrics*, **1**, 343–357.

Daniel, C. (1959). Use of half-normal plots in interpreting factorial two-level experiments. *Technometrics*, **1**, 311–341.

Hamada, M. and Balakrishnan, N. (1998). Analyzing unreplicated factorial experiments: a review with some new proposals. *Statistica Sinica*, **8**, 1-–41.

Holm, S., Mark, S. and Adolfsson, T. (2005). A step-down test for effects in unreplicated factorial trials. *Communication in Statistics–Theory and Methods,* **34(2)**, 405–416.

Mahamunulu, D. M. (1967). Some fixed-sample ranking and selection problems. *Annals of Mathematical Statistics*, **38**, 1079–1091.

Marcus, R., Peritz, E. and Gabriel, K. R. (1976). On closed testing procedures with special reference to ordered analysis of variance. *Biometrika*, **63**, 655–660.

Voss, D. T. (1988). Generalized modulus-ratio tests for analysis of factorial designs with zero degrees of freedom for error. *Communication in Statistics–Theory and Methods* **17**, 3345–3359.

Voss, D. T. (1999). Analysis of orthogonal saturated designs. *Journal of Statistical Planning and Inference*, 111–130.

Voss, D. T. and Wang, W. (2006a). On adaptive testing in orthogonal saturated designs. *Statistica Sinica*, **16**, 227–234.

Voss, D. T. and Wang, W. (2006b) Analysis of orthogonal saturated designs. In: Dean A., Lewis S. (eds) *Screening*. Springer, New York, NY.

Zahn, D. A. (1975a). Modifications of and revised critical values for the half-normal plots. *Ph.D. thesis*, Harvard Univ.

Zahn, D. A. (1975b). An empirical study of the half-normal plot. *Technometrics*, **17**, 201–211.

Table 3: Simulation for $k = 15$, $\nu = 8$, $\alpha = 0.01$, and up to seven active effects of size 3, 6 or 9

No. Null	θ_9–θ_{15}	$P(A)$	$P(CA)$	$P(IA)$	$P(MN)$
15	0 0 0 0 0 0 0	0.0102	0.0000	0.0102	0.0102
14	0 0 0 0 0 0 3	0.0791	0.0780	0.0081	0.0083
14	0 0 0 0 0 0 6	0.5852	0.5852	0.0099	0.0099
14	0 0 0 0 0 0 9	0.9477	0.9477	0.0101	0.0101
13	0 0 0 0 0 3 3	0.0891	0.0889	0.0071	0.0076
13	0 0 0 0 0 3 6	0.4899	0.4899	0.0078	0.0081
13	0 0 0 0 0 3 9	0.9102	0.9102	0.0082	0.0084
13	0 0 0 0 0 6 6	0.5986	0.5986	0.0104	0.0104
13	0 0 0 0 0 6 9	0.9098	0.9098	0.0104	0.0104
13	0 0 0 0 0 9 9	0.9478	0.9478	0.0102	0.0102
12	0 0 0 0 3 3 3	0.0726	0.0725	0.0049	0.0055
:	: : : : : : :	:	:	:	:
12	0 0 0 0 3 9 9	0.9043	0.9043	0.0078	0.0082
12	0 0 0 0 6 6 6	0.5386	0.5386	0.0094	0.0095
12	0 0 0 0 6 6 9	0.8515	0.8515	0.0101	0.0101
12	0 0 0 0 6 9 9	0.9031	0.9031	0.0097	0.0097
12	0 0 0 0 9 9 9	0.9258	0.9258	0.0100	0.0100
:	: : : : : : :	:	:	:	:
111	0 0 0 3 9 9 9	0.8569	0.8569	0.0075	0.0080
11	0 0 0 6 6 6 6	0.4408	0.4408	0.0100	0.0102
11	0 0 0 6 6 6 9	0.7567	0.7567	0.0099	0.0099
11	0 0 0 6 6 9 9	0.8243	0.8243	0.0099	0.0099
11	0 0 0 6 9 9 9	0.8531	0.8531	0.0100	0.0100
11	0 0 0 9 9 9 9	0.8730	0.8730	0.0104	0.0104
:	: : : : : : :	:	:	:	:
10	0 0 3 9 9 9 9	0.7623	0.7623	0.0071	0.0078
10	0 0 6 6 6 6 6	0.3137	0.3137	0.0094	0.0096
10	0 0 6 6 6 6 9	0.6143	0.6143	0.0101	0.0104
10	0 0 6 6 6 9 9	0.6960	0.6960	0.0096	0.0097
10	0 0 6 6 9 9 9	0.7321	0.7321	0.0098	0.0099
10	0 0 6 9 9 9 9	0.7557	0.7557	0.0095	0.0096
10	0 0 9 9 9 9 9	0.7721	0.7721	0.0096	0.0096
:	: : : : : : :	:	:	:	:
9	0 3 9 9 9 9 9	0.5901	0.5901	0.0062	0.0075
9	0 6 6 6 6 6 6	0.1789	0.1789	0.0082	0.0097
9	0 6 6 6 6 6 9	0.4178	0.4178	0.0084	0.0095
9	0 6 6 6 6 9 9	0.4958	0.4958	0.0099	0.0104
9	0 6 6 6 9 9 9	0.5361	0.5361	0.0099	0.0102
9	0 6 6 9 9 9 9	0.5658	0.5658	0.0096	0.0098
9	0 6 9 9 9 9 9	0.5814	0.5814	0.0094	0.0094
9	0 9 9 9 9 9 9	0.5989	0.5989	0.0100	0.0101
:	: : : : : : :	:	:	:	:
8	3 9 9 9 9 9 9	0.3261	0.3261	0.0012	0.0070
8	6 6 6 6 6 6 6	0.0614	0.0614	0.0004	0.0094
8	6 6 6 6 6 6 9	0.1839	0.1839	0.0010	0.0104
8	6 6 6 6 6 9 9	0.2340	0.2340	0.0014	0.0101
8	6 6 6 6 9 9 9	0.2611	0.2611	0.0018	0.0100
8	6 6 6 9 9 9 9	0.2839	0.2839	0.0023	0.0100
8	6 6 9 9 9 9 9	0.2995	0.2995	0.0026	0.0096
8	6 9 9 9 9 9 9	0.3051	0.3051	0.0030	0.0101
8	9 9 9 9 9 9 9	0.3173	0.3173	0.0029	0.0098

Table 4: Simulation for $k = 15$, $\nu = 8$, $\alpha = 0.01$, and up to three active effects of size 0.0001, 2, 4, 6, or 8

No. Null	θ_{13}	θ_{14}	θ_{15}	$P(A)$	$P(CA)$	$P(IA)$	$P(MN)$
15	0	0	0	0.0102	0.0000	0.0102	0.0102
14	0	0	0.0001	0.0099	0.0009	0.0094	0.0094
14	0	0	2	0.0257	0.0221	0.0072	0.0074
14	0	0	4	0.1998	0.1995	0.0097	0.0098
14	0	0	6	0.5826	0.5826	0.0100	0.0100
14	0	0	8	0.8767	0.8767	0.0102	0.0102
13	0	0.0001	0.0001	0.0102	0.0020	0.0094	0.0095
13	0	0.0001	2	0.0251	0.0221	0.0066	0.0069
13	0	0.0001	4	0.2007	0.2005	0.0083	0.0085
13	0	0.0001	6	0.5811	0.5811	0.0094	0.0094
13	0	0.0001	8	0.8774	0.8774	0.0094	0.0095
13	0	2	2	0.0293	0.0280	0.0048	0.0052
13	0	2	4	0.1561	0.1560	0.0065	0.0069
13	0	2	6	0.5067	0.5067	0.0070	0.0073
13	0	2	8	0.8309	0.8309	0.0073	0.0075
13	0	4	4	0.2146	0.2146	0.0087	0.0090
13	0	4	6	0.4970	0.4970	0.0096	0.0099
13	0	4	8	0.8180	0.8180	0.0098	0.0100
13	0	6	6	0.5988	0.5988	0.0100	0.0100
13	0	6	8	0.8252	0.8252	0.0101	0.0101
13	0	8	8	0.8831	0.8831	0.0102	0.0102
12	0.0001	0.0001	0.0001	0.0100	0.0030	0.0085	0.0086
12	0.0001	0.0001	2	0.0264	0.0235	0.0069	0.0071
12	0.0001	0.0001	4	0.2014	0.2012	0.0080	0.0082
12	0.0001	0.0001	6	0.5810	0.5810	0.0090	0.0090
12	0.0001	0.0001	8	0.8782	0.8782	0.0089	0.0090
12	0.0001	2	2	0.0280	0.0269	0.0049	0.0052
12	0.0001	2	4	0.1556	0.1555	0.0061	0.0064
12	0.0001	2	6	0.5055	0.5055	0.0063	0.0066
12	0.0001	2	8	0.8288	0.8288	0.0068	0.0071
12	0.0001	4	4	0.2182	0.2182	0.0079	0.0082
12	0.0001	4	6	0.4966	0.4966	0.0091	0.0092
12	0.0001	4	8	0.8151	0.8151	0.0093	0.0096
12	0.0001	6	6	0.5976	0.5976	0.0096	0.0096
12	0.0001	6	8	0.8237	0.8237	0.0095	0.0095
12	0.0001	8	8	0.8840	0.8840	0.0098	0.0098
:	:	:	:	:	:	:	:
12	4	4	4	0.1815	0.1815	0.0078	0.0085
12	4	4	6	0.3991	0.3991	0.0089	0.0092
12	4	4	8	0.7293	0.7293	0.0084	0.0089
12	4	6	6	0.4879	0.4879	0.0087	0.0089
12	4	6	8	0.7386	0.7386	0.0091	0.0093
12	4	8	8	0.8095	0.8095	0.0090	0.0093
12	6	6	6	0.5384	0.5384	0.0100	0.0101
12	6	6	8	0.7428	0.7428	0.0101	0.0101
12	6	8	8	0.8109	0.8109	0.0100	0.0101
12	8	8	8	0.8442	0.8442	0.0102	0.0102

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 319–336

Optimal Crossover Designs for Generalized Linear Models: An Application to Work Environment Experiment

Jeevan Jankar[1] **and Abhyuday Mandal**[1]
[1]*Department of Statistics*
University of Georgia, Athens, USA

Received: 11 December 2020; Revised: 26 February 2021; Accepted: 04 March 2021

Abstract

We consider locally D-optimal crossover designs for generalized linear models. Three different types of responses were recorded in a work environment experiment conducted at Booking.com. These responses follow Poisson, beta and gamma distributions. The responses from the same subjects are naturally correlated. To capture the dependence among these observations, we use six different types of correlation structures. The optimal allocations of subjects to each treatment sequence are obtained by minimizing the objective function, which is the variance of direct treatment effect estimates. We show that optimal allocations are reasonably robust to a different choice of correlation structures. Although uniform allocations are widely used in practice, we establish these designs are sub-optimal under certain conditions.

Key words: D-Optimality; Generalized Linear Models; Generalized estimating Equations; Latin Square Design.

1. Introduction

Crossover designs, also known as repeated measurements designs or change-over designs, have been used extensively in pharmaceutical and agriculture research. Most of the present work focuses on optimal crossover designs for normal responses. But, there are ample examples where responses are non-normal and needed to be described by generalized linear models (GLMs). The optimal designs obtained for normal responses can be quite inefficient for GLMs. The goal of this manuscript is to bridge this gap in the literature and obtain efficient designs for crossover experiments with responses under GLMs, including Poisson, beta, gamma responses, *etc.*

In crossover experiments, every subject is exposed to a sequence of treatments over different time periods, *i.e.*, subjects crossover from one treatment to another. Among different types of experiments available for treatment comparisons with multiple periods, the crossover designs are among the most important ones. We can get the same number of observations but with fewer subjects. The time at which the subject receives the treatment is known as a *period* and the order in which the particular subject receives treatments is known as a *sequence*. Each subject receives one treatment in each period, and the corresponding

Corresponding Author: Abhyuday Mandal
Email: amandal@stat.uga.edu

response is recorded. Naturally, crossover designs also provide within-subject information about treatment differences.

Most of the current literature in the crossover design deal with the continuous responses (see, for example, Kershner and Federer (1981), Laska and Meisner (1985), Matthews (1987), Carriere and Huang (2000), and the references therein). The problem of determining optimal crossover designs for continuous responses has been studied extensively (see, for example, Bose and Dey (2009) for a review of results). For examples of practical cases where the responses are discrete, such as binary responses, one may refer to Jones and Kenward (2014) and Senn (2003).

Different fixed effects models have been proposed in the literature, but the following linear model is used extensively to formulate crossover designs. For an experiment involving n subjects and p periods, the response is modeled as

$$Y_{ij} = \lambda + \beta_i + \alpha_j + \tau_{d(i,j)} + \rho_{d(i-1,j)} + \epsilon_{ij}, \tag{1}$$

where Y_{ij} is the observation from the jth subject in the ith time period, with $i = 1, \ldots, p$ and $j = 1, \ldots, n$. Here $d(i, j)$ stands for the treatment assigned to the jth subject at time period i and $\lambda, \beta_i, \alpha_j, \tau_{d(i,j)}, \rho_{d(i-1,j)}$ are the corresponding overall mean, the ith period effect, the jth subject effect, the direct treatment effect and the carryover treatment effect respectively. We define $\rho_{d(0,j)} = 0$. Here $\epsilon_{ij}s$ are the uncorrelated error terms which follow a normal distribution with zero mean and constant variance. Model (1) is commonly referred to as the traditional model due to its extensive use in the literature.

Note that the Fisher information matrix, for the linear model (1), is independent of model parameters because all the effects are fixed. Various optimality criteria such as A-, D-, E-optimality depend on this information matrix (see, for example, Pukelsheim (1993)). The optimality of crossover designs for linear models has been studied extensively in the literature. Hedayat and Afsarinejad (1978), Cheng and Wu (1980) and Kunert (1984b) considered the optimality of balanced, uniform designs. Optimality of designs when $p \leq t$ were first formulated in Dey *et al.* (1983). However, these results are not directly applicable to non-normal responses. In the case of GLMs Fisher information matrix depends on the model parameters (McCullagh and Nelder (1989), Stufken and Yang (2012)); hence the results on the optimality of crossover designs for linear models cannot be readily extended to other types of responses.

2. Preliminary Setup: Crossover Design for GLM

Most of the results available on optimal crossover designs deal with normal responses, so the results on crossover designs under GLMs are limited. Before presenting the results for optimal crossover designs, we formally introduce the associated generalized linear models for crossover designs.

Consider a crossover experiment with t treatments, n subjects, and p periods. The response from the jth subject is $Y_j = (Y_{1j}, \ldots, Y_{pj})'$ and the overall response for these n subjects are denoted as $Y_1, \ldots, Y_n$. The marginal distribution of Y_{ij} is described using GLMs as mentioned in Liang and Zeger (1986). Then the marginal mean μ_{ij} of Y_{ij} for crossover

trial is modeled as

$$\mathrm{g}(\mu_{ij}) = \eta_{ij} = \lambda + \beta_i + \tau_{d(i,j)} + \rho_{d(i-1,j)}, \tag{2}$$

where $i = 1, \ldots, p; j = 1, \ldots, n$; λ is the overall mean, β_i represents the effect of the ith period, τ_s is the direct effect due to treatment s assigned to subject j in period i, ρ_s is the carryover effect due to treatment s assigned to subject j in period $(i-1)$, where $s = 1, \ldots, t$ and g is a link function. We define $\rho_{d(0,j)} = 0$. For example, $\mu_{1,j}$ is modeled as $g(\mu_{1,j}) = \eta_{1,j} = \lambda + \tau_{d(1,j)}$.

In many situations interest lies mainly in the estimation of direct treatment effects, so we treat carryover effects as nuisance parameter. To ensure the estimability of the parameters, we set the baseline constraints as $\beta_1 = \tau_1 = \rho_1 = 0$.

Consider $\beta = (\beta_2, \ldots, \beta_p)'$, $\tau = (\tau_2, \ldots, \tau_t)'$ and $\rho = (\rho_2, \ldots, \rho_t)'$, which define the parameter vector $\theta = (\lambda, \beta, \tau, \rho)'$. Then the linear predictor corresponding to the jth subject, $\eta_j = (\eta_{1j}, \ldots, \eta_{pj})'$ can be written as

$$\eta_j \quad = \quad X_j\theta.$$

The corresponding design matrix X_j can be written as $X_j = [1_p, P_j, T_j, F_j]$, where P_j is $p \times (p-1)$ matrix such that $P_j = [0_{(p-1)1}, I_{p-1}]'$; T_j is a $p \times (t-1)$ matrix with its $(i,s)th$ entry equal to 1 if subject j receives the direct effect of the treatment s in the ith period and zero otherwise; F_j is a $p \times (t-1)$ matrix with its $(i,s)th$ entry equal to 1 if subject j receives the carryover effect of the treatment s in the ith period and zero otherwise. The columns of T_j and F_j are indexed by $2, \ldots, t$. Note that T_j and F_j have $t-1$ columns instead of t, because of the baseline constraints $\tau_1 = \rho_1 = 0$.

If the number of subjects is fixed to n and the number of periods is p, then we determine the proportion of subjects assigned to a particular treatment sequence. As the number of periods is fixed to p, each treatment sequence will be of length p and a typical sequence can be written as $\omega = (t_1, \ldots, t_p)'$ where $t_i \in \{1, \ldots, t\}$. Now, let Ω be the set of all such sequences and n_ω denote the number of subjects assigned to sequence ω. Then, the total number of subjects n can be written as $n = \Sigma_{\omega \in \Omega} n_\omega$, with $n_\omega \geq 0$. A crossover design ζ in approximate theory is specified by the set $\{p_\omega, \omega \in \Omega\}$, where $p_\omega = n_\omega/n$ is the proportion of subjects assigned to treatment sequence ω. Such a crossover design ζ can be denoted as follows:

$$\zeta = \left\{ \begin{array}{cccc} \omega_1 & \omega_2 & \ldots & \omega_k \\ p_{\omega_1} & p_{\omega_2} & \ldots & p_{\omega_k} \end{array} \right\},$$

where k is the number of treatment sequences involved, such that $\sum_{i=1}^{k} p_{\omega_i} = 1$. From the definitions of matrices T_j and F_j it can be noted that they depend only on the treatments sequence ω that subject j receives. Let T_ω be the matrix T and F_ω be the matrix F when subject receives sequence ω. Then it can be inferred that all the subjects receiving sequence ω have same T and F matrices. This implies, all the subjects receiving sequence ω have same design matrix *i.e.* $X_j = X_\omega$ as $P_j = [0_{(p-1)1}, I_{p-1}]'$.

Following Jankar *et al.* (2020), we use generalized estimating equations (GEEs) to estimate quasi-likelihood estimates of the model parameters. As mentioned in Zeger *et al.* (1988, equation (3.1)), it can be shown that for repeated measurement model, the GEEs are

$$\sum_{j=1}^{n} \frac{\partial \mu_j'}{\partial \theta} W_j^{-1} (Y_j - \mu_j) = 0,$$

where $\mu_j = (\mu_{1j}, \ldots, \mu_{pj})'$ and the asymptotic variance for the GEE estimator $\hat{\theta}$ (see Zeger *et al.*, 1988, equation (3.2)) is

$$\text{Var}(\hat{\theta}) = \left[\sum_{j=1}^{n} \frac{\partial \mu_j'}{\partial \theta} W_j^{-1} \frac{\partial \mu_j}{\partial \theta} \right]^{-1},$$

where $W_j = Cov(Y_j)$.

We can write the above equation in the form of approximate designs as follows,

$$\text{Var}(\hat{\theta}) = \sum_{\omega \epsilon \Omega} n p_\omega \frac{\partial \mu_\omega'}{\partial \theta} W_\omega^{-1} \frac{\partial \mu_\omega}{\partial \theta},$$

where W_ω corresponds to the covariance matrix of response Y_j when subject j receives treatment sequence ω.

Main interest usually lies in estimating the direct treatment effect contrasts. So, instead of working with the full variance-covariance matrix of parameter estimator $\hat{\theta}$, we concentrate only on the variance of the estimator of treatment effect $\text{Var}(\hat{\tau})$. Here

$$\text{Var}(\hat{\tau}) \quad = \quad H \text{Var}(\hat{\theta}) H', \tag{3}$$

where H is a $(t-1) \times m$ matrix given by $[0_{(t-1)1}, 0_{(t-1)(p-1)}, I_{t-1}, 0_{(t-1)(t-1)}]$, where $m = p + 2t - 2$ is the total number of parameters in θ and $0_{(t-1)(p-1)}$ is a $(t-1) \times (p-1)$ matrix of zeros.

Optimal proportions for crossover designs are obtained by minimizing the variances of estimators of treatment effect. We use the D-optimality criterion and use the determinant of $\text{Var}(\hat{\tau})$ as our objective function. Then an optimal design ζ^* minimizes the determinant of $\text{Var}(\hat{\tau})$ in equation (3) with respect to p_ω, such that $\sum_{w \in \Omega} p_w = 1$.

Note that the baseline constraints $\tau_1 = 0$ we set earlier results in the estimators for $\tau_i - \tau_1$ for $i \geq 2$. In the case of a D-optimality criterion, it is okay to use an above baseline constraint, but we must use different constraints in other optimality criteria. The above method has been discussed in detail in Jankar *et al.* (2020).

3. The Work Environment Experiment

We considered the data obtained from the work environment experiment conducted at Booking.com (Pitchforth *et al.* (2020)). In recent years, most corporate offices and organizations are adopting open office spaces over the traditional cubical office spaces. Since there

were no previous studies to examine the effects of office designs in workspaces, Booking.com conducted an experiment to assess different office spacing efficiency.

In the work environment experiment, there were a total of $n = 288$ participants. These participants were divided into four groups G_1, G_2, G_3, G_4 with each group having an equal number of (72) individual participants. This experiment is essentially an uniform crossover design with $p = 4$ periods and $t = 4$ treatments. Periods were named Wave1, Wave2, Wave3 and Wave4, where each Wave had a duration of 2 weeks. The four treatments involved in this experiment are office designs named as A (Activity-Based), B (Open Plan), C (Team Offices), and D (Zoned Open Plan), as shown in the figure below:

$A-$ ACT

$B-$ OPEN

$C-$ TEAM

$D-$ ZONE.

The images are reproduced from the manuscript Pitchforth *et al*. (2020), under Creative Commons Attribution license (https://creativecommons.org/licenses/by/4.0/).

During the experiment, each group is exposed to different treatments over different periods depending on the treatment sequence. At a given particular period, there was no interaction between subjects from different groups. A Latin square design (Wu and Hamada, 2009) of order four has been used to decide the sequence of exposure so that no group was exposed to the conditions in the same order as any other group. The design is shown below in Table 1. A total of $m = 23$ covariates was involved in the experiment, but we consider only the most important ones in our fitted model.

Table 1: Latin square design

Groups ⇒ Period ⇓	G_1	G_2	G_3	G_3
Wave 1	OPEN	TEAM	ZONE	ACT
Wave 2	ACT	ZONE	OPEN	TEAM
Wave 3	ZONE	ACT	TEAM	OPEN
Wave 4	TEAM	OPEN	ACT	ZONE

In the following analysis, we consider three different responses that were recorded during the experiment. We discuss these responses in more detail in the following sections.

These three responses follow three different types of distributions. We make an extra assumption that the responses from a particular subject are mutually correlated, while the responses from different subjects are uncorrelated. To capture the dependency among the observations coming from the same subject, we calculate optimal proportions for these different responses using six different correlation structures proposed in Section 2.3 of Jankar *et al.* (2020) and shown in the Appendix. For each correlation matrix that we consider, an optimal design ζ^* is the one minimizing the determinant of $\text{Var}(\hat{\tau})$ in equation (3) with respect to p_ω such that $\sum_{w\in\Omega} p_w = 1$.

We use different colors to represent different correlation structures. The color scheme that we use is as follows:

Correlation Structure	Color
Corr(1)	
Corr(2)	
Corr(3)	
Corr(4)	
Corr(5)	
Corr(6)	

4. Poisson Regression

In the case of Poisson response we calculate locally optimal design for the above example under the model,

$$\log(\mu_{ij}) = \eta_{ij} = \lambda + \beta_i + \tau_{d(i,j)} + \rho_{d(i-1,j)},$$

where notations have the same meaning as in equation (2). In the above experiment, there were many different types of responses recorded. We consider the response *commit count* to illustrate the optimal crossover design for the Poisson response. The commit counts were the number of commits submitted to the main git repository.

4.1. Analysis of data

We consider the three main predictors in the model, which are *area*, *wave* and *carryover* where *area* corresponds to the direct treatment effect, *wave* corresponds to the period effect, and *carryover* corresponds to the carryover effect of a treatment given in previous period. We use different kinds of correlation matrices and calculate the optimal proportions. As mentioned earlier we consider baseline constraints as $\beta_1 = \tau_1 = \rho_1 = 0$, so that all the parameters are estimable.

We fit the Poisson regression model to the commit data by using the `glm` function in **R** and calculate the parameter estimates. We use these parameter estimates to make a guess for values of unknown parameters. Our nominal guess for the parameter values is $\theta_1 = [2,$

0.3, 0.8, −0.1, −0.2, 0.04, −0.2, −0.6, 0.15, −0.4]. It is interesting to note that carryover effects are larger than direct effects even though θ_1 is calculated using experimental data. Now, we calculate the optimal designs for different correlation structures by minimizing the objective function. We also calculate optimal proportions for another parameter θ_2 = [2, 0.3, 0.8, −0.1, −2.0, 0.40, −2.0, −1.0, 0.30, −1.0], which is significantly different from θ_1.

4.2. Optimal designs

In the Table 2, we present the optimal proportions corresponding to Poisson response for six different choices of the correlation matrix.

Table 2: Optimal proportions in case of Poisson response

Correlation	θ_1				θ_2			
Structure	$BADC$	$CDAB$	$DBCA$	$ACBD$	$BADC$	$CDAB$	$DBCA$	$ACBD$
$Corr(1)$	0.2500	0.2500	0.2500	0.2500	0.2500	0.2500	0.2500	0.2500
$Corr(2)$	0.2500	0.2500	0.2500	0.2500	0.2747	0.3113	0.1841	0.2299
$Corr(3)$	0.2500	0.2500	0.2500	0.2500	0.2795	0.3074	0.1798	0.2333
$Corr(4)$	0.2500	0.2500	0.2500	0.2500	0.2562	0.3168	0.1860	0.2410
$Corr(5)$	0.2500	0.2500	0.2500	0.2500	0.2736	0.3138	0.1922	0.2204
$Corr(6)$	0.2500	0.2500	0.2500	0.2500	0.2537	0.3190	0.1844	0.2429

As seen from Table 2, in case of Poisson response the optimal proportions that we obtain using θ_1 are nearly uniform and that using θ_2 are non-uniform.

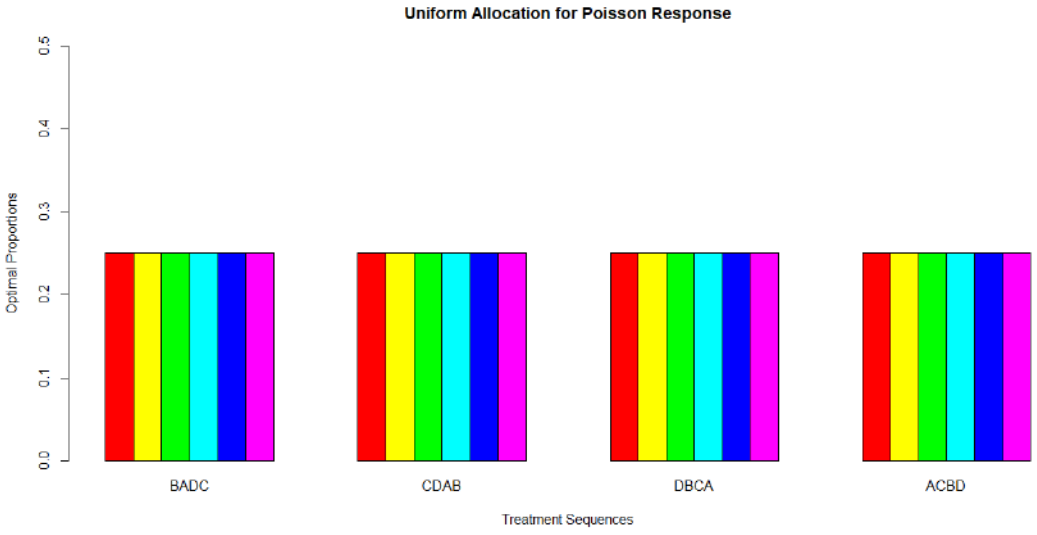


Figure 1: Uniform optimal proportions for Poisson response under θ_1

The plots in Figures 1 and 2 represent the optimal proportions for Poisson response under θ_1 and θ_2 respectively. It can be seen from these plots that the optimal proportions do not vary much when we use different correlation structures under θ_1 and θ_2. In most situations in practice, uniform, optimal designs (the same proportion for each treatment sequence) are used. It is clear from the above analysis that those uniform designs are sub-optimal under θ_2.

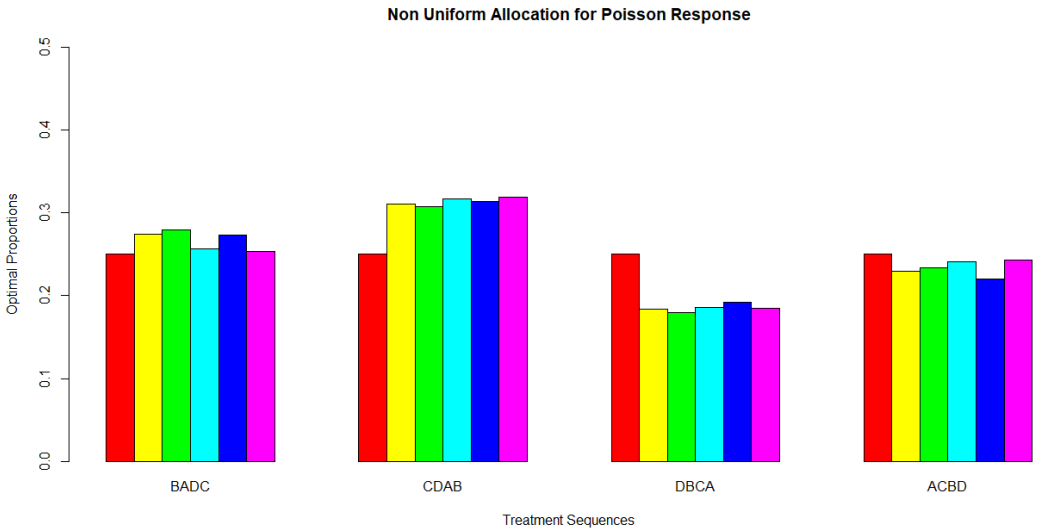


Figure 2: Non-uniform optimal proportions for Poisson response under θ_2

5. Beta Regression

In the beta response case, we calculate the locally optimal design for the response from the Booking.com example under two different models. We consider two different link functions to model the marginal mean of the response as follows:

$$\text{logit}(\mu_{ij}) = \log(\frac{\mu_{ij}}{1-\mu_{ij}}) = \eta_{ij} = \lambda + \beta_i + \tau_{d(i,j)} + \rho_{d(i-1,j)},$$

and,

$$\log(\mu_{ij}) = \eta_{ij} = \lambda + \beta_i + \tau_{d(i,j)} + \rho_{d(i-1,j)},$$

where notations have the same meaning as in equation (2).

To illustrate the optimal proportions in the beta response case, we consider the normalized response *engagement* from the work environment experiment. In the case of this experiment, *engagement* is a measure of the extent to which participants felt focused on and excited to complete regular work tasks.

5.1. Analysis of data

Similar to the Poisson response analysis, we consider three main predictors in the model for a beta response which are *area*, *wave* and *carryover*. We use six different kinds of correlation matrices as mentioned above and calculate optimal proportions under two different models with different link functions. As mentioned earlier, we consider baseline constraints so that all the parameters are estimable.

We get the initial estimates of parameters by fitting the beta regression model to the response. For two different link functions we need to guess two different sets of parameter values for θ_1 and θ_2. In case of *logit* link function, our nominal guess for the parameter values is $\theta_1 = [1.24, -0.035, 0.17, 0.078, -0.2, -0.3, 0.01, -0.35, -0.62, -0.329]$ and $\theta_2 = [1.24, -0.035, 0.17, 0.078, -4, -6, 2, -3.5, -3.1, -1.28]$. In case of *log* link function, our nominal guess for the parameter values is $\theta_1 = [-0.25, -0.01, 0.04, 0.02, -0.05, -0.08, -0.004, -0.088, -0.172, -0.08]$ and $\theta_2 = [-0.25, -0.01, 0.04, 0.02, -5, -8, -0.4, -2.2, -4.3, -2]$. Note that, as before, θ_1 is an educated guess based on the data at hand, whereas θ_2 has significantly different values for the parameters of interest than that of θ_1.

5.2. Optimal designs

In the Table 3, we present the optimal proportions corresponding to *logit* link case for six different choices of correlation matrix. As seen from Table 3, in case of beta response (*logit* link) the optimal proportions that we obtain using θ_1 are nearly uniform and that using θ_2 are non-uniform.

Table 3: Optimal proportions in case of beta response (*logit* link).

Correlation	θ_1				θ_2			
Structure	$BADC$	$CDAB$	$DBCA$	$ACBD$	$BADC$	$CDAB$	$DBCA$	$ACBD$
$Corr(1)$	0.2518	0.2563	0.2465	0.2454	0.3418	0.2085	0.1643	0.2854
$Corr(2)$	0.2525	0.2572	0.2453	0.2450	0.3316	0.2066	0.1690	0.2928
$Corr(3)$	0.2515	0.2568	0.2462	0.2455	0.3363	0.2058	0.1682	0.2897
$Corr(4)$	0.2405	0.2539	0.2419	0.2637	0.3205	0.2043	0.1739	0.3013
$Corr(5)$	0.2595	0.2542	0.2467	0.2396	0.3250	0.2070	0.1711	0.2969
$Corr(6)$	0.2366	0.2562	0.2423	0.2649	0.3218	0.2088	0.1668	0.3026

In Table 4, we present the optimal proportions corresponding to the *log* link case for six different choices of the correlation matrix. As before, in the beta response (*log* link) case, the optimal proportions that we obtain using θ_1 are nearly uniform and that using θ_2 are non-uniform.

The plots in Figures 3, 4 and Figures 5, 6 represent the optimal proportions for beta response under θ_1 and θ_2 for two different choices of link functions respectively.

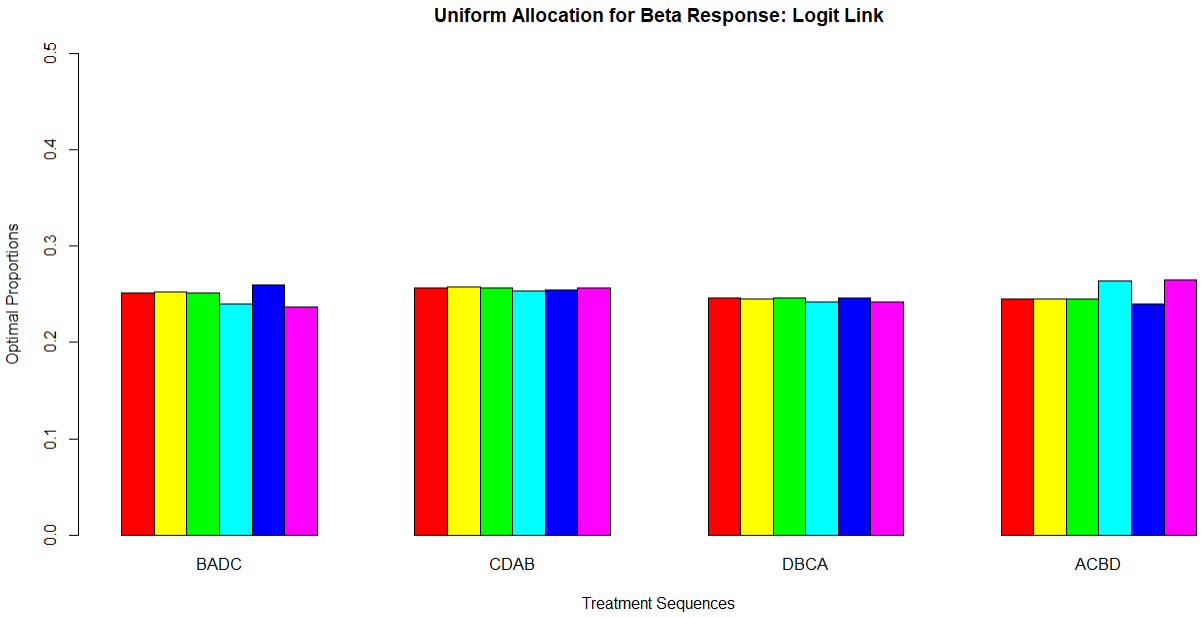


Figure 3: Uniform optimal proportions for beta response (*logit* link) under θ_1

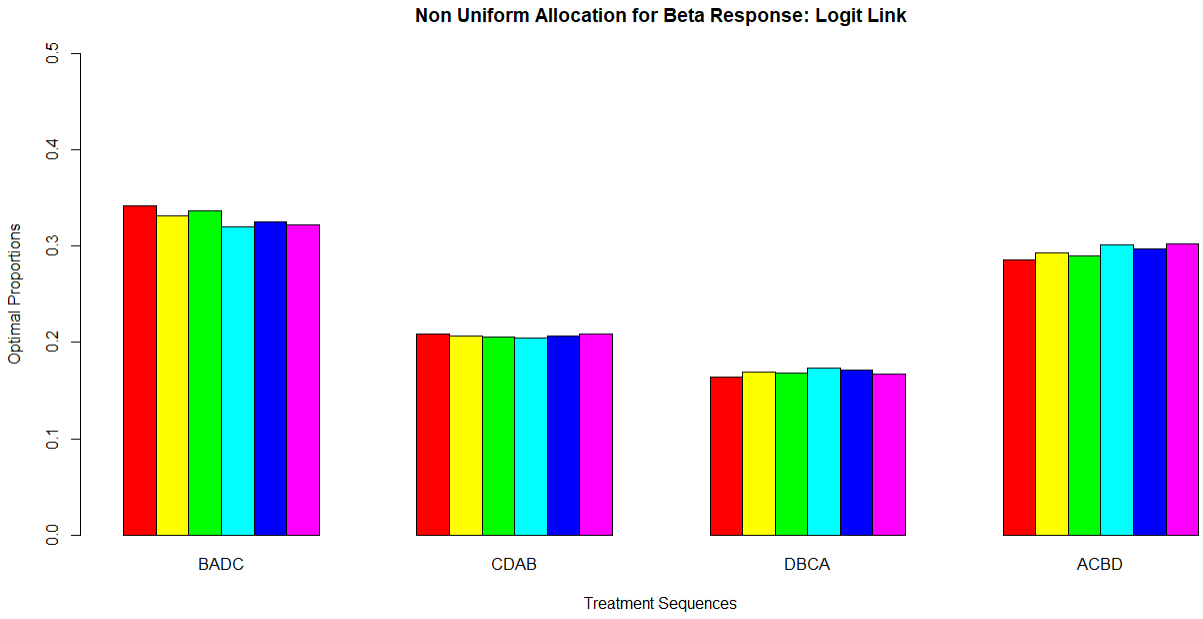


Figure 4: Non-uniform optimal proportions for beta response (*logit* link) under θ_2

Table 4: Optimal proportions in case of beta response (*log* link).

Correlation Structure	θ_1				θ_2			
	$BADC$	$CDAB$	$DBCA$	$ACBD$	$BADC$	$CDAB$	$DBCA$	$ACBD$
$Corr(1)$	0.2522	0.2560	0.2470	0.2448	0.3305	0.1470	0.1930	0.3295
$Corr(2)$	0.2529	0.2569	0.2458	0.2444	0.3270	0.1200	0.2084	0.3446
$Corr(3)$	0.2520	0.2564	0.2466	0.2450	0.3290	0.1210	0.2050	0.3450
$Corr(4)$	0.2410	0.2535	0.2425	0.2630	0.3271	0.1060	0.2137	0.3532
$Corr(5)$	0.2600	0.2540	0.2460	0.2400	0.3245	0.1101	0.2102	0.3552
$Corr(6)$	0.2371	0.2558	0.2428	0.2643	0.3272	0.1096	0.2120	0.3512

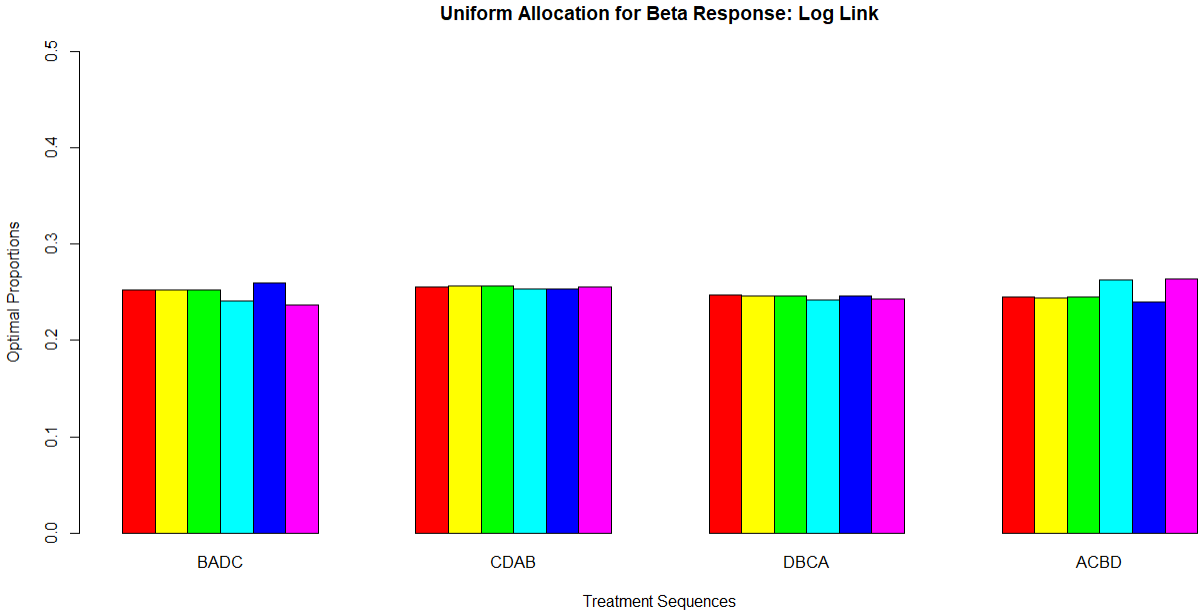


Figure 5: Uniform optimal proportions for beta response (*log* link) under θ_1

It can be seen from these plots that optimal proportions do not vary much when we use different correlation structures under θ_1 and θ_2. In most of the situations in practice uniform optimal designs are used. The above analysis shows that those uniform designs are sub-optimal under θ_2 irrespective of what link function is used.

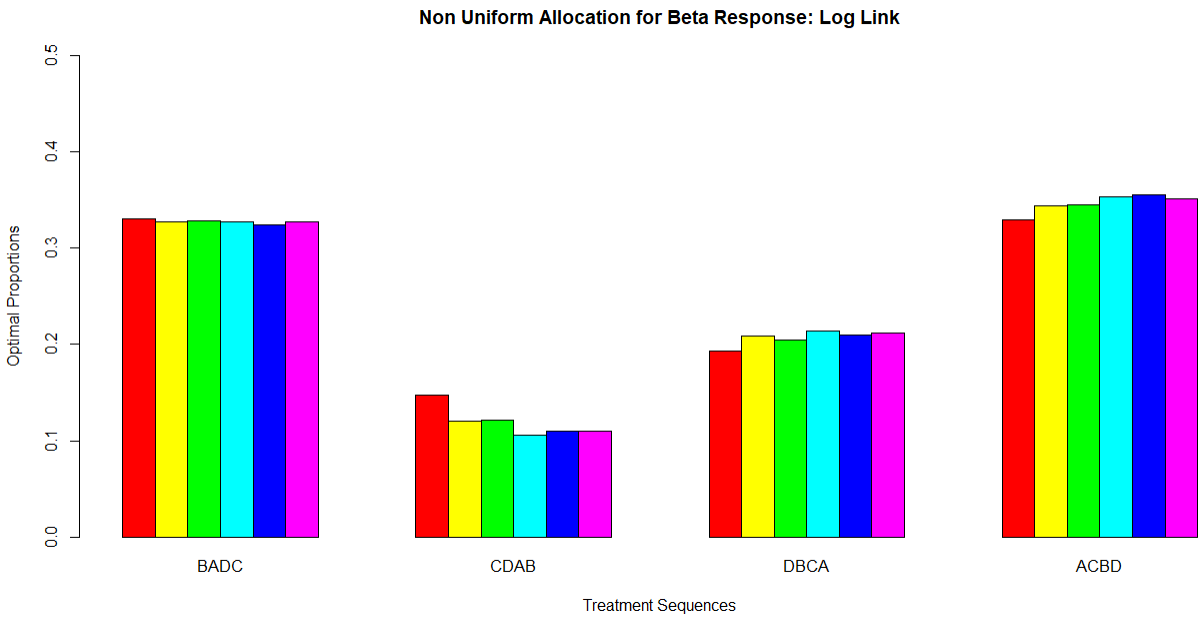


Figure 6: Non-uniform optimal proportions for beta response (*log* link) under θ_2

6. Gamma Regression

In the case of Gamma response, we calculate locally D-optimal design for the response from the same Booking.com example under two different models. Similar to the beta response, we consider two different link functions to model the marginal mean of the response. We use the *log*, and *inverse* link functions, and the two models are as follows:

$$\log(\mu_{ij}) = \eta_{ij} = \lambda + \beta_i + \tau_{d(i,j)} + \rho_{d(i-1,j)},$$

and,

$$inv(\mu_{ij}) = \frac{1}{\mu_{ij}} = \eta_{ij} = \lambda + \beta_i + \tau_{d(i,j)} + \rho_{d(i-1,j)},$$

where, as before, notations have the same meaning as in equation (2).

From the work environment experiment, we consider the response $satisfaction$. Satisfaction is an essential concept for organisational and office design research, and it is usually used to measure employees' sentiments. In the work environment experiment, the Leesman satisfaction index was used, which is useful for many benchmark purposes. Since the response is right-skewed, it is safe to assume that the response follows a gamma distribution.

6.1. Analysis of data

Similar to previous two cases, we consider three main predictors in the model for gamma response which are *area*, *wave* and *carryover*. As before, we consider six different kinds of correlation matrices and calculate optimal proportions under two different models with different link functions. We consider same baseline constraints as mentioned earlier. We fit the gamma regression model to the data with $satisfaction$ as response by using the `glm` function in **R** and calculate the parameter estimates.

In case of *log* link function, our nominal guess for the parameter values is $\theta_1 = [2.1, -0.19, -0.04, -0.04, -0.16, -0.4, -0.06, 0.05, 0.005, -0.05]$ and $\theta_2 = [2.1, -0.19, -0.04, -0.04, -1.6, -4.0, -0.6, 0.5, 0.05, -0.5]$. In case of *inverse* link function, our nominal guess for the parameter values is $\theta_1 = [0.13, 0.03, 0.003, 0.003, 0.025, 0.07, 0.008, -0.007, -0.0001, -0.01]$ and $\theta_2 = [0.13, 0.03, 0.003, 0.003, 2.5, 7, 0.8, -0.7, -0.01, -1]$. As before, θ_1 was motivated by the data provided by Pitchforth *et al.* (20202) and θ_2 is significantly different from θ_1.

6.2. Optimal designs

In the Table 5, we present the optimal proportions corresponding to *log* link case for six different choices of correlation matrix. As seen from Table 5, in case of gamma response (*log* link) the optimal proportions that we obtain using θ_1 are nearly uniform and that using θ_2 are non-uniform.

Table 5: Optimal proportions in case of gamma response (*log* link).

Correlation	θ_1				θ_2			
Structure	$BADC$	$CDAB$	$DBCA$	$ACBD$	$BADC$	$CDAB$	$DBCA$	$ACBD$
$Corr(1)$	0.2500	0.2500	0.2500	0.2500	0.1328	0.2775	0.3336	0.2561
$Corr(2)$	0.2500	0.2500	0.2500	0.2500	0.1248	0.2639	0.3527	0.2586
$Corr(3)$	0.2500	0.2500	0.2500	0.2500	0.1258	0.2582	0.3596	0.2564
$Corr(4)$	0.2500	0.2500	0.2500	0.2500	0.1206	0.2671	0.3451	0.2672
$Corr(5)$	0.2500	0.2500	0.2500	0.2500	0.1225	0.2770	0.3354	0.2656
$Corr(6)$	0.2500	0.2500	0.2500	0.2500	0.1195	0.2685	0.3416	0.2704

In Table 6, we present the optimal proportions corresponding to *inverse* link case for six different choices of correlation matrix. As before, Table 6 indicates that the optimal proportions that we obtain using θ_1 are nearly uniform and that using θ_2 are non-uniform in case of gamma response (*inverse* link).

The plots in Figures 7, 8 and Figures 9, 10 represent the optimal proportions for gamma response under θ_1 and θ_2 for two different choices of link functions respectively.

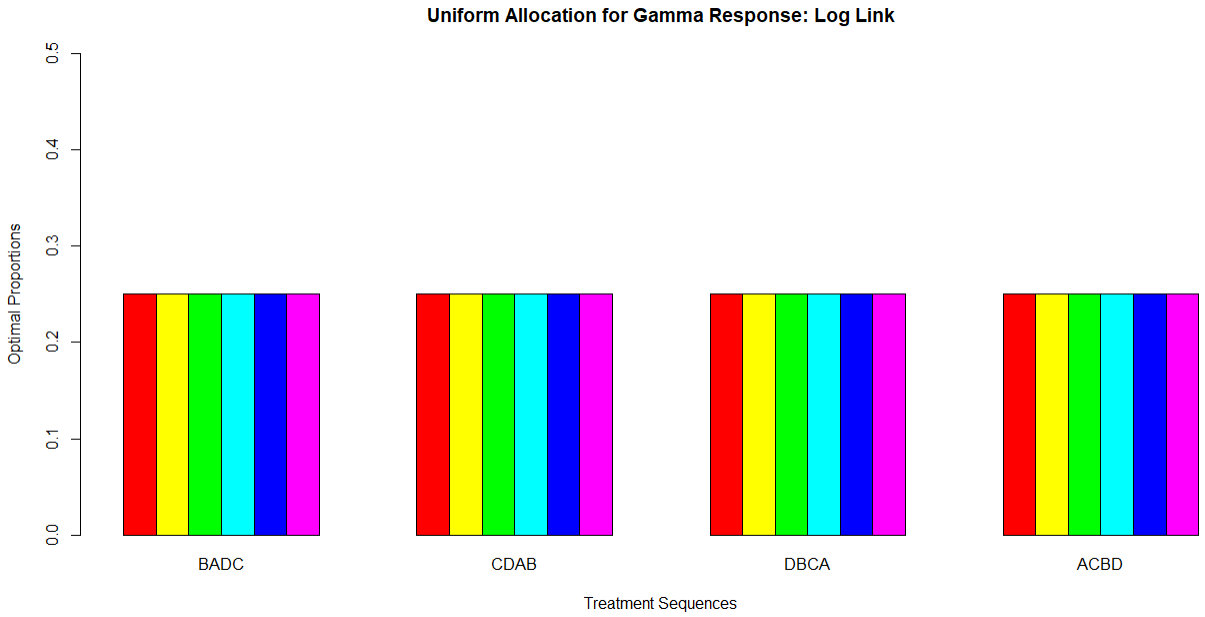


Figure 7: Uniform optimal proportions for gamma response (log link) under θ_1

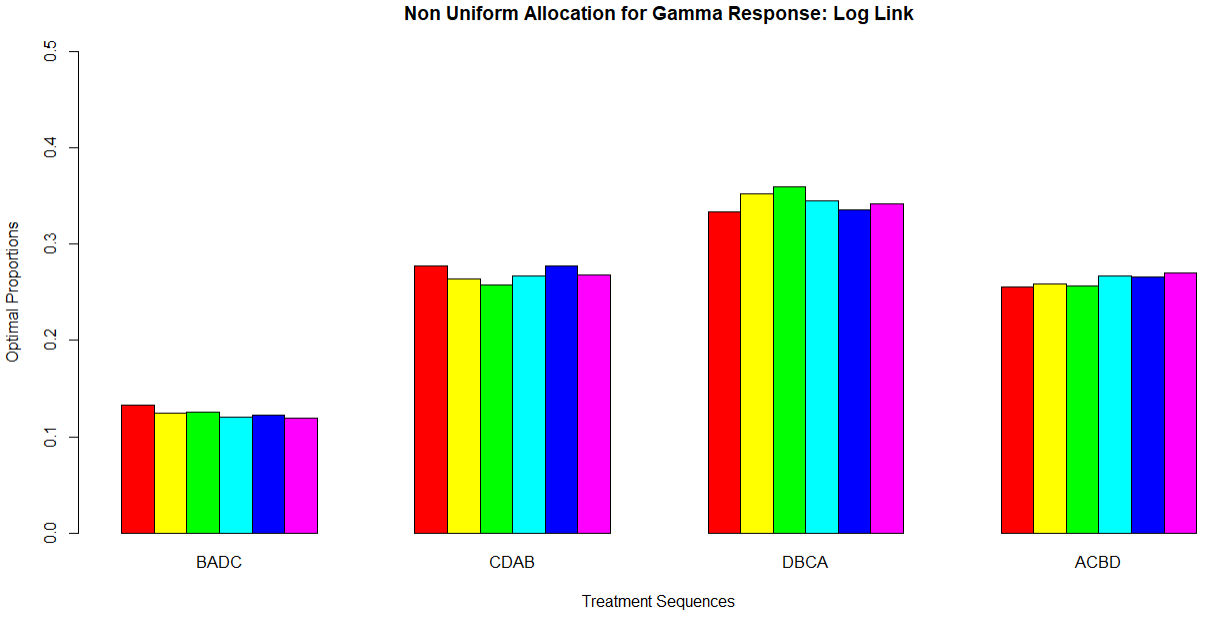


Figure 8: Non-uniform optimal proportions for gamma response (log link) under θ_2

Table 6: Optimal proportions in case of gamma response (*inverse* link).

Correlation Structure	θ_1				θ_2			
	$BADC$	$CDAB$	$DBCA$	$ACBD$	$BADC$	$CDAB$	$DBCA$	$ACBD$
$Corr(1)$	0.2500	0.2500	0.2500	0.2500	0.2650	0.3093	0.1828	0.2429
$Corr(2)$	0.2500	0.2500	0.2500	0.2500	0.2486	0.3031	0.1911	0.2572
$Corr(3)$	0.2500	0.2500	0.2500	0.2500	0.2588	0.3051	0.1879	0.2482
$Corr(4)$	0.2500	0.2500	0.2500	0.2500	0.2389	0.3087	0.1784	0.2740
$Corr(5)$	0.2500	0.2500	0.2500	0.2500	0.2406	0.3112	0.1762	0.2720
$Corr(6)$	0.2500	0.2500	0.2500	0.2500	0.2421	0.3146	0.1740	0.2729

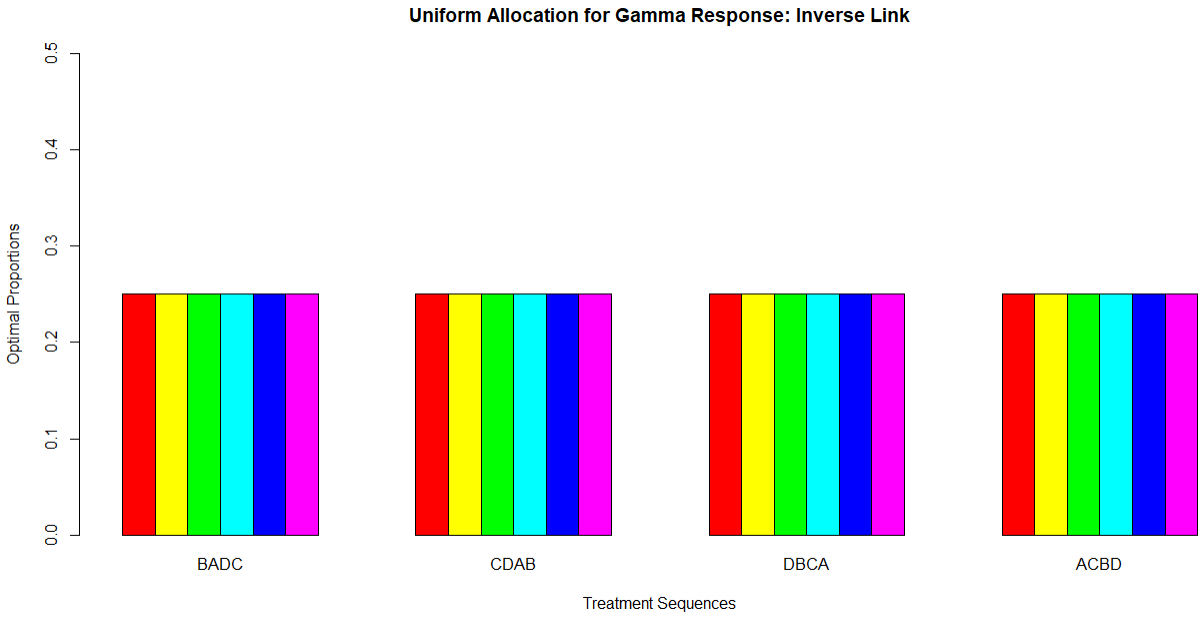


Figure 9: Uniform optimal proportions for gamma response (*inv* link) under θ_1

It can be seen from these plots that optimal proportions do not vary much when we use different correlation structures under θ_1 and θ_2. In most of the situations in practice uniform optimal designs are used. The above analysis shows that those uniform designs are sub-optimal under θ_2 irrespective of what link function is used.

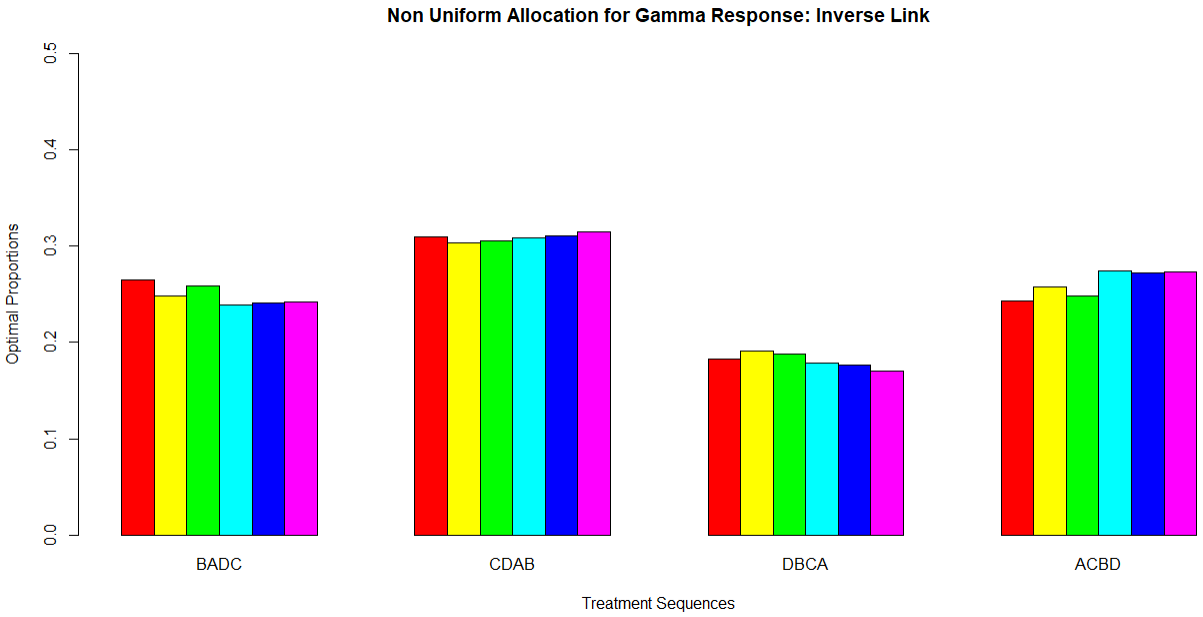


Figure 10: Non-uniform optimal proportions for gamma response (*inv* link) under θ_2

7. Summary and Conclusion

In many experiments in real life, uniform designs are often used. Uniform designs are those in which the same number of subjects are assigned to each treatment sequence. These uniform designs are optimal in the linear model case, *i.e.* when the response is normally distributed. But, in situations where responses are non-normal, the obtained optimal proportions are not necessarily uniform. In this paper's analysis, we identify locally optimal designs for responses belonging to Poisson, beta and gamma distributions. Two different link functions were considered in the case of beta and gamma responses. Tables 2 to 6 and plots in Figures 1 to 10 suggest that obtained optimal proportions are robust for different choice of correlations structures. These results also suggest that uniform designs are sub-optimal under θ_2 irrespective of the link function used or the response's distribution. Note that we are using the local optimality approach of Chernoff (1953). In real experiments, it is not always possible to guess the values of parameter estimates from prior knowledge. In that case, it is not easy to obtain locally optimal designs. In this paper we consider approximate designs in terms of optimal proportions. While conducting real life experiments, the practitioners must use exact designs where these proportions are to be converted into integers for determining the replication numbers of the sequences. The rounding off error might be significant unless the total number of observations is large. The Work Environment Experiment had 288 subjects and hence such issues do not arise.

References

Bose, M. and Dey, A. (2009). *Optimal Crossover Designs*. World Scientific.

Carriere, K. C. and Huang, R. (2000). Crossover designs for two-treatment clinical trials. *Journal of Statistical Planning and Inference*, **87**, 125−134.

Cheng, C. S. and Wu, C. F. J (1980). Balanced repeated measurements designs. *Annals of Statistics*, **8**, 1272−1283.

Chernoff, H. (1953) Locally optimal designs for estimating parameters. *Annals of Mathematical Statistics*, **24**, 586−602.

Dey, A., Gupta, V. K. and Singh, M. (1983). Optimal change-over designs. *Sankhya*, **B45**, 233−239.

Hedayat, A. and K. Afsarinejad (1978). Repeated measurements designs, I. In *A Survey of Statistical Designs and Linear Models (J. N. Srivastava,* ed.), pp. 229−242. Amsterdam, North-Holland.

Jankar, J., Mandal, A. and Yang, J. (2020). Optimal Crossover Designs for Generalized Linear Models. *Journal of Statistical Theory and Practice*, **14**, 23.

Jones, B. and Kenward, M. G. (2014). *Design and Analysis of Cross-over Trials*, 3rd ed. Chapman and Hall, London.

Kershner, R. P. and Federer, W. T. (1981). Two-treatment crossover designs for estimating a variety of effects. *Journal of American Statistical Association*, **76**, 612−619.

Kunert, J. (1984b). Optimality of balanced uniform repeated measurements designs. *Annals of Statistics*, **12**, 1006−1017.

Laska, E. and Meisner, M. (1985). A variational approach to optimal two treatment crossover designs: application to carryover effect models. *Journal of American Statistical Association*, **80**, 704−710.

Matthews, J. N. S. (1987). Recent developments in crossover designs. *International Statistical Review*, **56**, 117−127.

McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models,* 2nd ed. Chapman and Hall, London.

Pitchforth, J., Nelson-White, E., van den Helder, M., Oosting, W. (2020) The work environment pilot: An experiment to determine the optimal office design for a technology company. *PLoS ONE*, **15**, 5.

Pukelsheim, F. (1993). *Optimal Design of Experiments.* Wiley, New York..

Senn, S. (2003). *Cross-over Trials in Clinical Research,* 2nd ed. Wiley, Chichester, England.

Stufken, J. and Yang, M. (2012). Optimal designs for generalized linear models. In *Design and Analysis of Experiments*, Volume 3: Special Designs and Applications, (Edited by K.Hinkelmann). Wiley, New York.

Wu, C. F. J. and Hamada, M. (2009) Experiments: Planning, Analysis, and Optimization, 2nd Edition. Wiley, New York.

Zeger, S. L., Liang, K. Y.and Albert, P. S. (1988). Models for longitudinal data: A generalized estimating equation approach. *Biometrics* **44**, 1049−1060.

APPENDIX

Six Different Correlation Structures

The first correlation structure is a compound symmetric correlation structure, *i.e.*,

$$Corr(1) = (1-\rho)I_p + \rho J_p,$$

where I_p is the identity matrix of order p, and J_p is a $p \times p$ matrix with all elements unity.

The second correlation structure is the AR(1) correlation structure, *i.e.*,

$$Corr(2) = \left(\rho^{|i-i'|}\right),$$

so that the correlation between responses decreases as the time gap between responses increases.

The third correlation structure is as follows:

$$Corr(3) = \begin{pmatrix} 1 & \rho & 0 & \dots & 0 & 0 & 0 \\ \rho & 1 & \rho & \dots & 0 & 0 & 0 \\ \vdots & & & \vdots & & & \vdots \\ 0 & 0 & 0 & \dots & \rho & 1 & \rho \\ 0 & 0 & 0 & \dots & 0 & \rho & 1 \end{pmatrix}.$$

For each correlation structure different correlation matrices using different ρ values are considered.

To understand the other three correlation structures, we denote the correlation coefficient between the response when a subject receives treatment A first and the response when the same subject receives treatment B afterwards as ρ_{AB} and ρ_{BA} when the subject receives B first and A afterwards. Note that in general, ρ_{AB} is not necessarily the same as ρ_{BA}. In a similar manner we define ρ_{AA} and ρ_{BB}. To define the fourth type of correlation structure, we will use the same structure as $Corr(3)$ but with different values of the correlation coefficient for different treatment sequences.

To define the fifth and sixth type of correlation structures, we use AR(1) correlation structure with correlation coefficient depending on treatment sequence. For the fifth type, we use the same values for ρ_{AB} and ρ_{BA}, and for the sixth type of correlation structure, we use different values for ρ_{AB} and ρ_{BA}. For both fifth and sixth type of correlation structure we keep $\rho_{AA} = \rho_{BB}$. These values might vary from example to example and depend on what treatments A and B are. As the correlation matrix entries depend on which treatment the subject receives in a particular period, these correlation matrices are different for different treatment sequences.

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 337–365

On the Interrelationships between Lomax, Pareto and Exponentiality with Certain Extensions

Zhixin Lun[1] **and Ravindra Khattree**[2]
[1]*Department of Clinical Pharmacy*
University of California, San Francisco, California, USA

[2]*Department of Mathematics and Statistics and*
Center for Data Science and Big Data Analytics
Oakland University, Rochester, Michigan, USA

Received: 31 October 2020; Revised: 08 March 2021; Accepted: 11 March 2021

Abstract

We start with a recently published connection from Lomax to exponential distribution through the limiting distribution of a Lomax distribution scaled by its shape parameter. Motivated by that observation, we explore several relationships between (generalized) Lomax and other distributions of exponential family such as Gamma, Beta type II, Rayleigh and Weibull. As further extension, we introduce the generalized double Pareto distribution on the entire real line. Various properties of generalized double Pareto distribution are then studied including its representation as a mixture of Student's t and its connection to Laplace (double exponential) distribution. We then provide a simple approach to simulate random numbers from the double Pareto distribution and its implementation in R. Finally, we illustrate an application in a real biomedical research problem.

Key words: Double Pareto distribution; Exponential distribution; Khattree-Bahuguna skewness; Laplace distribution; Lomax distribution; Multivariate Lomax distribution.

AMS Subject Classifications: 62E10, 62E15, 62H05

1. Introduction

It is our distinct privilege to dedicate this paper in honor of Late Prof. Aloke Dey. One of us has learned a lot and personally benefited greatly in his research as well as in teaching from Prof. Dey's landmark book on block designs and although later he never got a chance to thank Prof. Dey in person, we hope our this article serves as our symbolic appreciation of Prof. Dey's contributions to the goal of advancing the knowledge for the betterment of society.

This article came into being due to an earlier simple curiosity described in Lun and Khattree (2020) about univariate and multivariate Lomax distributions. The relationship between Pareto/Lomax and exponential distributions has been well recorded in the literature (see Johnson, Kotz, and Balakrishnan (1994) and Kotz, Kozubowski and Podgórski (2001))

Corresponding Author: Ravindra Khattree
Email: khattree@oakland.edu

yet has not been very well publicized. Harris (1968) provides an approach to generate the Pareto variates through a mixture of exponential variates with parameter having a gamma distribution. Specifically and more generally, let X follow an exponential distribution with rate parameter η and allow η to have a gamma distribution with shape parameter β and scale parameter θ. Then, the unconditional probability density function of X is

$$f(x) = \int_0^\infty \eta e^{-\eta x} \frac{1}{\Gamma(\beta)\theta^\beta} \eta^{\beta-1} e^{-\eta/\theta} d\eta = \frac{\theta\beta}{(1+\theta x)^{\beta+1}}, \quad x > 0, \theta, \beta > 0, \tag{1}$$

which is the density function of Lomax (also called Pareto type II) distribution (hereon denoted by Lomax(β, θ)) with shape parameter β and rate parameter θ. Alternatively, given two independent standard exponential variates W_1 and W_2, the probability density function of $Y = W_1/W_2$ is a standard Lomax distribution (with $\theta = 1, \beta = 1$). On the other hand, assume that X has a Pareto type I distribution with density $g(x) = \lambda x^{-(\lambda+1)}$, $x \geq 1$, then the density function of $Y = \log X$ is $\lambda e^{-\lambda y}, y > 0$. To indicate the other connections, in Table 1, we summarize the expectations, variances, Pearson's coefficients of skewness and Khattree-Bahuguna's skewness [1] of Lomax and exponential distributions.

As $\beta \to \infty$, for Lomax$(\beta, \frac{\theta}{\beta})$, we observe that the variance approaches the square of expectation of an Exponential distribution with the rate parameter θ and the Pearson's coefficient of skewness approaches to 2. As we shall see later, as $\beta \to \infty$ the Khattree-Bahuguna's skewness of the above Lomax distribution also approaches that of exponential distribution. Proof is given in Section 2.3. A natural question then arises as to why does Lomax distribution have its distributional properties so similar to those of exponential distribution?

One of our main objectives in this article is to explore above question by showing and generalizing the connections between Lomax and exponential distributions. These have been motivated by a result by Lun and Khattree (2020) about the limiting distribution of a Lomax random variable scaled by its shape parameter. Based on this result, we discover many more results which establish many other connections between (generalized) Lomax and other distributions of exponential family. We further take upon introducing another connection between Laplace (double exponential) distribution and generalized double Pareto distribution, which is a popular choice of prior distribution in recent years for robust Bayesian shrinkage estimators.

The article is organized as follows. In Section 2, we show the Lomax-exponential and generalized Lomax-gamma connections. Then, we attempt to quantify the distance of Lomax to exponential distributions in terms of Patil-Patil-Bagkavos's η (2012) and by using the Kullback-Leibler divergence. In Section 3, we give some results pertaining to relationships between multivariate Lomax and other distributions of exponential family such as Gamma, Beta type II, Rayleigh and Weibull. In Section 4, we discuss a three-parameter generalized double Pareto distribution, including its simulation and connection to Laplace distribution. Section 5 includes an approach to simulation of generalized double Pareto variates. In Section 6, we give a real-world application of bivariate Lomax distribution. Section 7 includes some concluding remarks.

[1] Khattree-Bahuguna's skewness for a random variable is defined in Khattree and Bahuguna (2019) and reproduced here in Section 2.3. The values of this skewness for Lomax and Exponential distributions are also computed there.

Table 1: Comparison of summary parameters for Lomax(β, θ) and Exp(λ)

Summary Parameter	Lomax(β, θ)	Exp(λ)
E(X)	$\frac{1}{\theta(\beta-1)}, \quad \beta > 1$	$\frac{1}{\lambda}$
Var(X)	$\frac{\beta}{\theta^2(\beta-1)^2(\beta-2)}, \quad \beta > 2$	$\frac{1}{\lambda^2}$
Pearson's coefficient of skewness	$\frac{2(\beta+1)}{\beta-3}\sqrt{\frac{\beta-2}{\beta}}, \quad \beta > 3$	2
Khattree-Bahuguna's skewness	$\frac{1}{2}\left[1 + \left(B(1-\frac{1}{\beta}, 1-\frac{1}{\beta}) - \frac{\beta^2}{(\beta-1)^2}\right)\frac{(\beta-1)^2(\beta-2)}{\beta}\right], \quad \beta > 2$	$1 - \frac{\pi^2}{12} \simeq 0.177533$

2. Connections Between Generalized Lomax and Gamma Distributions

In this section, we first state a surprisingly simple connection between Lomax and exponential distributions partially given in Lun and Khattree (2020). This result is then generalized to connect the generalized Lomax with Gamma distribution in a straight forward way. To begin with, we first address the limiting distributions.

2.1. Limiting distributions

The following two theorems about limiting distributions can be easily proved.

Theorem 1: Let X be a univariate Lomax(β, θ) random variable with probability density function defined in (1). Define $Y = \beta X$. Then

(i) the distribution of Y is Lomax$(\beta, \frac{\theta}{\beta})$;

(ii) as $\beta \to \infty$, the distribution of Y approaches an exponential distribution with rate parameter θ;

(iii) as $\beta \to \infty$, the hazard function of Y approaches θ.

The result (ii) is given in Lun and Khattree (2020). It may be pointed out that an exponential distribution is characterized by its constant hazard function. Thus, (iii) in fact, provides an alternative proof of (ii). We leave it to the reader to calculate the hazard function of the indicated Lomax distribution and then verify the assertion in (iii). The importance of above theorem is that it allows us to be able to conveniently substitute, for large β, one distribution for another by approximating Lomax$(\beta, \frac{\theta}{\beta})$ by Exp(θ). Figure 1 shows a series of density function plots for Lomax$(\beta, \frac{\lambda}{\beta})$ for parameter $\beta = 3.01, 10, 20$ and for a fixed $\lambda = 0.25$, along with an exponential density with rate parameter λ. The closeness of two distributions for large β values is self evident.

Nayak (1987) has introduced a k-dimensional multivariate Lomax distribution by mixing k independent univariate exponential distributions with different failure rates with the mixing parameter η that has a gamma distribution with certain shape parameter β and the scale parameter 1. The Theorem 1 is easily extended to connect the multivariate Lomax and exponential distributions. Again, see Lun and Khattree (2020).

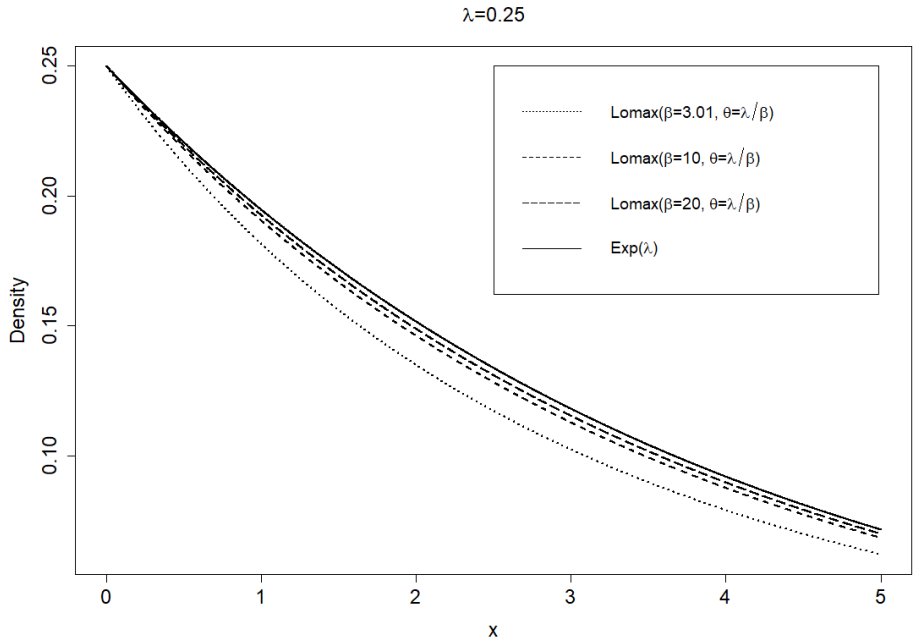


Figure 1: Density plots of a series of Lomax density for various β values ($= 3.01, 10, 20$) and with $\theta = \lambda/\beta$ where $\lambda = 0.25$, along with exponential density with rate parameter $\lambda = 0.25$.

Theorem 2: Let $X_1, X_2, \ldots, X_k$ jointly have k-dimensional multivariate distribution with probability density function (Nayak, 1987),

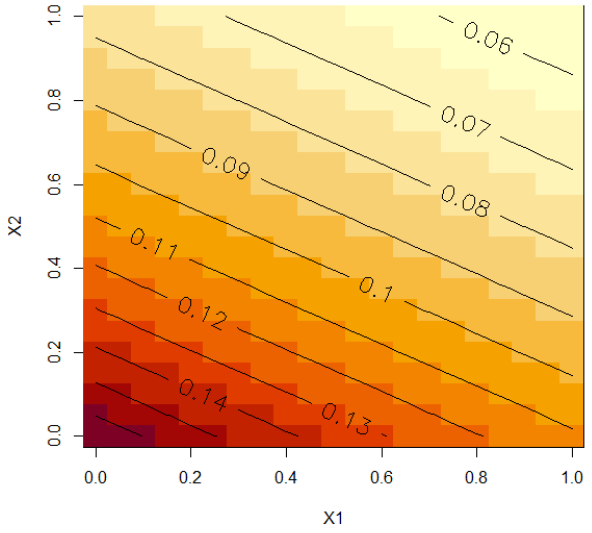
$$f(x_1, x_2, \cdots, x_k) = \frac{\left(\prod_{i=1}^{k} \theta_i\right) \prod_{i=1}^{k} (\beta + i - 1)}{\left(1 + \sum_{i=1}^{k} \theta_i x_i\right)^{\beta+k}}, \quad \beta > 0, \tag{2}$$

where $\theta_i, x_i > 0, i = 1, \cdots, k$. Define $Y_i = \beta X_i, i = 1, \cdots, k$. Then as $\beta \to \infty$, the joint probability distribution of $Y_1, Y_2, \ldots, Y_k$ approaches to that of k independent exponential variates with rate parameters $\theta_i, i = 1, 2, \cdots, k$, respectively.

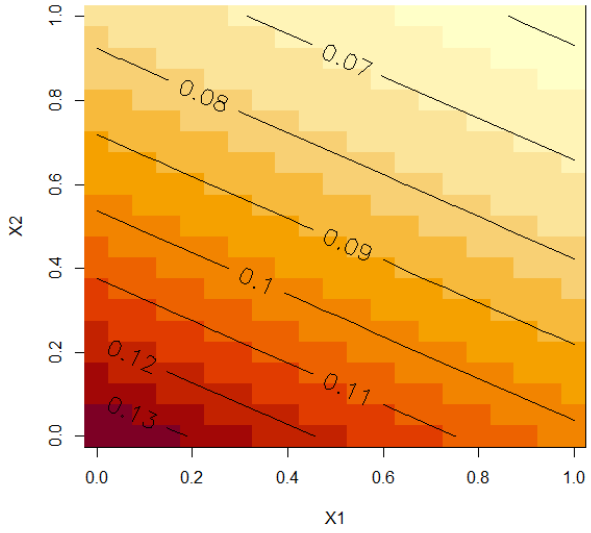
A series of density contour plots for bivariate Lomax distributions with parameters β $(= 3.01, 10, 200)$ and $\theta_1 = \lambda_1/\beta, \theta_2 = \lambda_2/\beta$ where $\lambda_1 = 0.25, \lambda_2 = 0.50$ are shown in Figure 2 (a), (b) and (c). As one can see, as β increases, the density contour plots of bivariate Lomax distribution more and more resemble the bivariate independent exponential density contour plot given in Figure 2 (d) with respective rate parameters $\lambda_1 = 0.25, \lambda_2 = 0.50$.

Note that the resulting limiting distribution involves independent exponential variates. In some ways, this is somewhat surprising even though the correlation matrix of a multivariate Lomax distribution has a compound symmetric structure with $\text{Corr}(X_i, X_i) = \frac{1}{\beta}$ which clearly goes to 0 when $\beta \to \infty$. On the other hand, this fact also underscores the well known difficulty that researchers have encountered to satisfactorily define a suitable multivariate exponential distribution with some kind of dependence among variables.

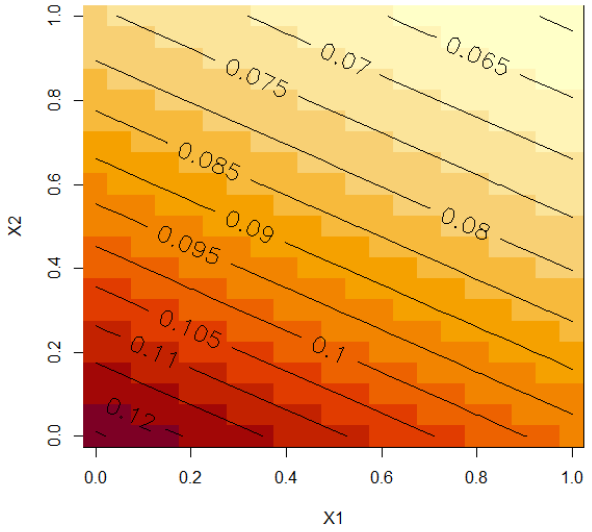
Nayak (1987) also generalized the distribution in (2) by mixing conditionally independent X_i having the Gamma$(l_i, \eta\theta_i)$ distribution, with a mixing variable $\eta \sim$ Gamma$(\beta, 1)$,



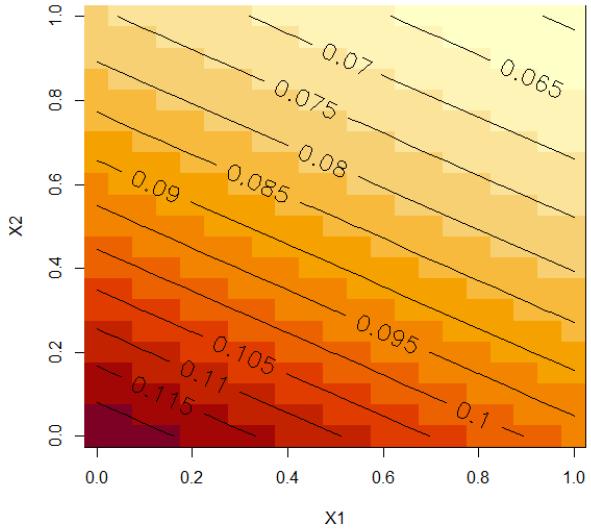
(a) Bivariate Lomax $(\beta = 3.01, \theta_1 = \frac{\lambda_1}{\beta}, \theta_2 = \frac{\lambda_2}{\beta})$



(b) Bivariate Lomax $(\beta = 10, \theta_1 = \frac{\lambda_1}{\beta}, \theta_2 = \frac{\lambda_2}{\beta})$



(c) Bivariate Lomax $(\beta = 200, \theta_1 = \frac{\lambda_1}{\beta}, \theta_2 = \frac{\lambda_2}{\beta})$



(d) Independent Bivariate Exponential (λ_1, λ_2)

Figure 2: Contour plots of a series of bivariate Lomax distributions with changing parameters $\beta, \theta_1 = \lambda_1/\beta, \theta_2 = \lambda_2/\beta$ **where** $\beta = 3.01, 10, 200$ **and** $\lambda_1 = 0.25, \lambda_2 = 0.50$ **and independent bivariate exponential density function with rate parameters** $\lambda_1 = 0.25$ **and** $\lambda_2 = 0.50$**, respectively.**

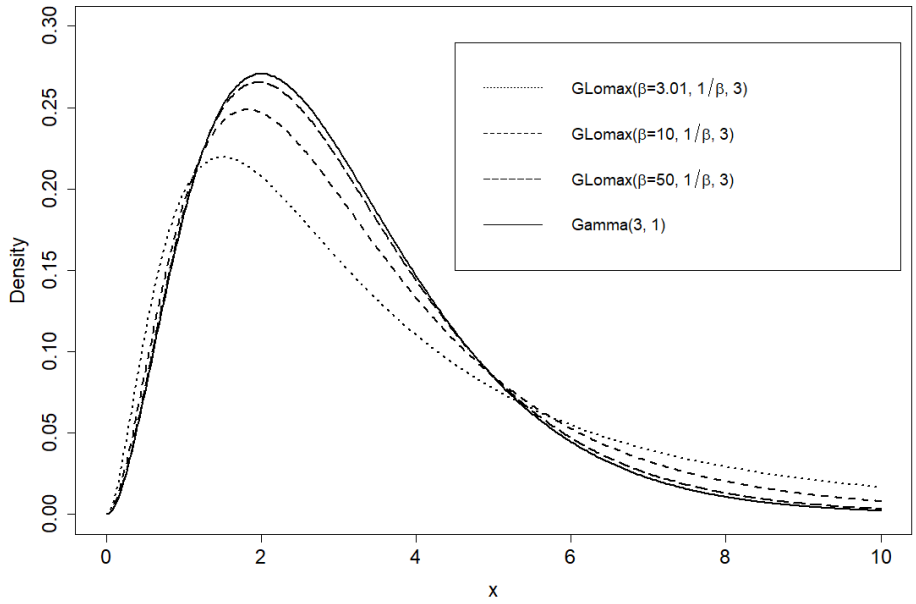


Figure 3: Density plots of a series of generalized Lomax distributions with changing parameters $\beta(= 3.01, 10, 50)$, $\theta = 1.0/\beta$ and a fixed parameter $l = 3.0$, and gamma density function with shape parameter $l = 3.0$ and rate parameter $\lambda = 1.0$.

$i = 1, \ldots, k$. The generalization of above multivariate Lomax-exponential connection to this generalized multivariate Lomax-gamma connection is stated in Theorem 3.

Theorem 3: Let $X_1, X_2, \ldots, X_k$ be a k-dimensional generalized multivariate Lomax random variable with probability density function,

$$f(x_1, x_2, \cdots, x_k) = \frac{\left(\prod_{i=1}^{k} \theta_i^{l_i}\right) \Gamma\left(\sum_{i=1}^{k} l_i + \beta\right) \left(\prod_{i=1}^{k} x_i^{l_i - 1}\right)}{\Gamma(\beta) \left[\prod_{i=1}^{k} \Gamma(l_i)\right] \left(1 + \sum_{i=1}^{k} \theta_i x_i\right)^{\sum_{i=1}^{k} l_i + \beta}},$$

where $\beta, l_i, \theta_i, x_i > 0, i = 1, \cdots, k$. Define $Y_i = \beta X_i, i = 1, \cdots, k$. Then as $\beta \to \infty$, the joint distribution of $Y_1, Y_2, \ldots, Y_k$ approaches that of k independent gamma random variables with shape parameter l_i and rate parameters θ_i, $i = 1, 2, \cdots, k$, respectively.

Clearly, in the special case of $l_i = 1, i = 1, \cdots, k$, the above generalized Lomax-gamma connection reduces to the previous Lomax-exponential connection. Again, to underscores the closeness, we give the density function plots in Figure 3 for a series of univariate generalized Lomax distributions for various values of the parameters $\beta, \theta = \lambda/\beta$ where $\beta = 3.01, 10, 50$, $\lambda = 1.0$, and for a fixed parameter $l = 3.0$, along with Gamma(3,1). To avoid being repetitive, we suppress the contour plots.

2.2. Some measures of closeness to exponential distribution

We have shown that the similarity between Lomax and exponential is due to the fact that the limiting distribution of Lomax$(\beta, \frac{\theta}{\beta})$ is Exp(θ) as $\beta \to \infty$. That begs the

question: How to quantify the closeness between a given Lomax distribution and an exponential distribution? To do so, we adopt several approaches. Specifically, we first use Patil-Patil-Bagkavos's η (2012) to measure the closeness of Lomax to exponential and then employ the Kullback-Leibler divergence measure. Later In Section 2.3, we also evaluate their Khattree-Bahuguna skewnesses for this purpose. In order to do all of this, we first define the Patil-Patil-Bagkavos's η (2012).

2.2.1. Patil-Patil-Bagkavos's η

Patil, Patil and Bagkavos (2012) attempted to propose a measure of (a)symmetry of a random variable X as,

$$\eta = \begin{cases} -\text{Corr}(f(X), F(X)) & \text{if } 0 < \text{Var}(f(x)) < \infty, \\ 0 & \text{if } \text{Var}(f(x)) = 0, \end{cases}$$

where $f(x)$ and $F(x)$ are the probability density function and the cumulative distribution function of X, respectively. Clearly, $-1 \leq \eta \leq 1$. However, these authors incorrectly claimed that the η defined above is a measure of the degree of (a)symmetry of a distribution and hence can be used as a measure of skewness of a distribution. Eberl and Klar (2019) disputed their claim and via several examples, they demonstrated that above as a measure of asymmetry is indeed a misleading measure. They further pointed out that η is instead, a measure of the closeness of a given distribution to the exponential distribution and based on their extensive discussion, we readily agree! They further pointed out that η equal to zero indicates a complete departure from exponential distribution while values of $+1$ (-1) show a complete similarity with the positive (negative) exponential. Therefore, in our context, η can be deemed as a tailor-made measure to evaluate the closeness of a Lomax distribution to the exponential distribution. The following theorem gives an explicit expression for η for the Lomax distribution.

Theorem 4: For the Lomax distribution, the Patil-Patil-Bagkavos's η is given by, $\eta = \frac{\sqrt{3\beta(3\beta+2)}}{(3\beta+1)}$.

Proof:. Let X be a Lomax random variable with pdf given in (1). The cumulative distribution function of X is then given by $F(x) = 1 - \frac{1}{(1+\theta x)^{\beta}}$. The covariance between $f(x)$ and $F(x)$ is $\text{Cov}\left[f(X), F(X)\right] = \int_0^{\infty} \left[\frac{\theta\beta}{(1+\theta x)^{\beta+1}}\right]^2 \left[1 - \frac{1}{(1+\theta x)^{\beta}}\right] dx - \int_0^{\infty} \left[\frac{\theta\beta}{(1+\theta x)^{\beta+1}}\right]^2 dx \cdot \frac{1}{2} = -\frac{\theta\beta^2(\beta+1)}{2(2\beta+1)(3\beta+1)}$. Also, $\text{Var}(F(x)) = \frac{1}{12}$ and $\text{Var}(f(x)) = \frac{\theta^2\beta^3(\beta+1)^2}{(3\beta+2)(2\beta+1)^2}$. Thus,

$$\eta = -\text{Corr}(f(X), F(X)) = \frac{\frac{\theta\beta^2(\beta+1)}{2(2\beta+1)(3\beta+1)}}{\sqrt{\frac{\theta^2\beta^3(\beta+1)^2}{(3\beta+2)(2\beta+1)^2}}\sqrt{\frac{1}{12}}} = \frac{\sqrt{3\beta(3\beta+2)}}{(3\beta+1)}. \qquad \square$$

As one would anticipate, η does not depend on scale parameter θ since correlation is invariant of any such scaling.

Straight forward calculations show that even for β as small as 3, $\eta = 0.995$, which indicates that Lomax distribution is generally very similar to exponential distribution even for small β value. Clearly, as $\beta \to \infty$, $\eta \to 1$, thereby reaffirming the previous result about a β multiple of Lomax distribution converging to the exponential distribution .

2.2.2. Kullback-Leibler divergence

Kullback-Leibler divergence is a measure of how different a given probability distribution described by a probability density function $f(x)$ is from another reference distribution with the probability density function $g(x)$ and is defined as

$$D_{KL}(f:g) = \mathrm{E}_f\left(\log\frac{f(x)}{g(x)}\right) = \int_{-\infty}^{\infty} f(x)\log\left(\frac{f(x)}{g(x)}\right)dx.$$

A zero value indicates that the two distributions are identical. In our context, the following theorem delivers the quantification of closeness.

Theorem 5: Let $f(x) = \frac{\theta\beta}{(1+\theta x)^{\beta+1}}$ and $g(x) = \lambda e^{-\lambda x}$. Then the Kullback-Leibler divergence measure for (Lomax : Exponential) pair is given by, $D_{KL}(f:g) = \ln\left(\frac{\beta\theta}{\lambda}\right) - \frac{\beta+1}{\beta} + \frac{\lambda}{\theta(\beta-1)}$.

Proof:. With $f(x)$ and $g(x)$ as given above, we have,

$$\begin{aligned} D_{KL}(f:g) &= \int_0^{\infty} \frac{\theta\beta}{(1+\theta x)^{\beta+1}} \ln\left(\frac{\frac{\theta\beta}{(1+\theta x)^{\beta+1}}}{\lambda e^{-\lambda x}}\right) dx \\ &= \int_0^{\infty} \frac{\theta\beta}{(1+\theta x)^{\beta+1}} \ln\left(\frac{\theta\beta}{(1+\theta x)^{\beta+1}}\right) dx - \int_0^{\infty} \frac{\theta\beta}{(1+\theta x)^{\beta+1}}[\ln(\lambda) - \lambda x]dx \\ &= \ln(\beta\theta) - \frac{\beta+1}{\beta} - \ln(\lambda) + \frac{\lambda}{\theta(\beta-1)} = \ln\left(\frac{\beta\theta}{\lambda}\right) - \frac{\beta+1}{\beta} + \frac{\lambda}{\theta(\beta-1)}. \end{aligned}$$

□

The measure does depend on the ratio of scale parameters $\frac{\lambda}{\theta}$. When $\frac{\lambda}{\beta} = \theta$ and thus $f(x)$ is the pdf of a Lomax$(\beta, \frac{\lambda}{\beta})$, we have $D_{KL}(f:g) = \frac{\beta}{\beta-1} - \frac{\beta+1}{\beta} = \frac{1}{\beta(\beta-1)}$. Clearly, the convergence to zero is of order $\frac{1}{\beta^2}$ which again reaffirms our assertion of the considerable closeness of the two distributions.

2.3. Khattree-Bahuguna's skewness for Lomax/Pareto and exponential distributions

Khattree and Bahuguna (2019) recently defined a measure of skewness of a probability distribution which for a quick reference, we state below.

Definition 6: Let X be a random variable possibly assumed to have been centered by mean and let $F(\cdot)$ be its cumulative distribution function. The Khattree-Bahuguna's skewness of X is defined as

$$\delta = \frac{\int_0^1 \left(\frac{F^{-1}(\alpha)+F^{-1}(1-\alpha)}{2}\right)^2 d\alpha}{\int_0^1 \left(\frac{F^{-1}(\alpha)+F^{-1}(1-\alpha)}{2}\right)^2 d\alpha + \int_0^1 \left(\frac{F^{-1}(\alpha)-F^{-1}(1-\alpha)}{2}\right)^2 d\alpha}.$$

When the second moment exists, the above simplifies to

$$\delta = \frac{1}{2}\left[1 + \frac{\int_0^1 F^{-1}(\alpha)F^{-1}(1-\alpha)d\alpha}{\mu_2}\right], \tag{3}$$

where μ_2 is the second central moment of the distribution. The sample skewness (after the sample has been scaled to have zero mean) can be computed as (see Khattree and Bahuguna, 2019),

Definition 7: Given a random sample of size n consist of observations $x_1, x_2, \ldots, x_n$, let $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$ be the order statistics of $x_1, x_2, \ldots, x_n$ after being centered by their sample mean. Define $y_i = \frac{x_{(i)}+x_{(n-i+1)}}{2}$ and $w_i = \frac{x_{(i)}-x_{(n-i+1)}}{2}$. The sample Khattree-Bahuguna univariate skewness is then defined as $\hat{\delta} = \frac{\sum y_i^2}{\sum y_i^2+\sum w_i^2}$.

Clearly being the measures of skewness, δ or $\hat{\delta}$ have no reference to any other distribution from which the distance of a given distribution can be measured. Nonetheless, δ is essentially a function of the (inverse) cumulative distribution function of the random variable and hence, intuitively speaking, if the two probability distributions are very similar, we expect that it would be reflected in their respective expressions of Khattree-Bahuguna's skewness. With this in mind, we evaluate the Khattree-Bahuguna's skewness of Lomax and exponential distributions and indicate how the former converges to the later. For this we first give, in Theorem 8, the explicit expressions for the two skewnesses. Theorem 9 then establishes the convergence.

Theorem 8: Let $X_1 \sim \text{Lomax}(\beta, \theta)$ and $X_2 \sim \text{Exp}(\lambda)$. Then the Khattree-Bahuguna's skewness of X_1 and X_2 are respectively given by,

Lomax: $\delta_{X_1} = \frac{1}{2}\left[1 + \left(B(1-\frac{1}{\beta}, 1-\frac{1}{\beta}) - \frac{\beta^2}{(\beta-1)^2}\right)\frac{(\beta-1)^2(\beta-2)}{\beta}\right], \quad \beta > 2;$

Exponential: $\delta_{X_2} = 1 - \frac{\pi^2}{12}$.

Note that neither δ_{X_1} nor δ_{X_2} depends on respective scale parameters.

Proof: Proof involves the corresponding evaluations of the expression given in (3). For exponential distribution this evaluation is straightforward. Thus, we will work out the details only for the Lomax distribution. We know that for the Lomax random variable X_1, $F(x_1) = 1 - (1+\theta x_1)^{-\beta}$ and hence $F^{-1}(\alpha) = \frac{(1-\alpha)^{-\frac{1}{\beta}}-1}{\theta}$. For $\beta > 2$, the mean and variance of X_1 are $\frac{1}{\theta(\beta-1)}$ and $\frac{\beta}{\theta^2(\beta-1)^2(\beta-2)}$, respectively. Therefore, $\int_0^1 [F^{-1}(\alpha) - \mu][F^{-1}(1-\alpha) - \mu]\, d\alpha = \int_0^1 \left[\frac{(1-\alpha)^{-\frac{1}{\beta}} - \frac{\beta}{\beta-1}}{\theta}\right]\left[\frac{\alpha^{-\frac{1}{\beta}} - \frac{\beta}{\beta-1}}{\theta}\right] d\alpha = \frac{1}{\theta^2}\left[B(1-\frac{1}{\beta}, 1-\frac{1}{\beta}) - \frac{\beta^2}{(\beta-1)^2}\right]$. Thus,

$$\delta_{X_1} = \frac{1}{2}\left[1 + \left(B(1-\frac{1}{\beta}, 1-\frac{1}{\beta}) - \frac{\beta^2}{(\beta-1)^2}\right)\frac{(\beta-1)^2(\beta-2)}{\beta}\right], \quad \beta > 2.$$

□

Theorem 9: As $\beta \to \infty$, Khattree-Bahuguna's skewness of Lomax distribution approaches that of exponential distribution.

Proof: We have,

$$\begin{aligned}
\lim_{\beta\to\infty}\delta_{X_1} &= \lim_{\beta\to\infty}\frac{1}{2}\left[1+\left(B(1-\frac{1}{\beta},1-\frac{1}{\beta})-\frac{\beta^2}{(\beta-1)^2}\right)\frac{(\beta-1)^2(\beta-2)}{\beta}\right] \\
&= \lim_{u\to 1}\frac{1}{2}\left[1+\left(B(u,u)\frac{u^2}{(1-u)^2}-\frac{1}{(1-u)^2}\right)(2u-1)\right] \qquad \left(u=1-\frac{1}{\beta}\right) \\
&= \frac{1}{2}+\frac{1}{2}\lim_{u\to 1}\frac{u^2B(u,u)-1}{(1-u)^2} \\
&= \frac{1}{2}+\frac{1}{2}\lim_{u\to 1}\frac{2uB(u,u)+u^2B'(u,u)}{2(u-1)} \\
&= \frac{1}{2}+\frac{1}{2}\lim_{u\to 1}\frac{2B(u,u)+2uB'(u,u)+2uB'(u,u)+u^2B''(u,u)}{2} \\
&= \frac{1}{2}+\frac{1}{2}\cdot\frac{2+2(-2)+2(-2)+8-\frac{\pi^2}{3}}{2} = 1-\frac{\pi^2}{12} = \delta_{X_2}.
\end{aligned}$$

The calculation of the last limit in above is rather complex and therefore is evaluated by using the Lemmas 21 and 22 which are given in ANNEXURE A. □

3. Relationships of Multivariate Lomax to Other Distributions in Exponential Family of Distributions

We now again consider the multivariate Lomax distribution of Nayak (1987). As pointed out by him, multivariate Lomax distribution is related to many other multivariate distributions such as Mardia's Pareto type I, Burr, Logistic, Cook-Johnson's uniform (alternatively called Clayton copula), and F. We will further observe here that Lomax distribution is also related to a few other univariate distributions of exponential family through the linear combinations of multivariate Lomax or via one-to-one transformation from univariate Lomax. We convey these facts via following Theorems. We skip the proofs for the sake of brevity.

Theorem 10: Let $(X_1, X_2, \cdots, X_k)$ follow a k-dimensional multivariate Lomax distribution as given by probability density function in (2). Define $X^* = \sum_{i=1}^{k}\theta_i X_i$. Then X^* is distributed as beta type II (also called inverted beta or beta prime) with shape parameters k and β, and therefore its probability density function is given by,

$$f^*(x^*) = \frac{x^{*k-1}(1+x^*)^{-(\beta+k)}}{B(k,\beta)}. \tag{4}$$

Theorem 11: Let $(X_1, X_2, \cdots, X_k)$ be k random variables jointly following multivariate Lomax distribution as defined in (2). Define $Y^* = \beta\sum_{i=1}^{k}\theta_i X_i$. Then

(i) Y^* is distributed as beta type II with shape parameters k and β and scale parameter β and therefore,

$$f(y^*) = \frac{y^{*k-1}(1+\frac{y^*}{\beta})^{-(\beta+k)}}{\beta^k B(k,\beta)}. \tag{5}$$

(ii) as $\beta \to \infty$, the distribution of Y^* approaches gamma distribution with shape parameter k and scale parameter 1.

Theorem 12: Let X be a random variable following Lomax distribution defined in (1). Define $Y = (\theta X/d)^{1/c}$. Then

(i) the probability density function of Y is a Burr density with shape parameters β and c, and rate parameter $d^{\frac{1}{c}}$, that is, $f(y) = \frac{cd\beta y^{c-1}}{(1+dy^c)^{\beta+1}}, y > 0$. See Nayak (1987);

(ii) let $W = \beta^{\frac{1}{c}} Y$, then the probability density function of W is $f(w) = \frac{cdw^{c-1}}{(1+\frac{d}{\beta}w^c)^{\beta+1}}, w > 0$;

(iii) as $\beta \to \infty$, the distribution of W approaches a Weibull distribution with shape parameter c and rate parameter d, density which is given by, $f(w) = cdw^{c-1}e^{-dw^c}, w > 0$.

It may be noted that when $c = 2$ and $d = \theta$, the distribution of W approaches to that of a Rayleigh random variable.

4. A Generalized Double Pareto-Laplace Connection

We take the previous discussion one step further by making the support of the respective random variables as the entire real line. Specifically, the role of Lomax will now be played by double Pareto distribution and that of exponential is now played by the generalized Laplace distribution. The probability density function for the two are given below.

$$\text{Double Pareto}: \quad f(x) = \begin{cases} \frac{\theta_1\theta_2\beta}{(\theta_1+\theta_2)(1-\theta_1 x)^{\beta+1}}, & \text{if } x \le 0,\ \beta, \theta_1, \theta_2 > 0 \\ \\ \frac{\theta_1\theta_2\beta}{(\theta_1+\theta_2)(1+\theta_2 x)^{\beta+1}}, & \text{if } x > 0,\ \beta, \theta_1, \theta_2 > 0. \end{cases}$$

$$\text{Generalized Laplace}: \quad g(x) = \frac{1}{\sigma}\frac{\kappa}{1+\kappa^2} \begin{cases} e^{\frac{1}{\sigma\kappa}(x-\mu)}, & \text{if } x \le 0,\ \sigma, \kappa > 0,\ \mu \in \mathbb{R} \\ \\ e^{-\frac{\kappa}{\sigma}(x-\mu)}, & \text{if } x > 0,\ \sigma, \kappa > 0,\ \mu \in \mathbb{R}. \end{cases}$$

A particular connection between the above two distributions is given by Kotz *et al.* (2001) who indicate that a double Pareto random variable can be generated by taking the ratio of two independent Laplace variates. Two especially attractive properties of double Pareto distribution are (i) its Laplace like spike of density function at zero and (ii) its Student's t-like heavy tails. See Armagan, Dunson and Lee (2013) and Pal, Khare and Hobert (2017). The double Pareto has recently received considerable attention as a choice of the prior distribution in the context of Bayesian robust shrinkage estimation (Armagan *et al.*, 2013) and thus its connection to generalized Laplace distribution is of special interest.

We will show that as in the case of Lomax-exponential connection, the properties (i) and (ii) stated above can be interpreted through the double Pareto-Laplace connection. We will also demonstrate that the double Pareto can be represented as a mixture of several t-distributions. However, to do so, we must first define a three-parameter generalized double Pareto distribution, which allows the possibility of asymmetry in the density. For this, it is convenient to pursue an approach where the bivariate Lomax distribution plays a central role. This is given by Theorem 13 that follows.

4.1. Three-parameter generalized double Pareto distribution

To set the stage, we observe that similar to the case of Laplace, a classical symmetric double Pareto distribution can be obtained by the difference of two independent Lomax variates. In order to incorporate asymmetric double Pareto distributions, we propose a three-parameter generalized double Pareto distribution defined via a bivariate Lomax distribution, where the variates naturally exhibit dependence.

Theorem 13: Assume that X_1 and X_2 are jointly distributed as the bivariate Lomax variables with parameters $\beta, \theta_1, \theta_2$ as given in (2). Then the probability density function of $X = X_2 - X_1$ is given by,

$$h_1(x) = \begin{cases} \frac{\theta_1\theta_2\beta}{(\theta_1+\theta_2)(1-\theta_1 x)^{\beta+1}}, & \text{if } x \leq 0,\ \beta, \theta_1, \theta_2 > 0 \\ \\ \frac{\theta_1\theta_2\beta}{(\theta_1+\theta_2)(1+\theta_2 x)^{\beta+1}}, & \text{if } x > 0,\ \beta, \theta_1, \theta_2 > 0. \end{cases} \tag{6}$$

Proof: Let $X = X_2 - X_1$ and $Y = \theta_1 X_1 + \theta_2 X_2$, then the Jacobian of the transformation is $|J| = \frac{1}{\theta_1+\theta_2}$. The joint probability density of X and Y is thus, $h(x, y) = \frac{\theta_1\theta_2\beta(\beta+1)}{(\theta_1+\theta_2)(1+y)^{\beta+2}}, \quad y > \max\{\theta_2 x, -\theta_1 x\}$. Integrating over y gives the marginal density function of X as given above. □

Note that for bivariate Lomax variate (X_1, X_2), $X = X_1 + X_2$ does not follow a Lomax distribution. We state the result about this sum as follows.

Theorem 14: Let (X_1, X_2) follow a bivariate Lomax distribution with parameters $\beta, \theta_1, \theta_2$ and let $X = X_1 + X_2$. Assuming $\theta_1 \neq \theta_2$, the probability density function of X is

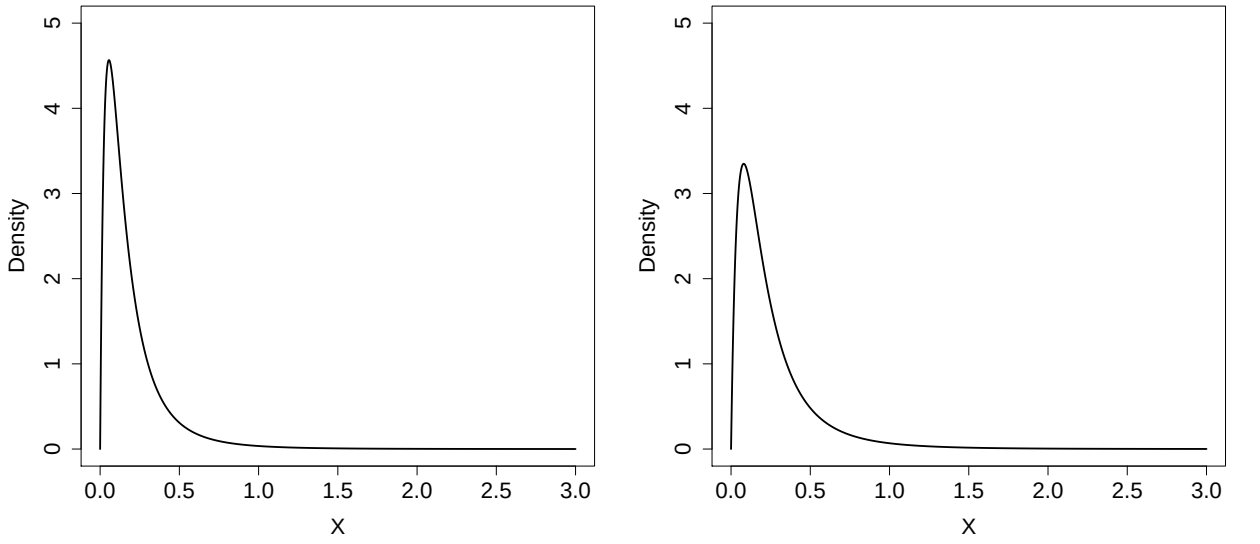
$$h_2(x) = \frac{\theta_1\theta_2\beta}{(\theta_2 - \theta_1)} \left[\frac{1}{(1+\theta_1 x)^{\beta+1}} - \frac{1}{(1+\theta_2 x)^{\beta+1}} \right], \quad x > 0.$$

When $\theta_1 = \theta_2 = \theta$, the probability density function of X is beta type II distribution given in (4) with shape parameters $k = 2$ and β, and rate parameter θ.

Proof follows by letting $Y = X_1$ and $X = X_1 + X_2$ and integrating over y. For the case $\theta_1 = \theta_2 = \theta$, the distribution of $Y = X_1 + X_2$ is a beta type II distribution with rate parameter θ, as already stated in Theorem 10. Figure 4 gives the density plots for the sum of the components of a bivariate Lomax vector with parameters $\beta = 4, \theta_1 = 2.5, \theta_2 = 5$ in (a) and parameters $\beta = 4, \theta_1 = \theta_2 = 2.5$ in (b). Clearly, both density plots exhibit a very different shape compared to Lomax distribution. The nth raw moment of $X = X_1 + X_2$ is given by $E(X^n) = \frac{\theta_1\theta_2 n!}{(\theta_2-\theta_1)(\beta-1)\cdots(\beta-n)} \left[\frac{1}{\theta_1^{n+1}} - \frac{1}{\theta_2^{n+1}} \right], n < \beta$, and $E(X^n) = \infty$ when $n \geq \beta$. Specifically, for $\beta > 2$, we have, $E(X) = \frac{\theta_1+\theta_2}{\theta_1\theta_2(\beta-1)}$ and $\mathrm{Var}(X) = \frac{\beta\theta_1^2+\beta\theta_2^2+2\theta_1\theta_2}{\theta_1^2\theta_2^2(\beta-1)^2(\beta-2)}$.

Definition 15: We define a generalized double Pareto distribution as that for a real-valued random variable X whose probability density function is given by (6). We will denote it by $\mathbf{GDP}(\beta, \theta_1, \theta_2)$.

Thus, in the case of generalized double Pareto distribution, β is the shape parameter and θ_1, θ_2 are two rate parameters. It is clear that if X is $\mathbf{GDP}(\beta, \theta_1, \theta_2)$ then $-X$ is



(a) $\beta = 4, \theta_1 = 2.5, \theta_2 = 5.0$

(b) $\beta = 4, \theta_1 = \theta_2 = 2.5$ **(Beta type II: shape parameters** $k = 2$ **and** $\beta = 4$ **and rate parameter** $\theta = 2.5$**)**

Figure 4: Examples of sum variable for bivariate Lomax distribution

$\mathbf{GDP}(\beta, \theta_2, \theta_1)$. When $\theta_1 = \theta_2 = \theta$, the generalized double Pareto distribution reduces to the classical symmetric double Pareto distribution, denoted by $\mathbf{CDP}(\beta, \theta)$ and with the density function $f(x) = \frac{\theta}{2} \frac{1}{(1+\theta|x|)^{\beta+1}}, -\infty < x < \infty, \beta, \theta > 0$. Letting $\theta = \frac{1}{\xi}$ and scaling X by shape parameter β in this density results in the density of the double Pareto distribution defined by Armagan *et al.* (2013) with probability density function, $f(x) = \frac{1}{2\xi} \left(1 + \frac{|x|}{\beta\xi}\right)^{-(\beta+1)}, -\infty < x < \infty, \beta, \xi > 0$.

Figure 5 contrasts the behavior of the density function of the generalized double Pareto distribution random variables for the symmetric ($\theta_1 = \theta_2$) and asymmetric ($\theta_1 \neq \theta_2$) cases.

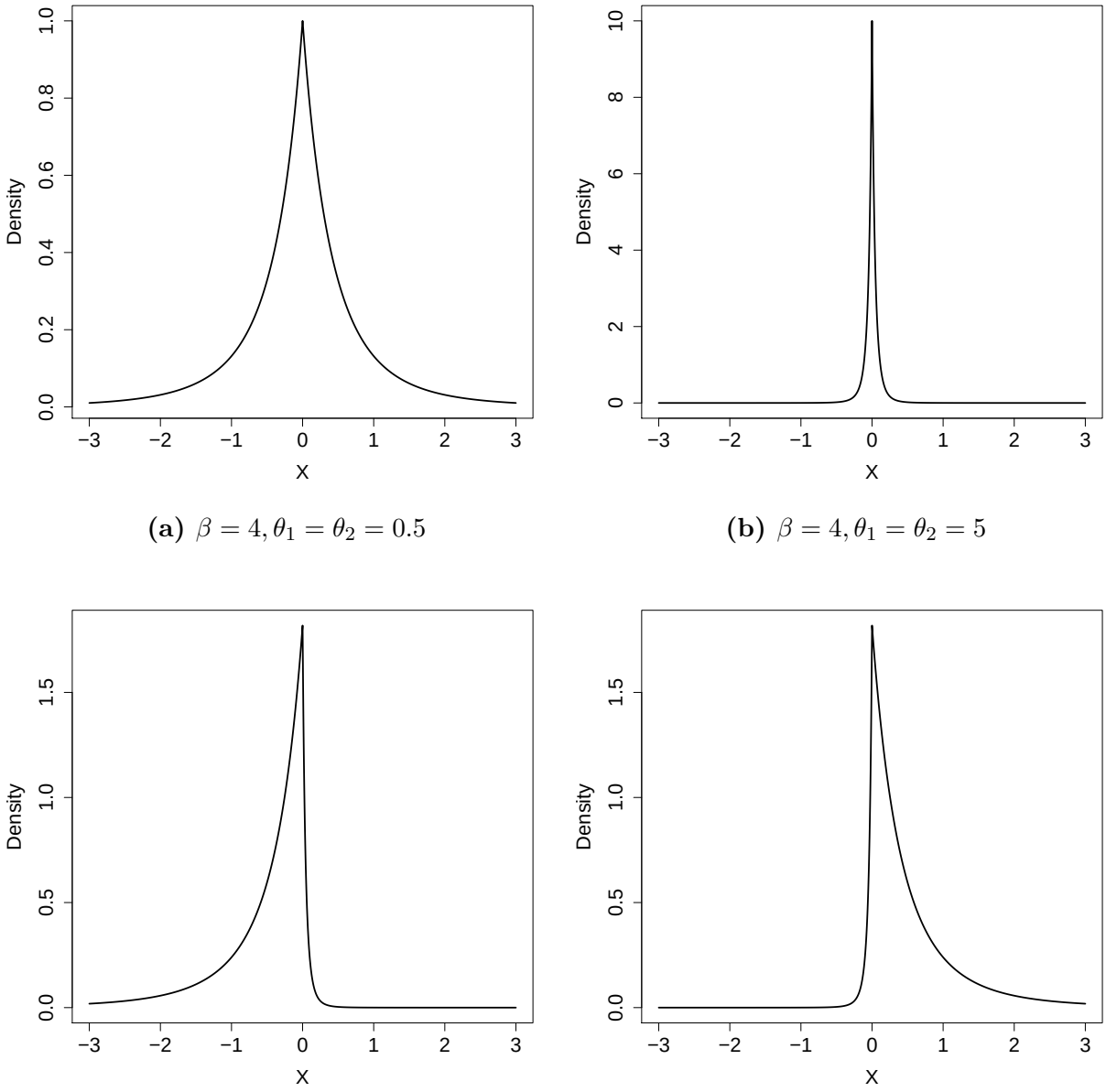
The cumulative distribution function of $\mathbf{GDP}(\beta, \theta_1, \theta_2)$ is given by

$$F(x) = \begin{cases} \frac{\theta_2}{(\theta_1+\theta_2)(1-\theta_1 x)^{\beta}}, & \text{if } x \leq 0,\ \beta, \theta_1, \theta_2 > 0 \\ \\ 1 - \frac{\theta_1}{(\theta_1+\theta_2)(1+\theta_2 x)^{\beta}}, & \text{if } x > 0,\ \beta, \theta_1, \theta_2 > 0. \end{cases}$$

It is easy to observe that $P(X \leq 0) = \frac{\theta_2}{\theta_1+\theta_2}$ and $P(X > 0) = \frac{\theta_1}{\theta_1+\theta_2}$, which can be interpreted as the weights of the two rate parameters. There is a larger proportion of negative values whenever $\theta_2 > \theta_1$. This observation is evident in Figure 5 (c).

The quantile function $F^{-1}(\alpha)$ of $\mathbf{GDP}(\beta, \theta_1, \theta_2)$ is,

$$F^{-1}(\alpha) = \begin{cases} \frac{1}{\theta_1} \left[1 - \left(\frac{\theta_2}{\alpha(\theta_1+\theta_2)}\right)^{\frac{1}{\beta}}\right], & \text{if } 0 < \alpha \leq \frac{\theta_2}{\theta_1+\theta_2}, \\ \\ \frac{1}{\theta_2} \left[\left(\frac{\theta_1}{(1-\alpha)(\theta_1+\theta_2)}\right)^{\frac{1}{\beta}} - 1\right], & \text{if } \frac{\theta_2}{\theta_1+\theta_2} < \alpha < 1. \end{cases} \tag{7}$$



(a) $\beta = 4, \theta_1 = \theta_2 = 0.5$

(b) $\beta = 4, \theta_1 = \theta_2 = 5$

(c) $\beta = 4, \theta_1 = 0.5, \theta_2 = 5$

(d) $\beta = 4, \theta_1 = 5, \theta_2 = 0.5$

Figure 5: Examples of symmetric (a)-(b) and asymmetric (c)-(d) double Pareto distributions

The nth raw moment of $\mathbf{GDP}(\beta, \theta_1, \theta_2)$ is given by $E(X^n) = \frac{n!}{(\theta_1+\theta_2)(\beta-1)\cdots(\beta-n)}\left[\frac{(-1)^n\theta_2}{\theta_1^n} + \frac{\theta_1}{\theta_2^n}\right]$, if $n < \beta$ and $E(X^n) = \infty$ when $n \geq \beta$. We defer this straightforward yet a bit tedious calculation to ANNEXURE B. When $\beta > 2$, $E(X) = \frac{\theta_1-\theta_2}{\theta_1\theta_2(\beta-1)}$ and $\text{Var}(X) = \frac{\beta\theta_1^2+\beta\theta_2^2-2\theta_1\theta_2}{\theta_1^2\theta_2^2(\beta-1)^2(\beta-2)} = \frac{\beta(\theta_1-\theta_2)^2+2\theta_1\theta_2(\beta-1)}{\theta_1^2\theta_2^2(\beta-1)^2(\beta-2)}$. Also assuming $\beta > 3$, the Pearson's coefficient of skewness is given by,

$$\gamma = \frac{2(\beta+1)\sqrt{(\beta-2)}(\theta_1-\theta_2)[\beta\theta_1^2+\beta\theta_2^2+(\beta-3)\theta_1\theta_2]}{(\beta-3)(\beta\theta_1^2+\beta\theta_2^2-2\theta_1\theta_2)^{3/2}}, \qquad \beta > 3.$$

Clearly, $\gamma = 0$ for the symmetric double Pareto as in that case, $\theta_1 = \theta_2$. By letting $\frac{\theta_1}{\theta_2} = \kappa$, we can further simplify the Pearson's γ as $\gamma = \frac{2(\beta+1)\sqrt{(\beta-2)}(\kappa-1)[\beta\kappa^2+\beta+(\beta-3)\kappa]}{(\beta-3)(\beta\kappa^2+\beta-2\kappa)^{3/2}}$. When $\kappa \to 0$, that is, when $\theta_1 \ll \theta_2$, the skewness approaches to that of negative univariate Lomax distribution. That is,

$$\lim_{\kappa\to 0}\gamma = \frac{-2(\beta+1)}{\beta-3}\sqrt{\frac{\beta-2}{\beta}}.$$

Similarly, let $\frac{\theta_2}{\theta_1} = \kappa^*$, we have $\gamma = \frac{2(\beta+1)\sqrt{(\beta-2)}(1-\kappa^*)[\beta\kappa^{*2}+\beta+(\beta-3)\kappa^*]}{(\beta-3)(\beta\kappa^{*2}+\beta-2\kappa^*)^{3/2}}$. When $\kappa^* \to 0$, that is, $\theta_2 \ll \theta_1$, the skewness approaches that of positive univariate Lomax distribution. Specifically, $\lim_{\kappa^*\to 0}\gamma = \frac{2(\beta+1)}{\beta-3}\sqrt{\frac{\beta-2}{\beta}}$.

By using (7), Khattree-Bahuguna's skewness is evaluated to be

$$\delta = \frac{1}{2}\left[1 + \frac{2\mathcal{I}_1 + \mathcal{I}_2 - \mu^2}{\mu_2}\right],$$

where

$$\mathcal{I}_1 = \frac{d_1}{\theta_1\theta_2}\left[\frac{1}{\beta-1} - \left(\frac{d_2}{d_1}\right)^{\frac{1}{\beta}}\frac{\beta}{(\beta-1)(\beta-2)}\right]$$

$$\mathcal{I}_2 = \frac{1}{\min(\theta_1,\theta_2)^2}\left\{d_2^{\frac{2}{\beta}}\left[B\left(d_2; 1-\frac{1}{\beta}, 1-\frac{1}{\beta}\right) - B\left(d_1; 1-\frac{1}{\beta}, 1-\frac{1}{\beta}\right)\right] + 2\left(\frac{d_2}{d_1}\right)^{\frac{1}{\beta}}\frac{\beta}{\beta-1}d_1 - 2\frac{\beta}{\beta-1}d_2 + d_2 - d_1\right\}$$

$$\mu = \frac{\theta_1-\theta_2}{\theta_1\theta_2(\beta-1)}, \qquad \mu_2 = \frac{\beta(d_2-d_1)^2}{d_1^2d_2^2(\beta-1)^2(\beta-2)} + \frac{2}{\theta_1\theta_2(\beta-1)(\beta-2)}$$
$$d_1 = \min\left(\frac{\theta_1}{\theta_1+\theta_2}, \frac{\theta_2}{\theta_1+\theta_2}\right), \qquad d_2 = 1-d_1,$$

and $B(x; a, b) = \int_0^x t^{a-1}(1-t)^{b-1}dt$. For the symmetric double Pareto distribution, $d_1 = d_2 = \frac{1}{2}$, and then $\mathcal{I}_1 = \frac{-1}{\theta_1\theta_2(\beta-1)(\beta-2)}, \mathcal{I}_2 = 0$. Accordingly, $\delta = 0$. Detailed and cumbersome calculations for all of these facts are deferred to ANNEXURE B.

Like Laplace distribution, the probability density function shown in Figure 5 is also spiked. Thus, it is natural to explore any double Pareto-Laplace connection by using the

similar technique as used earlier to obtain the Lomax-exponential connection. In fact, this double Pareto-Laplace connection was indirectly hinted by Armagan *et al.* (2013) where their Laplace prior (Bayesian lasso) was treated as the limiting case of double Pareto prior. The following theorem formalizes it.

Theorem 16: Let X be a $\mathbf{GDP}(\beta, \theta_1, \theta_2)$ random variable. Define $Y = \beta X$. Then

(i) the distribution of Y is $\mathbf{GDP}(\beta, \frac{\theta_1}{\beta}, \frac{\theta_2}{\beta})$;

(ii) as $\beta \to \infty$, the distribution of Y approaches a two-parameter Laplace distribution with parameters θ_1, θ_2, ($\mathbf{Laplace}(\theta_1, \theta_2)$) with probability density function

$$g^*(y) = \begin{cases} \frac{\theta_1\theta_2}{\theta_1+\theta_2} e^{\theta_1 y}, & \text{if } y \leq 0, \\ \\ \frac{\theta_1\theta_2}{\theta_1+\theta_2} e^{-\theta_2 y}, & \text{if } y > 0. \end{cases} \tag{8}$$

Therefore, we surmise that $\mathbf{GDP}(\beta, \frac{\theta_1}{\beta}, \frac{\theta_2}{\beta})$ can be approximated by $\mathbf{Laplace}(\theta_1, \theta_2)$ if the shape parameter β is large. Moreover, both $\mathbf{GDP}(\beta, \frac{\theta_1}{\beta}, \frac{\theta_2}{\beta})$ and $\mathbf{Laplace}(\theta_1, \theta_2)$ have the probability density spiked at zero. As pictorially demonstrated in Figure 6, as β increases, the density plot of $\mathbf{GDP}(\beta, \theta_1, \theta_2)$ approaches that of Laplace distribution. Also with reparameterization $\theta_1 = \frac{1}{\sigma\kappa}$ and $\theta_2 = \frac{\kappa}{\sigma}$ the Laplace density in (8) reduces to the form introduced by Hinkley and Revankar (1977) with zero location parameter ($\mu = 0$) as

$$g^*(y) = \frac{1}{\sigma} \frac{\kappa}{1+\kappa^2} \begin{cases} e^{\frac{1}{\sigma\kappa} y}, & \text{if } y \leq 0,\ \sigma, \kappa > 0 \\ \\ e^{-\frac{\kappa}{\sigma} y}, & \text{if } y > 0,\ \sigma, \kappa > 0. \end{cases}$$

4.2. A representation of double Pareto distribution

The next result shows that any symmetric double Pareto random variable can be thought of as a Student's t random variable when scaled by an independent Lomax random variate. This results in a symmetric yet a Student's t-like heavy tails of double Pareto distribution. Due to this heavy-tail property, it has been widely used in Bayesian shrinkage as a choice of prior. More formally,

Theorem 17: A symmetric double Pareto random variate X with shape parameter $(\nu - 1)$ and scale parameter $\sqrt{\nu}$ can be represented as

$$X \stackrel{d}{=} \sqrt{Y} T, \tag{9}$$

where the random variable Y has a standard Lomax distribution with shape parameter $\beta > 0$ and T has an independent Student's t distribution with degrees of freedom $\nu = 2\beta + 1$. The notation $\stackrel{d}{=}$ indicates the equivalence of distributions.

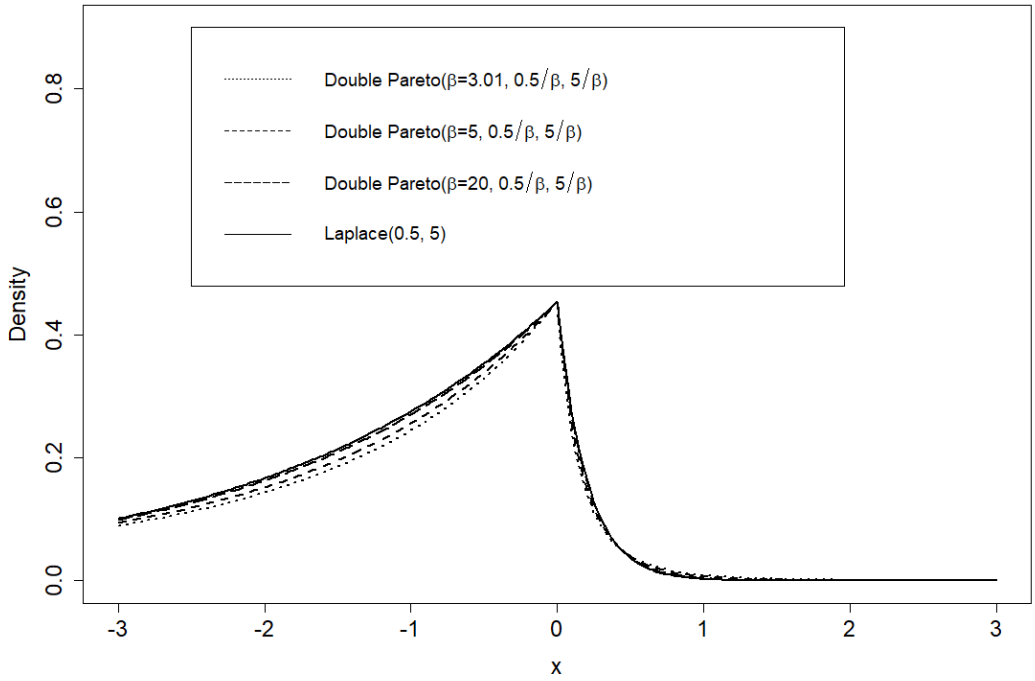


Figure 6: Density plots of GDP$(\beta, \frac{0.5}{\beta}, \frac{5}{\beta})$ **with changing parameter** $\beta = 3.01, 5, 20$ **and Laplace**$(\theta_1 = 0.5, \theta_2 = 5)$

Proof:. Given $Y \sim$ Lomax$(\beta, 1)$ and $T \sim$ Student's t with degrees of freedom ν, the probability density function of X, is

$$\begin{aligned} f_X(x) &= \int_0^\infty f_T(\frac{x}{\sqrt{y}})\frac{1}{\sqrt{y}} f_Y(y)dy = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})}\int_0^\infty \frac{1}{(1+\frac{x^2}{\nu y})^{\frac{\nu+1}{2}}}\frac{1}{\sqrt{y}}\frac{\beta}{(1+y)^{\beta+1}}dy \\ &= \frac{2\beta\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})}\int_0^\infty \frac{1}{(1+\frac{x^2}{\nu u^2})^{\frac{\nu+1}{2}}(1+u^2)^{\beta+1}}du \qquad (u=\sqrt{y}). \\ &= \frac{2\beta}{\sqrt{\nu}B(\frac{\nu}{2},\frac{1}{2})}\int_0^\infty \frac{1}{(1+\frac{x^2}{\nu u^2})^{\frac{\nu+1}{2}}(1+u^2)^{\beta+1}}du. \end{aligned}$$

The above integral is difficult to evaluate for general values of ν and β. However, when $\nu = 2\beta + 1$ and hence $\frac{\nu+1}{2} = \beta + 1$, (as stated in Theorem) simplification occurs. In this case, by using the Lemma 23 which is stated and proved in ANNEXURE, we have,

$$f_X(x) = \frac{(\nu-1)}{\sqrt{\nu}B(\frac{\nu}{2},\frac{1}{2})}\frac{B(\frac{\nu+1}{2}-\frac{1}{2},\frac{1}{2})}{2(1+\frac{|x|}{\sqrt{\nu}})^{2\cdot\frac{\nu+1}{2}-1}} = \frac{1}{2\sqrt{\nu}}\frac{(\nu-1)}{(1+\frac{|x|}{\sqrt{\nu}})^{\nu}}, \quad -\infty < x < \infty,$$

which is a symmetric double Pareto distribution with shape parameter $(\nu - 1)$ and scale parameter $\sqrt{\nu}$. □

Theorem 18: A symmetric double Pareto random variable X with shape parameter $(\nu-1)$ and scale parameter ν can be represented as $X \stackrel{d}{=} \sqrt{\nu Y}T$, where the random variable Y has

a standard Lomax distribution with shape parameter $\beta > 0$ and T has an independent Student's t distribution with degrees of freedom $\nu = 2\beta + 1$. Moreover, as $\nu \to \infty$, X approaches the standard Laplace distribution.

Proof: Through the linear transformation used on Theorem 17, we easily obtain the probability density function of X as $f_X(x) = \frac{1}{2\nu}\frac{\nu-1}{(1+\frac{|x|}{\nu})^{\nu}}, -\infty < x < \infty$. Clearly $\lim_{\nu\to\infty} f_X(x) = \frac{1}{2}e^{-|x|}, -\infty < x < \infty$. □

The above two results actually describe another remarkable feature of double Pareto-Laplace connection where Laplace distribution also has the similar representation but of mixture of normal distributions instead of Student's t. For completeness and for the sake of comparison, we restate the representation of Laplace distribution mentioned in Kotz *et al.* (2001), in the following result.

Theorem 19: A standard classical (symmetric) Laplace random variable X has the representation $X \stackrel{d}{=} \sqrt{2W}Z$, where the random variables W and Z are independent and have the standard exponential and normal distributions, respectively.

The above result establishes the Laplace distribution as a mixture of normal distribution with a scale parameter having exponential distribution. Actually, the above proposition can be viewed as the limiting case of both sides of (9) as $\beta, \nu \to \infty$ where the double Pareto, Lomax and Student's t distributions respectively approach Laplace, exponential and normal. This is perhaps the reason as to why the double Pareto appears to be a better choice for prior distribution than the Laplace distribution in Bayesian shrinkage estimation when we require a prior with heavy tails.

5. Random Number Generation from Double Pareto Distribution

The R (R Core Team, 2019) package **NonNorMvtDist** (Lun and Khattree, 2020) is a recent versatile package which implements the simulation and probability computations for a large number of non-normal multivariate distributions including the Lomax. See Figure 10. By Theorem 13 and by using the aforementioned package, random numbers from the $\mathbf{GDP}(\beta, \theta_1, \theta_2)$ can be easily generated in barely two steps as follows.

1. Generate a sample of size n bivariate Lomax random vector (X_1, X_2) with shape parameter β and the vector of rate parameters (θ_1, θ_2) using the function `rmvlomax()`.
2. Return $X = X_2 - X_1$.

As an example, we generate $\mathbf{GDP}(3.5, 1.5, 5)$ of size 5000, using the following R code.

```
library(NonNorMvtDist)
beta = 3.5; theta1 = 1.5; theta2 = 5
set.seed(2020)
bivLomax = rmvlomax(n = 5000, parm1 = beta, parm2 = c(theta1, theta2))
x = bivLomax[,2] - bivLomax[,1]
hist(x, breaks=30, freq = FALSE)
```

With suitable similar extensions of other distributions as shown in Figure 10, one can implement random number generations in other cases as well.

6. An Illustrative Biomedical Application of Bivariate Lomax distribution

We consider a data set from a breast cancer study from University of California Irvine Machine Learning Repository to highlight the usefulness of Lomax distribution for modeling the non-negative skewed data. The data are attributed to Patrício *et al.* (2018) and consist of nine quantitative clinical features (age, BMI, glucose, insulin, HOMA, leptin, adiponectin, resistin, and MCP-1), and a binary classification variable (Patients *vs.* Healthy controls) observed for 64 patients with breast cancer and 52 healthy control subjects recruited from the University Hospital Centre of Coimbra. With substantial skewness present in all clinical features and hence the lack of normality assumptions, Crisóstomo *et al.* (2016) analyzed the data by applying the nonparametric methods (specifically the Kruskal-Wallis test). Among other things, insulin was identified as a significant discriminator between the two groups but only for corresponding subsets with BMI $> 25\text{kg/m}^2$. However, for the group with BMI $\leq 25\text{kg/m}^2$, the significance of insulin seemed inconclusive. Nonparametric approach with relatively low power may be one of the reason for not so clear a conclusion.

We will choose insulin as the variable of interest for our work. Instead of choosing a nonparametric approach which usually has low power especially when the sample size is not very large, and forgoing the normality based methods due to absence of normality, we here propose a Lomax model for this data. The high skewness in the data, as shown in Figures 7 (a) and (b) for healthy group and the breast cancer group respectively and the shape of the distributions justify our use of this model. Using BMI as a matching variable to match pairs of one healthy subject and one breast cancer subject, from each of the two groups, we obtain the bivariate data on insulin measurements for $n = 52$ such pairs. These values after discarding unmatched subjects are presented in Table 4 in ANNEXURE. The corresponding R code for this application can be obtained from the authors.

Sample descriptive statistics summary for the respective marginal distributions are given in the columns 2-4 of Table 2. With skewed marginal distributions as shown in Figure 7 (a) and (b), we fit the bivariate Lomax distribution on this data, using the maximum likelihood (ML) approach. This results in, $\hat{\beta} = 283.8444, \hat{\theta}_{health} = 0.000508, \hat{\theta}_{cancer} = 0.000299$. The values of descriptive statistics based on these estimates are given as columns 5-7 of Table 2. The agreement between the sample descriptive statistics and ML based descriptive statistics is quite good, even though standard deviation for the latter is somewhat higher. Also the ML estimates of correlation between the paired variables, is equal to $1/\hat{\beta} = 1/283.8444 = 0.0035$ (the sample correlation $= 0.0465$), which is low (as we must expect since the patients as well as two samples are independent). The bivariate Lomax seems to fit the data very well. This is further justified by Lomax Q-Q plots given in Figure 8.

Clearly as the ML estimates of mean and variance in Table 2, show, breast cancer group does indeed have not only much higher mean value of insulin, its values also vary much more greatly within the group, compared to those for healthy subject group. Further, we note that large variability is persistent even in the group with BMI $< 25\text{kg/m}^2$ – a fact obscured and hence lost in the nonparametric analysis done by the original authors.

Table 2: Descriptive and theoretical statistics for healthy controls and breast cancer patients, respectively.

	Descriptive Statistics ($n = 52$)					
	Sample			ML-Based		
Group	Mean	SD	Skewness	Mean	SD	Skewness
Healthy controls	6.9340	4.86	2.2765	6.9540	6.98	2.0213
Breast cancer patients	11.8420	10.20	1.4753	11.8194	11.86	2.0213

It may also be mentioned that in view of large $\hat{\beta} = 283.8444$, the exponential model, for insulin levels scaled by $\hat{\beta}$, may be applicable for both groups – healthy as well as the breast cancer group. Thus, the estimates of the two rate parameter are

$$\hat{\lambda}_{health} = \hat{\beta}\hat{\theta}_{health} = 0.1442 \quad (1/\bar{X}_{health} = 0.1441 \text{ if fitting the exponential distribution})$$
$$\hat{\lambda}_{cancer} = \hat{\beta}\hat{\theta}_{cancer} = 0.0849 \quad (1/\bar{X}_{cancer} = 0.0845 \text{ if fitting the exponential distribution})$$

Clearly the estimated value of λ_{cancer} is smaller than that for λ_{health} then again reconfirming the higher levels of insulin for the cancer group.

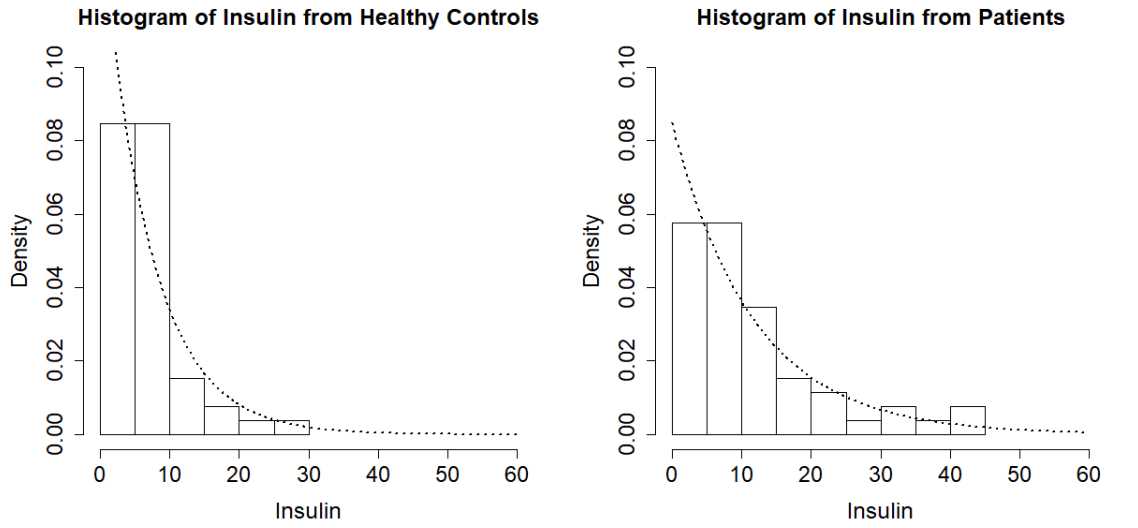
Returning to Lomax context, we may be interested in formally testing the null hypothesis $H_0 : \theta_{health} = \theta_{cancer}$ vs. $H_a : \theta_{health} > \theta_{cancer}$ which aims to test if two mean insulin levels are same for the two groups against the alternative that it is higher for healthy control groups. To do so, we consider the variable representing the difference $Y = X_{cancer} - X_{health}$. With bivariate Lomax assumption on (X_1, X_2) in place, under the null hypothesis Y must follow the symmetric double Pareto distribution.

We take Khattree-Bahuguna's skewness $\hat{\delta}$ as the test statistic. Clearly, under H_0 and hence under symmetry, $\delta = 0$. Thus we must reject H_0 for large values of $\hat{\delta}$, where $\hat{\delta}$ is the estimate of δ obtained by using the sample Khattree-Bahuguna's univariate skewness. For our data, $\hat{\delta} = 0.1164$, which is considerably larger than the one-sided cutoff value $\delta_{0.95} = 0.0742$ calculated under the null hypothesis via large number of simulations ($nsim = 1000$) and by using the R packages of Lun and Khattree (2020). The null hypothesis is thus rejected.

We may also be interested in those pairs with BMI $< 25\text{kg/m}^2$ (this is the data, which original authors had discarded as they analyzed only subjects with BMI $\geq 25\text{kg/m}^2$). Thus, we may try to fit the bivariate Lomax distribution only on $n = 17$ pairs of healthy control

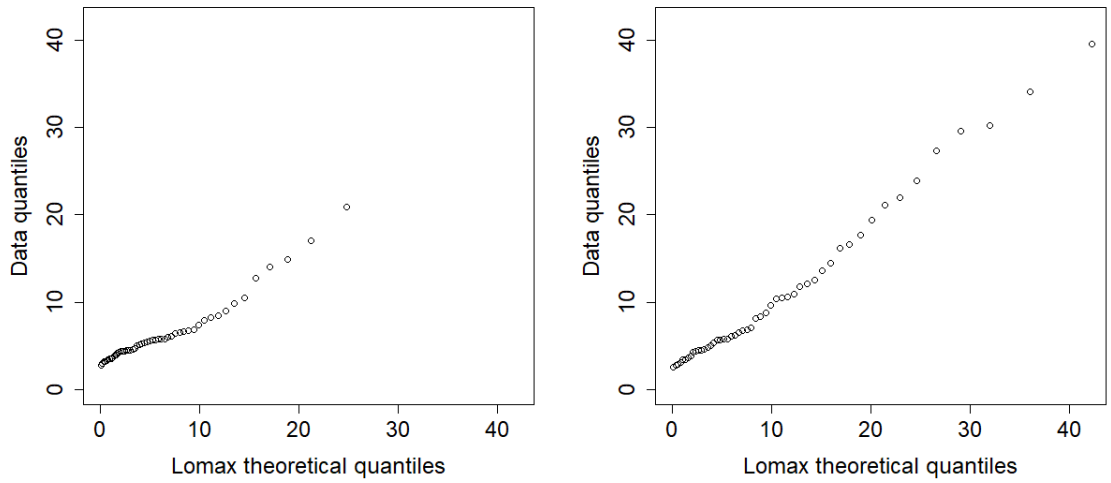
Table 3: Descriptive and theoretical statistics for healthy controls and breast cancer patients with BMI $< 25\mathbf{kg/m^2}$, respectively.

	Descriptive Statistics ($n = 17$)					
	Sample			ML-Based		
Group	Mean	SD	Skewness	Mean	SD	Skewness
Healthy controls	4.4304	1.44	0.5955	4.4100	4.4194	2.0128
Breast cancer patients	8.2047	8.76	2.2348	8.2512	8.2689	2.0128



(a) Healthy Controls (skewness = 2.2765) **(b) Breast Cancer Patients (skewness = 1.4753)**

Figure 7: Empirical distributions of healthy controls and breast cancer patients and density plots (dashed lines) obtained by maximum likelihood estimation based on bivariate Lomax model

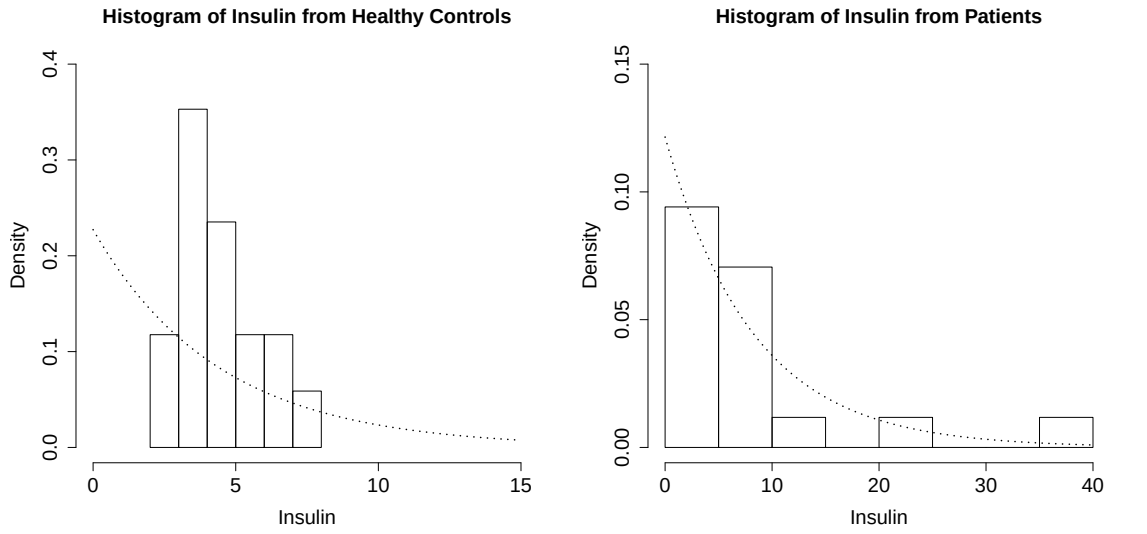


(a) Healthy controls **(b) Breast cancer patients**

Figure 8: Quantile-Quantile plot of insulin levels for healthy controls and for breast cancer patients.

and breast cancer patient, using the maximum likelihood (ML) approach. This results in, $\hat{\beta} = 471.0522, \hat{\theta}_{health} = 0.000482, \hat{\theta}_{cancer} = 0.000258$.

However, as the histograms and summary statistics show, the bivariate Lomax distribution may not fit this subset of data as satisfactorily as the whole data since the group of healthy controls has relatively low skewness while the breast cancer group is highly skewed. See the sample descriptive statistics in Table 3 and histogram in Figure 9. Therefore, as-



(a) **Healthy Controls (skewness** $= 0.5955$**)** **(b)** **Breast Cancer Patients (skewness** $= 2.2348$**)**

Figure 9: Empirical distributions of healthy controls and breast cancer patients with BMI $< 25\text{kg/m}^2$ **and density plots (dashed lines) obtained by maximum likelihood estimation based on bivariate Lomax model**

sumptions for the corresponding hypothesis testing are not met and hence hypothesis testing is not performed for this subset of data. It is difficult to determine if the poor fit to Lomax distribution is due to small number ($n = 17$) observations.

7. Concluding Remarks

As the title suggests, this article revolves around connections between Lomax and exponential distributions and between the extensions thereof. Various relationships between multivariate Lomax and several other univariate and multivariate distributions are known to exist and these relationships are graphically reproduced in Figure 10. Via these interrelationships one can possibly establish many more similar connections. For example, generalized double Pareto distribution can also be conveniently obtained via appropriate transformations of many of these bivariate distributions. The same can be said about the representation of generalized double Pareto by a mixture of Student's t distributions. Compared with scale mixture of normal distributions, Generalized double Pareto distribution provides a possibly more robust and more flexible choice of prior in practice, such as robust Bayesian shrinkage estimation and biomedical data modeling.

Data analysis presented here exemplifies the potential applications which distributions presented in this work may have. Comparisons such as that presented in our illustration, require the distributions of random variables which are linear functions of such variables and may not result in nice symmetric distributions with support on the entire real line. Generalized double Pareto distribution with asymmetry, skewness and fat tail is one such distribution which may be a flexible enough choice to accommodate such situations.

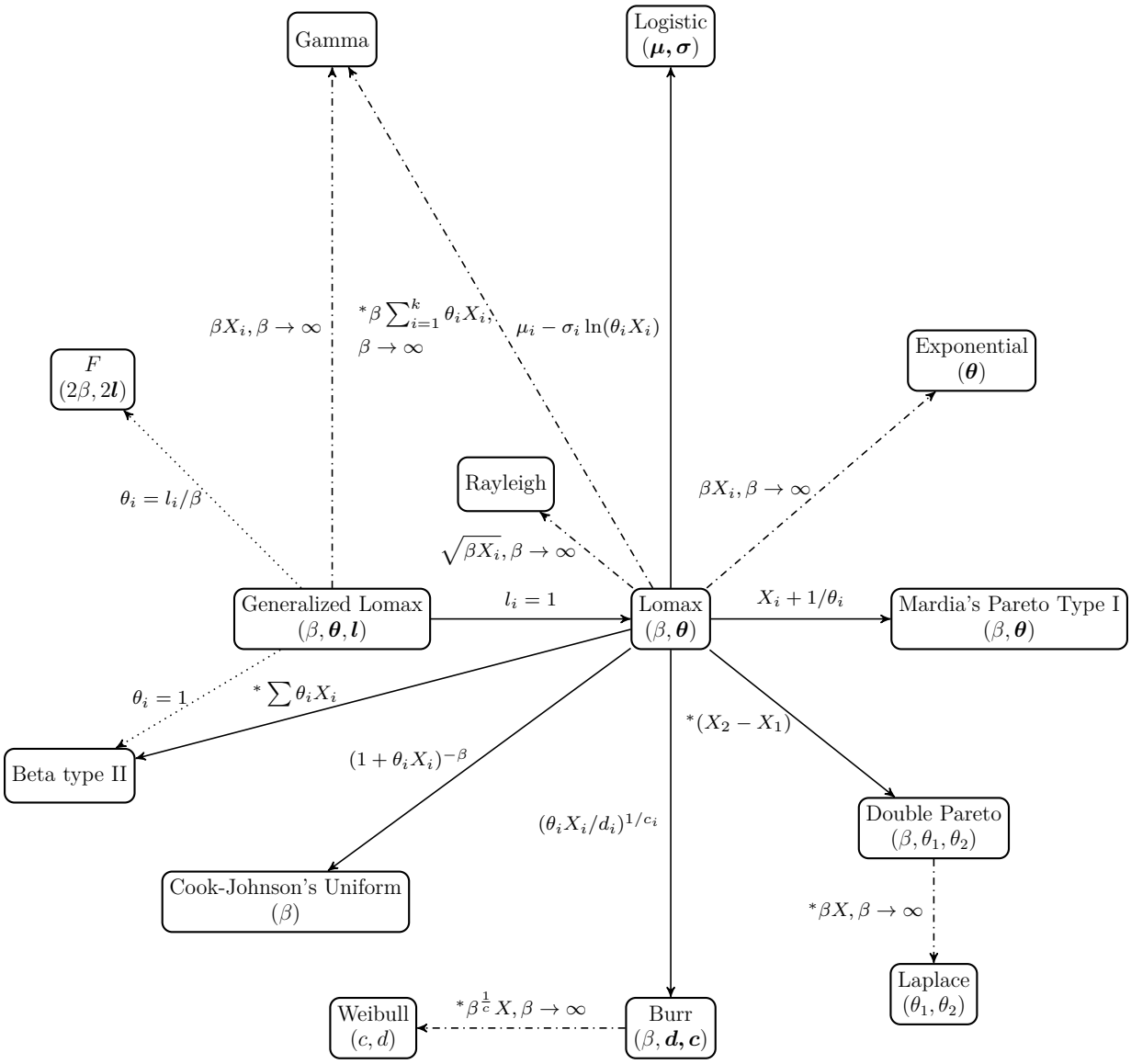


Figure 10: Lomax and its relationships through transformation (solid lines), parameter substitution (dotted lines) and limiting distribution (dash-dotted lines). Univariate distributions are marked with *.

Acknowledgments

Content of this article constitutes a part of the PhD dissertation of Zhixin Lun at Oakland University. Financial research support from Oakland University during the summers of Years 2018 and 2019 is acknowledged. We are also grateful to an anonymous referee and editor for their valuable comments resulting in several improvements in this article.

References

Armagan. A, Dunson, D. and Lee J. (2013). Generalized double Pareto shrinkage. *Statistica Sinica*, **23**, 119–143.

Crisóstomo, J., Matafome, P., Santos-Silva, D., Gomes, A. L., Gomes, M., Patrício, M., Letra, L., Sarmento-Ribeiro, A. B., Santos, L. and Seiça, R. (2016). Hyperresistinemia and metabolic dysregulation: a risky crosstalk in obese breast cancer. *Endocrine*, **53**, 433–442.

Eberl, A. and Klar, B. (2019). A note on a measure of asymmetry. *Statistical Papers*. https://doi.org/10.1007/s00362-019-01145-4

Harris, C. M. (1968). The Pareto distribution as a queue service discipline. *Operations Research*, **16**, 307–313.

Hinkley, D. V. and Revankar, N. S. (1977). Estimation of the Pareto law from underreported data: A further analysis. *Journal of Econometrics*, **5**, 1–11.

Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994). *Continuous Univariate Distributions, Volume 1, 2nd Edition*. John Wiley & Sons, New York.

Khattree, R. and Bahuguna, M. (2019). An alternative data analytic approach to measure the univariate and multivariate skewness. *International Journal of Data Science and Analytics*, **7**, 1–16.

Kotz, S., Kozubowski, T. J., and Podgórski, K. (2001). *The Laplace Distribution and Generalizations*. Birkhäuser, Boston.

Lun, Z. and Khattree, R. (2020). *KbMvtSkew: Khattree-Bahuguna's Univariate and Multivariate Skewness*. R package version 1.0.2. https://CRAN.R-project.org/package=KbMvtSkew.

Lun, Z. and Khattree, R. (2020). *NonNorMvtDist: Multivariate Lomax (Pareto Type II) and Its Related Distributions*. R package version 1.0.2. https://CRAN.R-project.org/package=NonNorMvtDist.

Lun, Z. and Khattree, R. (2020). An overlooked Lomax-exponential connection. *Communications in Statistics - Theory and Methods*. https://doi.org/10.1080/03610926.2020.1734836

Nayak, T. K. (1987). Multivariate Lomax distribution: properties and usefulness in reliability theory. *Journal of Applied Probability*, **24(1)**, 170–177.

Pal, S., Khare, K., and Hobert, J. P. (2017). Trace class markov chains for Bayesian inference with generalized Double Pareto shrinkage priors. *Scandinavian Journal of Statistics*, **44**, 307–323.

Patil, P. N., Patil, P. P. and Bagkavos, D. (2012). A measure of asymmetry. *Statistical Papers*, **53**, 971–985.

Patrício, M., Pereira, J., Crisóstomo J., Matafome, P., Gomes, M., Seiça, R. and Caramelo, F. (2018). Using Resistin, glucose, age and BMI to predict the presence of breast cancer. *BMC Cancer*, **18(29)**, 1–8.

R Core Team. (2019). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

ANNEXURE

A. Proofs of Some of the Lemmas Used

Lemma 20: For $0 < \alpha < 1$, $\int_0^1 \ln(1-\alpha)\ln(\alpha)d\alpha = 2 - \frac{\pi^2}{6}$.

Proof: Clearly, the integral is convergent. The orders of integration and summation can be interchanged. Using Taylor series to expand $\ln(1-\alpha) = -\sum_{n=1}^{\infty} \frac{\alpha^n}{n}$, we have

$$\begin{aligned}\int_0^1 \ln(1-\alpha)\ln(\alpha)d\alpha &= -\int_0^1 \ln(\alpha)\sum_{n=1}^{\infty}\frac{\alpha^n}{n}d\alpha \\ &= -\sum_{n=1}^{\infty}\frac{1}{n}\left[-\frac{1}{(n+1)^2}\right] \text{(by using integration by parts)} \\ &= \sum_{n=1}^{\infty}\left[\frac{1}{n} - \frac{1}{n+1} - \frac{1}{(n+1)^2}\right] \\ &= \sum_{n=2}^{\infty}\left[\frac{1}{n-1} - \frac{1}{n}\right] - \sum_{n=2}^{\infty}\left[\frac{1}{n^2}\right] = 1 - (\frac{\pi^2}{6} - 1) = 2 - \frac{\pi^2}{6}. \quad \square \qquad (10)\end{aligned}$$

□

Lemma 21: For the first derivative of beta function $B(u,u)$, with respect to u, say $B'(u,u)$, $\lim_{u\to 1} B'(u,u) = -2$.

Proof. Given $B(u,u) = \int_0^1 x^{u-1}(1-x)^{u-1}dx$, we have by using the Leibniz's Rule, the derivative of $B(u,u)$ with respect to u,

$$B'(u,u) = \int_0^1 \left[x^{u-1}\ln(x)(1-x)^{u-1} + x^{u-1}(1-x)^{u-1}\ln(1-x)\right]dx.$$

Accordingly, $\lim_{u\to 1} B'(u,u) = \int_0^1 \ln(x)dx + \int_0^1 \ln(1-x)dx = -1 + -1 = -2$. □

Lemma 22: For the second derivative of beta function $B(u,u)$, say $B''(u,u)$, $\lim_{u\to 1} B''(u,u) = 8 - \frac{\pi^2}{3}$.

Proof: Again, by applying Leibniz's Rule, we have

$$\begin{aligned}B''(u,u) = \int_0^1 [&x^{u-1}\ln(x)^2(1-x)^{u-1} + x^{u-1}\ln(x)(1-x)^{u-1}\ln(1-x) \\ &+ x^{u-1}\ln(x)(1-x)^{u-1}\ln(1-x) + x^{u-1}(1-x)^{u-1}\ln(1-x)^2]dx\end{aligned}$$

Thus,

$$\lim_{u\to 1} B''(u,u) = \int_0^1 \ln(x)^2 dx + 2\int_0^1 \ln(1-x)\ln(x)dx + \int_0^1 \ln(1-x)^2 dx$$
$$= 2 + 2(2 - \frac{\pi^2}{6}) + 2 = 8 - \frac{\pi^2}{3}. \quad \square$$

The middle integral is evaluated by using the Lemma 20 and integration by parts. $\square$

Lemma 23:

$$\int_0^\infty \frac{1}{(1+\frac{x^2}{y^2})^p(1+y^2)^p}dy = \frac{B(p-\frac{1}{2},\frac{1}{2})}{2(1+|x|)^{2p-1}}, \quad -\infty < x < \infty, \quad p > \frac{1}{2}.$$

Proof:. Let $\mathcal{I} = \int_0^\infty \frac{1}{(1+\frac{x^2}{y^2})^p(1+y^2)^p}dy$. Define $t = \frac{|x|}{y}$, then $dy = -\frac{|x|}{t^2}dt$. So $\mathcal{I} = \int_0^\infty \frac{1}{(1+\frac{x^2}{y^2})^p(1+y^2)^p}dy = \int_0^\infty \frac{\frac{|x|}{t^2}}{(1+t^2)^p(1+\frac{x^2}{t^2})^p}dt$. Thus,

$$2\mathcal{I} = \int_0^\infty \frac{1}{(1+\frac{x^2}{t^2})^p(1+t^2)^p}dt + \int_0^\infty \frac{\frac{|x|}{t^2}}{(1+t^2)^p(1+\frac{x^2}{t^2})^p}dt$$
$$= \int_0^\infty \left(1+\frac{|x|}{t^2}\right)\frac{1}{(1+\frac{x^2}{t^2})^p(1+t^2)^p}dt = \int_0^\infty \left(1+\frac{|x|}{t^2}\right)\frac{1}{(1+\frac{x^2}{t^2}+x^2+t^2)^p}dt.$$

Now define, $s = t - \frac{|x|}{t}$, then $ds = 1 + \frac{|x|}{t^2}dt$, $\quad s^2 = t^2 - 2|x| + \frac{x^2}{t^2}$, and hence we have

$$2\mathcal{I} = \int_{-\infty}^\infty \frac{1}{(1+s^2+2|x|+x^2)^p}ds = \int_{-\infty}^\infty \frac{1}{[s^2+(1+|x|)^2]^p}ds.$$

Thus, $\mathcal{I} = \int_0^\infty \frac{1}{[s^2+(1+|x|)^2]^p}ds = \frac{B(p-\frac{1}{2},\frac{1}{2})}{2(1+|x|)^{2p-1}}$. $\square$

B. Moments and Skewness of Generalized Double Pareto Distribution

B.1. Moments

Consider the nth raw moment for univariate Lomax distribution with shape parameter β and rate parameter θ:

$$\mu_n' = \int_0^\infty x^n f(x)dx = \int_0^\infty x^n \frac{\beta\theta}{(1+\theta x)^{\beta+1}}dx$$
$$= \frac{\beta}{\theta^n}\int_0^1 u^{(\beta-n)-1}(1-u)^{(n+1)-1}du \quad \left(\text{Let } u = \frac{1}{1+\theta x}, \text{ then } dx = -\frac{1}{\theta u^2}du\right)$$
$$= \frac{\beta}{\theta^n}B(\beta-n, n+1) \quad = \frac{n!}{\theta^n(\beta-1)\cdots(\beta-n)}, \qquad n < \beta.$$

Consider the nth raw moment for generalized double Pareto distribution with probability density function given by

$$h(x) = \begin{cases} h_1(x) = \frac{\theta_1\theta_2\beta}{(\theta_1+\theta_2)(1-\theta_1 x)^{\beta+1}}, & \text{if } x \leq 0,\ \beta, \theta_1, \theta_2 > 0 \\ \\ h_2(x) = \frac{\theta_1\theta_2\beta}{(\theta_1+\theta_2)(1+\theta_2 x)^{\beta+1}}, & \text{if } x > 0,\ \beta, \theta_1, \theta_2 > 0. \end{cases}$$

Then we have

$$\begin{aligned} \mu'_n &= \int_{-\infty}^{0} x^n h_1(x)dx + \int_{0}^{\infty} x^n h_2(x)dx \\ &= \frac{\theta_2}{\theta_1+\theta_2}(-1)^n \frac{n!}{\theta_1^n(\beta-1)\cdots(\beta-n)} + \frac{\theta_1}{\theta_1+\theta_2}\frac{n!}{\theta_2^n(\beta-1)\cdots(\beta-n)} \\ &= \frac{n!}{(\theta_1+\theta_2)(\beta-1)\cdots(\beta-n)}\left[\frac{(-1)^n\theta_2}{\theta_1^n} + \frac{\theta_1}{\theta_2^n}\right], \qquad n < \beta. \end{aligned}$$

B.2. Pearson's Coefficient of Skewness

The third central moment is given by $\mu_3 = \mu'_3 - 3\mu_2\mu'_1 - \mu'^3_1$ while the third raw moment is given by

$$\mu'_3 = \frac{3!}{(\theta_1+\theta_2)(\beta-1)(\beta-2)(\beta-3)}\left[\frac{(-1)^3\theta_2}{\theta_1^3} + \frac{\theta_1}{\theta_2^3}\right] = \frac{6(\theta_1^2+\theta_2^2)(\theta_1-\theta_2)}{\theta_1^3\theta_2^3(\beta-1)(\beta-2)(\beta-3)}.$$

Also, $\mu_2 = \text{Var}(X) = \frac{\beta\theta_1^2+\beta\theta_2^2-2\theta_1\theta_2}{\theta_1^2\theta_2^2(\beta-1)^2(\beta-2)}$ and $\mu'_1 = \frac{\theta_1-\theta_2}{\theta_1\theta_2(\beta-1)}$. Upon substitution, the third central moment is given by

$$\begin{aligned} \mu_3 &= \frac{6(\theta_1^2+\theta_2^2)(\theta_1-\theta_2)}{\theta_1^3\theta_2^3(\beta-1)(\beta-2)(\beta-3)} - \frac{3(\beta\theta_1^2+\beta\theta_2^2-2\theta_1\theta_2)}{\theta_1^2\theta_2^2(\beta-1)^2(\beta-2)}\frac{(\theta_1-\theta_2)}{\theta_1\theta_2(\beta-1)} - \frac{(\theta_1-\theta_2)^3}{\theta_1^3\theta_2^3(\beta-1)^3} \\ &= \frac{2(\theta_1-\theta_2)(\beta+1)[\beta\theta_1^2+\beta\theta_2^2+(\beta-3)\theta_1\theta_2]}{\theta_1^3\theta_2^3(\beta-1)^3(\beta-2)(\beta-3)}. \end{aligned}$$

Finally, the Pearson's skewness becomes

$$\begin{aligned} \gamma = \frac{\mu_3}{\mu_2^{3/2}} &= \frac{2(\theta_1-\theta_2)(\beta+1)[\beta\theta_1^2+\beta\theta_2^2+(\beta-3)\theta_1\theta_2]}{\theta_1^3\theta_2^3(\beta-1)^3(\beta-2)(\beta-3)} \times \left(\frac{\theta_1^2\theta_2^2(\beta-1)^2(\beta-2)}{\beta\theta_1^2+\beta\theta_2^2-2\theta_1\theta_2}\right)^{3/2} \\ &= \frac{2(\beta+1)\sqrt{(\beta-2)}(\theta_1-\theta_2)[\beta\theta_1^2+\beta\theta_2^2+(\beta-3)\theta_1\theta_2]}{(\beta-3)(\beta\theta_1^2+\beta\theta_2^2-2\theta_1\theta_2)^{3/2}}, \qquad \beta > 3. \end{aligned}$$

B.3. Khattree-Bahuguna's Skewness

Recall the quantile function for GDP$(\beta, \theta_1, \theta_2)$

$$F^{-1}(\alpha) = \begin{cases} \frac{1}{\theta_1}\left[1 - \left(\frac{\theta_2}{\alpha(\theta_1+\theta_2)}\right)^{\frac{1}{\beta}}\right], & \text{if } 0 < \alpha \leq \frac{\theta_2}{\theta_1+\theta_2}, \\ \\ \frac{1}{\theta_2}\left[\left(\frac{\theta_1}{(1-\alpha)(\theta_1+\theta_2)}\right)^{\frac{1}{\beta}} - 1\right], & \text{if } \frac{\theta_2}{\theta_1+\theta_2} < \alpha < 1. \end{cases}$$

For simplification, we let $d_1 = \frac{\theta_2}{\theta_1+\theta_2}$, $\quad d_2 = \frac{\theta_1}{\theta_1+\theta_2}$. Clearly, $d_1 + d_2 = 1$. Thus,

$$F^{-1}(\alpha) = \begin{cases} h_1(\alpha) = \frac{1}{\theta_1}\left[1 - \left(\frac{d_1}{\alpha}\right)^{\frac{1}{\beta}}\right], & \text{if } 0 < \alpha \le d_1, \\ h_2(\alpha) = \frac{1}{\theta_2}\left[\left(\frac{d_2}{1-\alpha}\right)^{\frac{1}{\beta}} - 1\right], & \text{if } d_1 < \alpha < 1, \end{cases}$$

and

$$F^{-1}(1-\alpha) = \begin{cases} g_1(1-\alpha) = \frac{1}{\theta_2}\left[\left(\frac{d_2}{\alpha}\right)^{\frac{1}{\beta}} - 1\right], & \text{if } 0 < \alpha \le d_2, \\ g_2(1-\alpha) = \frac{1}{\theta_1}\left[1 - \left(\frac{d_1}{1-\alpha}\right)^{\frac{1}{\beta}}\right], & \text{if } d_2 < \alpha < 1. \end{cases}$$

As the above quantile functions have not been centered, the Khattree-Baguhuna's skewness is computed by $\delta = \frac{1}{2}\left[1 + \frac{\int_0^1 F^{-1}(\alpha)F^{-1}(1-\alpha)d\alpha - \mu^2}{\mu_2}\right]$. We consider the computation of the kernel term $\int_0^1 F^{-1}(\alpha)F^{-1}(1-\alpha)$. If $d_1 < d_2$, we have $\int_0^1 F^{-1}(\alpha)F^{-1}(1-\alpha)d\alpha = \int_0^{d_1} h_1(\alpha)g_1(1-\alpha)d\alpha + \int_{d_1}^{d_2} h_2(\alpha)g_1(1-\alpha)d\alpha + \int_{d_2}^1 h_2(\alpha)g_2(1-\alpha)d\alpha$. Now, consider the first and third integrals,

$$\begin{aligned} \mathcal{I}_1 &= \int_0^{d_1} h_1(\alpha)g_1(1-\alpha)d\alpha = \frac{1}{\theta_1\theta_2}\int_0^{d_1}\left[\left(\frac{d_2}{\alpha}\right)^{\frac{1}{\beta}} - 1 - \left(\frac{d_1d_2}{\alpha^2}\right)^{\frac{1}{\beta}} + \left(\frac{d_1}{\alpha}\right)^{\frac{1}{\beta}}\right]d\alpha \\ &= \frac{1}{\theta_1\theta_2}\left[\left(\frac{d_2}{d_1}\right)^{\frac{1}{\beta}}\frac{\beta}{\beta-1}d_1 - d_1 - \left(\frac{d_2}{d_1}\right)^{\frac{1}{\beta}}\frac{\beta}{\beta-2}d_1 + \frac{\beta}{\beta-1}d_1\right] = \frac{d_1}{\theta_1\theta_2}\left[\left(\frac{d_2}{d_1}\right)^{\frac{1}{\beta}}\left(\frac{\beta}{\beta-1} - \frac{\beta}{\beta-2}\right) + \frac{1}{\beta-1}\right] \\ &= \frac{d_1}{\theta_1\theta_2}\left[\frac{1}{\beta-1} - \left(\frac{d_2}{d_1}\right)^{\frac{1}{\beta}}\frac{\beta}{(\beta-1)(\beta-2)}\right]. \end{aligned}$$

Similarly,

$$\begin{aligned} \mathcal{I}_3 &= \int_{d_2}^1 h_2(\alpha)g_2(1-\alpha)d\alpha = \frac{1}{\theta_1\theta_2}\int_{d_2}^1\left[\left(\frac{d_2}{1-\alpha}\right)^{\frac{1}{\beta}} - \left(\frac{d_1d_2}{(1-\alpha)^2}\right)^{\frac{1}{\beta}} - 1 + \left(\frac{d_1}{1-\alpha}\right)^{\frac{1}{\beta}}\right]d\alpha \\ &= \frac{1}{\theta_1\theta_2}\left[\left(\frac{d_2}{d_1}\right)^{\frac{1}{\beta}}\frac{\beta}{\beta-1}d_1 - \left(\frac{d_2}{d_1}\right)^{\frac{1}{\beta}}\frac{\beta}{\beta-2}d_1 - d_1 + \frac{\beta}{\beta-1}d_1\right] = \frac{d_1}{\theta_1\theta_2}\left[\frac{1}{\beta-1} - \left(\frac{d_2}{d_1}\right)^{\frac{1}{\beta}}\frac{\beta}{(\beta-1)(\beta-2)}\right]. \end{aligned}$$

Thus, $\mathcal{I}_1 = \mathcal{I}_3$. Now we consider the second integral term:

$$\begin{aligned} \mathcal{I}_2 &= \int_{d_1}^{d_2} h_2(\alpha)g_1(1-\alpha)d\alpha = \int_{d_1}^{d_2}\frac{1}{\theta_2}\left[\left(\frac{d_2}{1-\alpha}\right)^{\frac{1}{\beta}} - 1\right]\frac{1}{\theta_2}\left[\left(\frac{d_2}{\alpha}\right)^{\frac{1}{\beta}} - 1\right]d\alpha \\ &= \frac{1}{\theta_2^2}\int_{d_1}^{d_2}\left[d_2^{\frac{2}{\beta}}(1-\alpha)^{-\frac{1}{\beta}}\alpha^{-\frac{1}{\beta}} - \left(\frac{d_2}{1-\alpha}\right)^{\frac{1}{\beta}} - \left(\frac{d_2}{\alpha}\right)^{\frac{1}{\beta}} + 1\right]d\alpha \\ &= \frac{1}{\theta_2^2}\left\{d_2^{\frac{2}{\beta}}\left[B\left(d_2; 1-\frac{1}{\beta}, 1-\frac{1}{\beta}\right) - B\left(d_1; 1-\frac{1}{\beta}, 1-\frac{1}{\beta}\right)\right] - d_2^{\frac{1}{\beta}}\frac{\beta}{\beta-1}\left[-d_1^{-\frac{1}{\beta}+1} + d_2^{-\frac{1}{\beta}+1}\right] - d_2^{\frac{1}{\beta}}\frac{\beta}{\beta-1}\left[d_2^{-\frac{1}{\beta}+1} - d_1^{-\frac{1}{\beta}+1}\right] + d_2 - d_1\right\} \\ &= \frac{1}{\theta_2^2}\left\{d_2^{\frac{2}{\beta}}\left[B\left(d_2; 1-\frac{1}{\beta}, 1-\frac{1}{\beta}\right) - B\left(d_1; 1-\frac{1}{\beta}, 1-\frac{1}{\beta}\right)\right] + 2\left(\frac{d_2}{d_1}\right)^{\frac{1}{\beta}}\frac{\beta}{\beta-1}d_1 - 2\frac{\beta}{\beta-1}d_2 + d_2 - d_1\right\} \end{aligned}$$

Recall that $\mu = \frac{\theta_1-\theta_2}{\theta_1\theta_2(\beta-1)} = \frac{d_2-d_1}{\theta_2 d_2(\beta-1)}$ and $\mu_2 = \frac{\beta(d_2-d_1)^2}{d_1^2 d_2^2(\beta-1)^2(\beta-2)} + \frac{2}{\theta_1\theta_2(\beta-1)(\beta-2)}$. Therefore, the Khattree-Bahuguna's skewness is $\delta = \frac{1}{2}\left[1 + \frac{2\mathcal{I}_1+\mathcal{I}_2-\mu^2}{\mu_2}\right]$. We omit the similar derivation for $d_2 < d_1$.

Table 4: Data set of healthy controls and breast cancer patients (Suitably adjusted from the source data reported at *https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Coimbra*)

	Healthy controls		Breast cancer patients	
Pair	BMI	Insulin	BMI	Insulin
1	18.67	6.11	18.37	6.03
2	20.69	3.12	20.83	4.56
3	20.76	7.55	20.83	3.42
4	21.11	3.55	21.08	6.20
5	21.37	3.23	21.30	13.85
6	21.47	3.47	21.36	3.00
7	22.00	3.35	21.51	6.68
8	22.03	2.87	22.21	36.94
9	22.70	4.69	22.22	5.70
10	22.85	3.23	22.50	5.26
11	22.86	4.09	22.66	3.48
12	23.00	4.95	22.83	6.86
13	23.01	5.66	22.89	2.74
14	23.12	4.50	23.14	4.90
15	23.34	5.78	23.62	4.42
16	23.50	2.71	24.22	3.73
17	23.80	6.47	24.24	21.70
18	25.30	3.51	25.51	10.39
19	25.70	8.08	25.59	2.82
20	25.90	4.58	26.56	10.55
21	26.35	5.14	26.56	6.52
22	26.60	4.46	26.67	41.61
23	27.10	26.21	26.67	22.03
24	27.20	14.07	26.84	4.53
25	27.30	5.20	26.85	3.33
26	27.69	3.85	27.18	19.91
27	27.70	6.04	27.64	2.43
28	28.58	4.34	28.44	8.81
29	29.22	5.38	29.14	10.95
30	29.40	10.70	29.15	16.58
31	29.61	5.82	29.30	4.17
32	30.28	4.38	29.38	4.71
33	30.30	8.34	29.67	14.65
34	30.48	5.54	29.78	8.40
35	31.24	4.18	30.48	7.01
36	31.45	9.24	30.80	30.21
37	31.98	4.53	30.84	41.89
38	32.04	18.08	30.92	10.49
39	32.27	5.81	31.22	18.08
40	32.50	5.43	31.23	30.13
41	34.17	6.59	31.25	4.33
42	34.42	23.19	31.25	12.16
43	34.53	4.43	31.64	9.67
44	35.09	5.65	31.98	16.64
45	35.25	6.82	32.05	5.73
46	35.59	3.88	32.46	28.68
47	35.86	8.58	32.46	24.89
48	36.21	15.53	33.18	5.75
49	36.51	14.03	34.84	12.55
50	36.79	10.18	35.56	8.15
51	37.04	6.76	36.05	11.91
52	38.58	6.70	37.11	5.64

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 367–386

Association of Socioeconomic and Demographic Factors With COVID-19 Related Health Outcomes in SAARC Nations

Sangeeta Chakravarty[1], Gurprit Grover[2] and Sanya Aggarwal[2]
[1]*Institute of Economic Growth, Delhi University Enclave, Delhi*
[2]*Department of Statistics, Faculty of Mathematical Sciences, University of Delhi, Delhi*

Received: 11 December 2020; Revised: 11 March 2021; Accepted: 13 March 2021

Abstract

COVID-19 pandemic has reshaped our world in a timescale much shorter than what we can understand and is now a major global health threat. As there was no preparedness on this virus, authorities around the world took restrictive policy measures to control the spread to ensure the wellbeing of the people. This pandemic affected both developed and underdeveloped countries equally. Moreover, existing socioeconomic and demographic characteristics of the countries may be contributing to the variation in health outcomes between countries. This study aims to analyse the influence of socioeconomic and demographic factors on COVID-19 related health outcomes in SAARC nations. The study is important as the objectives behind SAARC are regional integration and economic development of its member countries.

Panel regression analysis and Negative binomial regression are used to identify country specific factors that are associated with COVID-19 related Case Fatality Rate (CFR) and count data, such as, daily cases and active cases, respectively. The findings of the study indicate that increasing CFR are associated with countries having higher cardiovascular death rates, diabetes prevalence, health expenditure (percentage of GDP) and life expectancy. It is also found that co-morbidities such as cardiovascular disease, Tuberculosis and diabetes prevalence are associated with increased national caseloads and mortality, respectively. The study may help government to evaluate policies that can aid in managing the effects of the pandemic by utilizing resources and capabilities in an efficient way.

Key words: COVID-19; Case Fatality Rate; SAARC nations; Socioeconomic factors; Demographic factors; Negative binomial regression; Panel data analysis.

1. Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) causing coronavirus disease 2019 (COVID-19), was first reported in Wuhan, Hubei Province, China in December 2019 [Yang *et al*. (2020)]. It was declared global pandemic by World Health Organisation (WHO) on 11th March 2020 [Cucinotta and Vanelli (2020)]. As on 31st January 2021, more than 100 million people were infected with COVID-19 and 2.2 million have already died [WHO (2020)]. After initial breakout of COVID-19 in China, the epicentre changed to Italy, United Kingdom (UK) and then to United States of America (USA) [Gupta and Misra (2020)]. Most infected cases were in USA followed by India and Brazil [WHO (2020)]. SARS-CoV-2 has a stronger transmission capacity as compared with the SARS-CoV that caused an outbreak

Corresponding Author: Sanya Aggarwal
Email: sanya.aggarwal.stats@gmail.com

of SARS in 2003 [Ma *et al*. (2020)]. Possible modes of transmission of virus causing COVID-19 includes animal-to-human transmission, human-to-human through casual contact, droplets, airborne, fomite, fecal-oral, bloodborne and mother-to-child transmission [World Health Organization (2020)]. Although most people infected with the SARS-COV-2 virus will experience mild to moderate respiratory illness and recover without requiring special treatment. Older people, on the other hand, with underlying medical problems like cardiovascular disease, diabetes, chronic respiratory disease, and cancer are more likely to develop serious illness [World Health Organization (2020)].

The South Asian Association for Regional Cooperation (SAARC) countries comprises 3% of the world's area and home to 21% of the world's total population and comprising of eight nations—Afghanistan, Bangladesh, Bhutan, India, Maldives, Nepal, Pakistan, and Sri Lanka—has experienced the wave of pandemic much later than Europe and America [Shohan *et al*. (2020), WHO (2020)]. SAARC nations share a common regional space with similar geographical conditions and population, yet they differ significantly in the prevalence, severity, mortality and management of the pandemic. The first case of COVID-19 in this region was reported in Nepal on 23rd January 2020 [WHO (2020)]. As on 31st January 2021, India has the highest prevalence of COVID-19 in the region and ranks second globally, on contrary, Bhutan records one death due to the virus [WHO (2020)]. Some of these variations could be ascribed to demographic, social and economic factors, as well as health infrastructure, access to healthcare, political and public health response. Although SAARC countries had gained upper hand in demarcating the initial entry of COVID-19 into the countries, the region is much more vulnerable to its severe impacts. Infectious diseases are the major cause of mortality and morbidity in South Asia [Zaidi *et al*. (2020)]. Recently, World Bank has warned that South Asia faces its worst economic performance in ten years due to this deadly SARS-COV-2 virus. An emergency fund in response to the pandemic has been set up by these nations where each country has voluntarily contributed to secure the people of the region [Augustine (2020)]. But the region is less prepared against pandemic due to poverty, poor medical infrastructure and medical care facilities, as well as the lower number of physicians. An evidence-based study thus becomes imperative to assist policy makers and government in limiting the impact of COVID-19.

Good health improves learning, working production and income and as such health contributes to economic growth and development of the nation. For an unprecedented epidemic such as COVID-19 where individual level data is not available, frequency level estimation such as number of cases, number of deaths, number of active cases, *etc*., becomes a viable choice. Various studies have been conducted on COVID-19 related impacts on SAARC nations. Sultana and Reza (2020) studied the impact of COVID-19 from the perspective of working population of SAARC nations. Shohan *et al*. (2020) examined the onset and transmission of the virus in each SAARC country at an early stage and critically appraised their response with respect to their medical facilities for diagnosis and management. Awasthi (2020) discussed challenges faced by SAARC countries in the wake of COVID-19 pandemic and how India's endeavour is bringing all the nations together in combating the pandemic. Deo *et al*. (2020) predicted the dynamics of COVID-19 pandemic in India. Some studies have reported the effect of country specific factors on COVID-19 around the world. Chaudhry *et al*. (2020) conducted a country level exploratory analysis to assess the impact of timing and type of national health policy/actions undertaken towards COVID-19 mortality and related health outcomes. Yang *et*

al. (2020) studied the impact of COVID-19 in Wuhan, China and suggests that older patients with co-morbidities had increased risk of death.

This paper aims to examine how country-specific socioeconomic and demographic factors effect health outcomes related to COVID-19 in SAARC nations. The importance of selecting SAARC for this study is that the geographical position of some of the member countries is such that they share their borders with China, where the cases first reported. Panel regression analysis and Negative Binomial (NB) regression modelling are utilised to model COVID-19 related health outcomes such as, Case Fatality Rate (CFR), daily infected cases and total active cases against country specific factors.

2. Materials and Methodology

2.1. Data

The study includes twelve country specific factors of eight countries. COVID-19 related health outcomes included in this study are Case fatality rate (CFR), number of reported cases and number of active cases. Publicly available information on COVID-19 related health outcomes such as number of cases, recovered cases and total deaths were extracted from various websites [Roser *et al.* (2020), COVID (2019)]. CFR is defined as the ratio of number of deaths by number of infected cases due to disease over a certain period of time. For any disease to be less severe, the CFR should be less than 1 % [Global Health Observatory (2020)]. The higher CFR suggests that the disease is severe and requires measures by government and individuals to minimise the fatalities. Daily cases are calculated by subtracting total number of cases at time t to total number of cases at time (t–1). Active cases are the number of cases which are neither dead nor recovered but are still infected. It is calculated by subtracting recovered and dead cases from the number of infected cases. Various other rates that are utilized in the study to measure the severity of COVID-19 related health outcomes are recovery rate, percentage of active cases and infection cases per capita. Recovery rate and percentage of active cases are calculated similarly as CFR with numerator changed to number of recovered cases and number of active cases, respectively. Infection per capita is another measure for understanding severity of the disease. It is calculated as the number of infections in each region to the total population in that region over a certain period of time.

Data on country level variables and indices were captured through various sources (see Appendix Table A.1). These includes total population (2019), population density, life expectancy, cardiovascular death rates, diabetes prevalence, GDP per capita and handwashing facilities [Roser *et al*. (2020)]. Other factors included were health expenditure (% of GDP), Tuberculosis (TB) prevalence, age dependency ratio, hospital beds per ten thousand population and proportion of employed population below poverty line [Asian Development Bank (ADB) (2020)]. Global health security (GHS) is another factor that is included in the study for each country [GHS Index Project Team (2019)].

The proportion of the employed population below the international poverty line of US$1.90 per day, also referred to as the working poverty rate, reveals the proportion of the employed population living in poverty despite being employed, implying that their employment-related incomes are not enough to lift them and their families out of poverty and ensure decent living conditions [United Nations SDG indicators (2020)]. Age dependency ratio (% of working population) is the ratio of dependents of people younger than 15 or older than

64, to the working-age population, that is, between 15-64 years of age. Data are shown as the proportion of dependents per 100 working-age population [World Bank, (2019)]. Similarly, the data on stringency index was also obtained for each country at each point of time (Hale *et al.* 2020). The index is published and updated real time by a research group from Oxford university on nine response indicators including school closures, workplace closures, and travel bans, rescaled to a value from 0 to 100 (100 = strictest) [Jayatilleke *et al*. (2020), Hale *et al.* (2020)].

Another index utilised in the study is Global Health Security index (GHS). The index is the comprehensive assessment of 195 countries' health security and related capabilities cross six categories, 34 indicators, and 85 sub-indicators. The six categories are as follows: Prevention of the emergence or release of pathogens; early detection and reporting for epidemics of potential international concern; rapid response to and mitigation of the spread of an epidemic; Sufficient and robust health system to treat the sick and protect health workers; commitments to improving national capacity, financing plans to address gaps and adhering to global norms; and; overall risk environment and country vulnerability to biological threats [GHS Index Project Team (2019)].

As the data is continuously evolving, the period for the study considered is from 25th January 2020 to 31st January 2021. The data is divided into two parts and the point of partition is obtained by plotting the average stringency index per day and recording the date when stringency index was below 60. The date thus obtained is 15th September 2020. For this study, the first phase is considered from 25th January 2020-14th September 2020 and the second phase is considered from 15th September 2020-31st January 2021. We then determine the impact of the socioeconomic and demographic factors on COVID-19 health outcomes in these two periods of the pandemic.

2.2. Statistical models

The descriptive analysis was conducted on COVID-19 related health outcomes of SAARC nations. For modelling the relationship between CFR and country specific variables, Panel regression modelling technique is utilized. Panel regression modelling is used to model longitudinal data.

2.2.1. Panel regression modelling

The basic linear panel models can be described through suitable restrictions of the following general model:

$$y_{it} = \alpha_{it} + \beta'_{it} x_{it} + \mu_{it} \tag{1}$$

where, $i = 1,2,\ldots,n$ is the individual country index, and, $t = 1,2,\ldots,T$ is the time index and μ_{it} is a random disturbance term of mean 0 [Menard (2007), Croissant and Millo (2008)]. When t is same for all countries, it is called balanced data, otherwise it is unbalanced data. The data is recorded from the occurrence of first COVID-19 case in the each of the SAARC country, the data set is thus unbalanced. When the assumption of parameter homogeneity is taken, that is, $\alpha_{it} = \alpha$ for all i, t and $\beta_{it} = \beta$ for all i, t; the resulting model is standard linear pooled model, written as,

$$y_{it} = \alpha + \beta' x_{it} + \mu_{it} \tag{2}$$

To model individual heterogeneity, the error term assumes two separate components, one of which is specific to the individual and does not change over time. This is called the unobserved effects model which can be represented as:

$$y_{it} = \alpha + \beta' x_{it} + \mu_{it} + \varepsilon_{it} \tag{3}$$

The appropriate estimation method for this model depends on the properties of the two error components. If the individual component is missing altogether, pooled OLS is the most efficient estimator for β. To check 'poolability' of the data, pooling tests is conducted i.e., the hypothesis that the same coefficients apply across all individuals. It is a standard F test, based on the comparison of a model obtained for the full sample and a model based on the estimation of an equation for each country [Croissant and Millo (2008)]. Rejection of null hypothesis implies the rejection of poolability and other techniques should be utilized to analyse the data.

2.2.2. Poisson model

For studying the relationship between frequency type dependent variable and other independent variables, Poisson and Negative Binomial modelling are recommended. Poisson regression is typically used to evaluate count data in public health. It is often assumed that the number of events follows a Poisson distribution with a conditional mean μ depending upon a set of regressors x and corresponding parameters β for a participant's linear predictor. Using a log link, we can express the expected number of events for country i as $\mu_i = E(y_i|x_i) = e^{\beta' x_i}$. The Poisson probability distribution of y_i given x_i can be expressed as:

$$P(Y_i = y_i) = \frac{e^{-\mu_i}\mu_i^{y_i}}{y_i!} \tag{4}$$

where, y_i is a non- negative integer. The log likelihood for the model can be expressed as:

$$LL(\beta) = \sum_{i=1}^{n} y_i X'\beta - \mu_i - \log(y_i!) \tag{5}$$

However, this model assumes the variance is equal to mean, an assumption which is often violated [Rose *et al*. (2006)]. The most common alternative for over dispersion of dependent variable over Poisson regression is Negative Binomial (NB) model, which has a built-in dispersion parameter and can account for variance greater than mean [Agresti (2003)].

2.2.3. Negative Binomial (NB) model

The NB regression model allows for over dispersion by introducing an unobserved heterogeneity term for observation [Sheu *et al.* (2004)], *i.e.*, $\mu_i = e^{(\beta' x_i + e_i)}$. We normally assume that $\exp(e_i)$ has a gamma distribution with mean 1 and variance a so that the conditional mean of y_i is still μ_i but the conditional variance of y_i becomes $\mu_i(1 + a\mu_i)$. As a approaches zero, y becomes a Poisson distribution and as a becomes larger the distribution becomes more dispersed. The NB probability distribution for country i is given by:

$$P(Y_i = y_i) = \frac{\Gamma(y_i + a^{-1})}{\Gamma(y_i+1)\Gamma(a^{-1})} \left(\frac{1}{1+a\mu_i}\right)^{a^{-1}} \left(\frac{a\mu_i}{1+a\mu_i}\right)^{y_i} \tag{6}$$

where, $\mu_i, a, and\ \Gamma(.)$ refer to the mean of the count distribution, the NB dispersion parameter, and the gamma function [Rose *et al*. (2006)]. The log likelihood for the model can be expressed as:

$$LL(\beta, a) = \sum_{i=1}^{n}\{\log[\Gamma(y_i + a^{-1})] - \log[\Gamma(y_i + 1)] - \log[\Gamma(a^{-1})] - a^{-1}\log(1 + a\mu_i) + y_i\log(\mu_i) + y_i\log(a) - y_i\log(1 + a\mu_i)\} \quad (7)$$

which can be maximized by iterative methods (preferably Newton–Raphson) to obtain the estimates of β and a.

2.2.4. Model comparison

To compare the predictive performance of NB regression model with that of Poisson regression, common model selection criterion, Akaike information criterion (AIC) is used. AIC is calculated as $-AIC = -2\,log\,L + k$, where L denotes the likelihood function of the model evaluated at maximum likelihood estimates and k is the total number of parameters in the model. The models which had a higher log-likelihood, or a lower AIC value are considered to be the best. Model's goodness of fit was accessed by AIC and Cox and Snell pseudo R-squared statistic. Cox and Snell pseudo R-squared that uses likelihood ratio to assess overall fit compared to null model. It is calculated as :

$$Cox\ and\ Snell\ pseudo\ R - square = 1 - \left[\frac{LR(full\ model)}{LR(null\ model)}\right]^{2/n} \quad (8)$$

where n is the sample size and LR is the likelihood ratio of the model. NB regression does not have an equivalent to the R-squared measure found in ordinary least squares (OLS) regression, pseudo R-square measure are utilised. Its value ranges from 0 to l higher value indicates a better fitting model [Allison (2014)].

Given the limited sample size of 8 countries, the potential independent variables included in the models were identified using forward selection process. Population density was adjusted on logarithmic scale for ease of calculation. The results of the selected regression models were reported in the incidence rate ratio (IRR) where a value less than one suggests a decreased likelihood and a value of greater than one denotes an increased likelihood of the event under investigation. Similar analysis is then carried out on the two parts of the data as explained in Section 2.1.The data was managed in excel and the statistical analysis was carried out using R software.

3. Results

The situation of COVID-19 related health outcomes of 8 SAARC countries as on 31st January 2021 are presented in Table A.2 (see Appendix). India has recorded highest number of cases with 10,757,610 infected individuals, followed by Pakistan 546,428 and Bangladesh 535,139. It is evident that death toll was highest in India with 154,392 people dying due to COVID-19 followed by Pakistan (11,683) and Bangladesh (8,127). Bhutan on the other hand had only one death due to COVID-19. Highest number of recovered and active cases were seen in India. Pakistan and Bangladesh recorded recovered cases at 501,252 and 479,744, respectively.

Only 814 patients were recovered in Bhutan with 44 patients still active. Sri Lanka reports 57,159 recovered and 6,682 active cases. Maldives stands at 14,139 recovered patients with Nepal at 266,336 and Afghanistan at 47,679. Figure A.1 (see Appendix) shows the progression of the epidemic from the first reported case in each of the SAARC nations. Table

1 enlists the percentage of COVID-19 related outcomes. Even though India records highest COVID-19 related cases, but it is observed that infection per capita (percentage of people infected with COVID-19) is highest in Maldives with 2.93 per cent of people infected, followed by Nepal at 0.93 and India at 0.78 per cent. Bangladesh has 0.32 per cent of population infected and Sri Lanka with 0.3 per cent.

Table 1: COVID-19 related events rate as on 31st January, 2021

Countries	Infection per capita	CFR	Recovery rate	Active cases rate
Afghanistan	0.14	4.36	86.65	8.99
Bangladesh	0.32	1.52	89.65	8.83
Bhutan	0.11	0.12	94.76	5.12
India	0.78	1.44	97.00	1.56
Maldives	2.93	0.33	89.26	10.42
Nepal	0.93	0.75	98.29	0.96
Pakistan	0.25	2.14	91.73	6.13
Sri Lanka	0.30	0.49	89.09	10.42

Bhutan reports lowest per capita infection of 0.11 percent, with Afghanistan at 0.14 percent and Pakistan 0.25. It can be observed that percentage of deaths out of total infected cases, also known as CFR, was highest in Afghanistan (4.36) followed by Pakistan (2.14), Bangladesh (1.52) and India (1.44). Maldives and Sri Lanka record the fatality rate at 0.33 and 0.49 percent, respectively. Bhutan, on the other hand, witnessed 0.12 percent fatalities. The highest recovery rate has been recorded in Nepal of 98.29 percent followed by India where 97 percent of COVID-19 infected patients have gained recovery. Bhutan has also witnessed a recovery rate of 94.76 percent while the rate of recuperation in the Bangladesh stands at 89.65 percent. Maldives and Sri Lanka both have 10.42 percent of active cases out of total infected cases whereas Nepal has only 0.96 percent of active cases and India with 1.56 active cases. Figure 1 illustrates these rates from the first reported cases in each of the SAARC nations.

Table 2: Descriptive Statistics of the variables

Variable	N	Mean	Standard Deviation	Minimum	Maximum
Number of cases	8	582574	1963378	1	10757610
Number of recovered cases	8	516435	1827691	0	10434983
Number of deaths	8	9108	29157.71	0	154392
Number of active cases	8	57031	158712.3	0	1017754
Number of daily cases	8	4391	14244.32	0	97894
Case Fatality Rate (CFR)	8	1.27	1.42	0	12.82
Recovery rate (RR)	8	61.47	32.21	0	100
Stringency Index	7	37.26	25.26	2.78	100

Descriptive analysis of the COVID-19 related health outcomes and stringency index recorded for each of the SAARC nations are recorded in the Table 2. On an average 582,574 infected cases, 516,435 recovered cases and 9,108 deaths were recorded for all nations. Average daily cases recorded were 4,391 with total active cases average being 57,031. Average

CFR was 1.27, reaching maximum at 12.82. Average recovery rate recorded was 61.47. The countries' stringency index average was 37.26 with minimum being 2.78 and highest recorded value being 100 (Table 2).

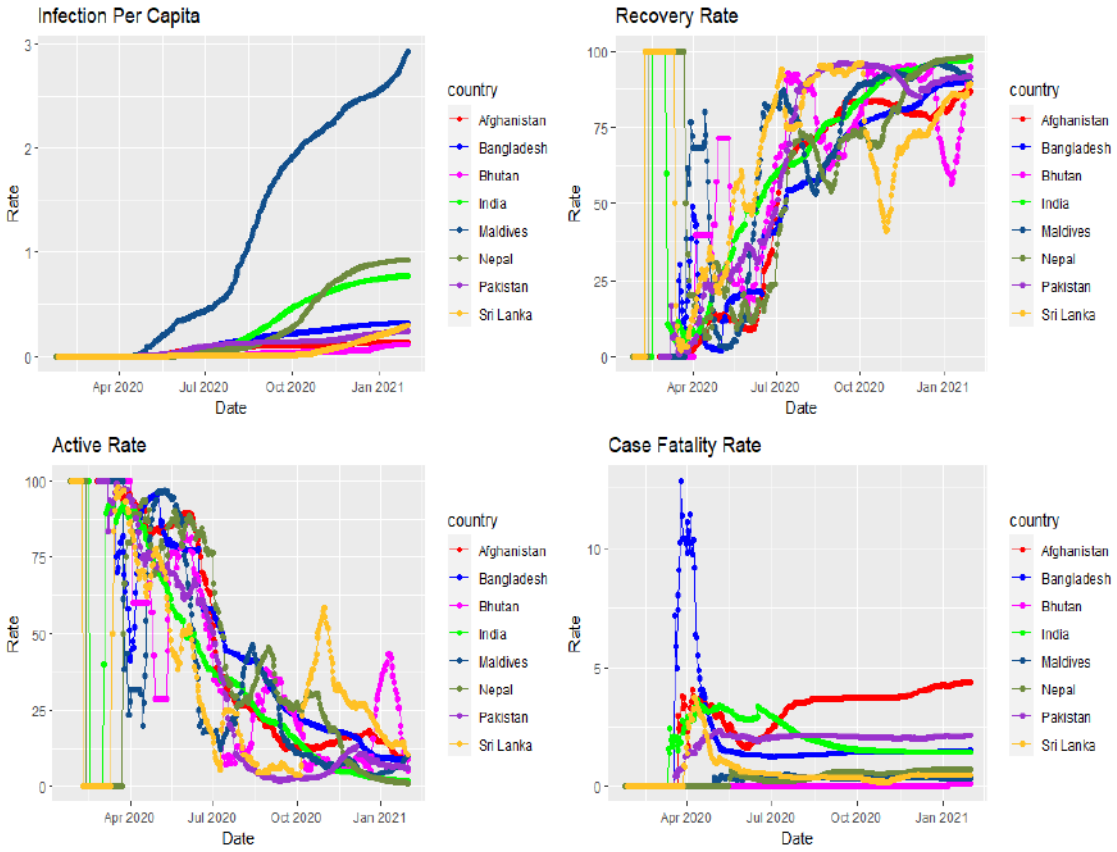


Figure 1: COVID-19 related rates as on 31st January 2021 in SAARC nations

Results of the tests to validate the choice of the regression models are presented in Table 3. The pooling tests to check 'poolability' of the data, *i.e.*, the hypothesis that the same coefficients apply across all individuals, has the normal test statistic –1.369 and *p*-value 0.9145. Null hypothesis is not rejected implying that the individual effect of coefficient is missing, and pooling technique is the most suitable for CFR. The suitable model for daily cases and active cases was assessed using AIC. For daily reported cases, NB model has smaller AIC (40376.93) than Poisson model (49386782), suggesting that NB model is better fit. Similarly, for active cases, NB model shows better fit over Poisson model.

Table 3: Tests to validate the choice of models for each dependent variable

Dependent variable 1: CFR		**Dependent variable 2 : Daily cases**		**Dependent variable 3 : Active cases**	
	Pooling test		AIC		AIC
Test statistic	–1.369	Poisson	49386782	Poisson	537267252
p-value	0.9145	NB	40376.93	NB	56942.78

Table 4: Multiple regression models for the dependent variable 'daily cases'

	Poisson distribution			NB distribution		
Variables	**Estimate**	**S.E.**	***p*-value**	**Estimate**	**S.E.**	***p*-value**
(Intercept)	–18.00	0.02131	<.0001	–27.37	0.6842	<.0001
GHS	0.3421	0.00015	<.0001	0.3344	0.0105	<.0001
Population density	1.1070	0.00163	<.0001	2.168	0.0335	<.0001
TB prevalence	0.0032	0.00003	<.0001	0.0022	0.0008	<.0001
Age dependency ratio	0.0798	0.00023	<.0001	0.1165	0.0079	<.0001
Health expenditure (% of GDP)	0.0379	0.00079	<.0001	0.2194	0.0303	<.0001
Stringency index	0.0119	0.00002	<.0001	0.02441	0.0015	<.0001
AIC	12159689			33821		
Residual deviance	12143840			2998.6		
Degrees of freedom	2452			2452		

Note: S.E. stands for Standard Error; GHS stands for Global Health security index

Poisson and NB regression models fitted to the data with daily cases as dependent variable are presented in Table 4. The AIC of NB regression (33,821) is much lower than that of Poisson regression (12,159,689). It can be observed that residual deviance of NB regression (2,998.6) is much lower suggesting that the model estimates decent amount of variation than the Poisson regression model. The same results can be verified from regression plots in Figure A.2.(see Appendix). As can be observed from Q-Q plot, the points fall relatively closer to the dashed line for NB regression model than Poisson regression. The residual vs leverage plot shows that there are several problematic points in the Poisson model and fewer in the NB regression model. In general, NB regression model provides a better fit for daily reported cases.

Table 5: Multiple regression model for dependent variable 'active cases'

	Poisson distribution			NB distribution		
Variables	**Estimate**	**S.E.**	***p*-value**	**Estimate**	**S.E.**	***p*-value**
(Intercept)	–3.771	0.03597	<0.0001	–3.7711	0.6587	<0.0001
Population density	1.981	0.00290	<0.0001	1.9810	0.0394	<0.0001
Cardiovascular death rate	0.009	0.00003	<0.0001	0.0094	0.0012	<0.0001
Age dependency ratio	–0.095	0.00043	<0.0001	–0.0947	0.0127	<0.0001
Employed people BPL	0.093	0.00010	<0.0001	0.0938	0.0051	<0.0001
Hospital beds(per 10,000 people)	–0.330	0.00005	<0.0001	–0.3301	0.0067	<0.0001
Handwashing facilities	0.121	0.00007	<0.0001	0.1211	0.0043	<0.0001
AIC	147431462			47480		
Residual deviance	147408287			3042.4		
Degrees of freedom	2411			2411		

Note: BPL stands for Below Poverty Line, S.E. stands for Standard Error

Similarly, the results from Poisson regression model and NB regression model for 'Active cases' as the dependent variable are recorded in Table 5. It can be observed that the NB regression model has comparatively smaller residual deviance (3,042.4) and AIC (47,480)

value than Poisson regression model implying that NB regression model has better fit. The results can be verified from the regression plots (Figure A.3, see Appendix). The Q-Q plot shows that the NB regression model should be preferred over the Poisson model. The analysis thus suggests that NB regression model has better than Poisson regression model for active cases variable.

Table 6: Panel regression analysis on COVID-19 Case fatality rate (CFR)

Variables	IRR (95%CI)	IRR (95%CI)	IRR (95%CI)
Case Fatality Rate (CFR)	**25/01/2020-31/01/2021**	**25/01/2020-14/09/2020**	**15/09/2020-31/1/2021**
Cardiovascular death rate	1.025 (1.023-1.027)	1.029 (1.026-1.032)	1.019 (1.018-1.019)
Diabetes prevalence	2.735 (2.469-3.029)	3.551 (3.024-4.157)	1.792 (1.757-1.829)
Hospital beds (per 10,000 people)	0.835 (0.821-0.849)	0.792 (0.771-0.813)	0.905 (0.902-0.908)
Employed people BPL	0.905 (0.891-0.920)	0.882 (0.861-0.905)	0.943 (0.940-0.946)
Health expenditure (% GDP)	1.148 (1.110-1.187)	1.143 (1.085-1.205)	1.145 (1.137-1.152)
Life expectancy	2.249 (2.101-2.409)	2.885 (2.582-3.201)	1.555 (1.535-1.576)
R-squared	0.53	0.42	0.99

Note: IRR stands for incidence rate ratio, CI stands for confidence interval, BPL stands for Below Poverty Line

The findings for association between CFR and country specific factors from Panel regression modelling are presented in Table 6. The significant factors associated with the CFR are cardiovascular death rates, prevalence of diabetes, hospital beds per ten thousand people, employed persons below poverty line, health expenditure (% of GDP) and life expectancy. There was negative association between hospital bed per ten thousand people (IRR = 0.835; 95% CI: 0.821-0.849) and CFR. People employed below poverty line, earning less than US$1.99 per day was also negatively associated with CFR (IRR = 0.905; 95% CI: 0.891-0.920).

Table 7: Negative Binomial regression analysis on COVID-19 daily reported cases

Variables	IRR (95%CI)	IRR (95%CI)	IRR (95%CI)
Daily Cases	**25/01/2020-31/01/2021**	**25/01/2020-14/09/2020**	**15/09/2020-31/1/2021**
GHS	1.397 (1.366-1.430)	1.436 (1.398-1.477)	1.472 (1.435-1.510)
Population density	8.731 (8.082-9.409)	8.633 (7.788-9.544)	8.967 (8.301-9.665)
TB prevalence	1.002 (1.0003-1.004)	1.009 (1.006-1.012)	0.991 (0.989-0.994)
Age dependency ratio	1.123 (1.103-1.143)	1.171 (1.144-1.199)	1.218 (1.187-1.251)
Health expenditure (% GDP)	1.249 (1.184-1.318)	1.037 (0.952-1.141)	1.212 (1.147-1.281)
Stringency index	1.024 (1.019-1.029)	1.057 (1.050-1.063)	1.057 (1.048-1.065)
Cox and Snell pseudo R-square	0.65	0.70	0.84

Note: IRR stands for incidence rate ratio, CI stands for confidence interval, GHS stands for Global Health security index

In contrast, countries with higher cardiovascular death rates (IRR = 1.025; 95% CI: 1.023-1.027), higher diabetes prevalence (IRR= 2.735; 95% CI: 2.469-3.029), spends higher

percentage of GDP on healthcare (IRR= 1.148; 95% CI:1.110-1.187) and have higher life expectancy (IRR = 2.249; 95% CI: 2.101-2.409) had significantly higher CFR. The R-squared value is 0.53 reveals that model explains 53% of the variation in the response variable CFR. This implies that the model has decent fit.

The results of NB regression for daily reported cases are presented in Table 7 in terms of IRR. Socioeconomic and demographic factors positively associated with the increasing daily cases are Global health score (GHS) (IRR = 1.397; 95% CI: 1.366 -1.430), population density (IRR = 8.731; 95% CI: 8.082-9.409), higher prevalence of Tuberculosis (IRR = 1.002; 95% CI: 1.0003-1.004), higher age dependency ratio (% of working population) (IRR = 1.123; 95% CI: 1.103-1.143) and higher stringency index (IRR = 1.024; 95% CI: 1.019-1.029). Higher healthcare expenditure as percentage of GDP (IRR = 1.249; 95% CI: 1.184-1.318) was associated also with higher number of daily reported infected cases. Cox & Snell's pseudo R-squared value is 0.65 which implies that the model has decent fit.

The findings of NB regression analysis of the total active cases on each day (Table 8) suggests that factors significantly associated with increased active cases are: population density (IRR = 7.254 95% CI: 6.710- 7.832), higher cardiovascular death rates (IRR = 1.009; 95% CI: 1.007- 1.012), higher employed people working below poverty line (with less than US$1.99 per day) (IRR = 1.099; 95% CI: 1.088- 1.109) and higher handwashing facilities in the country (IRR = 1.129; 95% CI: 1.120- 1.138). In contrast, higher age dependency ratio (IRR = 0.910; 95% CI: 0.887- 0.933) and more hospital bed available per ten thousand people (IRR = 0.719; 95% CI: 0.709- 0.728) were associated with lower number of active cases in the country. Cox & Snell's pseudo R squared value reported as 0.72 implies that the model has decent fit.

Table 8: Negative Binomial regression analysis on COVID-19 active cases

Variables	IRR (95%CI)	IRR (95%CI)	IRR (95%CI)
Active Cases	**25/01/2020-31/01/2021**	**25/01/2020-14/09/2020**	**15/09/2020-31/1/2021**
Population density	7.254 (6.710-7.832)	7.749 (6.862-8.722)	7.064 (6.727-7.414)
Cardiovascular death rate	1.009 (1.007-1.012)	1.017 (1.013-1.021)	1.003 (1.002-1.005)
Age dependency ratio	0.910 (0.887-0.933)	0.869 (0.837-0.904)	0.934 (0.919-0.949)
Employed people BPL	1.099 (1.088-1.109)	1.086 (1.069-1.103)	1.114 (1.106-1.120)
Hospital beds(per 10,000 people)	0.719 (0.709-0.728)	0.724 (0.709-0.738)	0.708 (0.702-0.714)
Handwashing facilities	1.129 (1.120-1.138)	1.129 (1.114-1.143)	1.132 (1.126-1.138)
Cox & Snell pseudo R-square	0.72	0.63	0.95

Note: IRR stands for incidence rate ratio, CI stands for confidence interval, BPL stands for Below Poverty Line

Further, the results of partitioned data to ascertain any differential change in impact of socioeconomic and demographic variables on the health outcomes are presented in Table 6-8. The point of partition is fixed on 15th September 2020, when the average stringency index was below 60 for the first time (see Appendix Figure A.4) Notable difference appears in the association of Tuberculosis prevalence and daily reported cases. In the phase from 15th September 2020 to 31st January 2021 there is negative association between TB prevalence and daily reported cases (Table 7). The result contrasts with the positive association during the first

phase and full data analysis. All the other results are in line with the complete data analysis (Table 6-Table 8).

4. Discussion

It is important to analyse the significant association between country specific socioeconomic and demographic factors and COVID-19 related health outcomes. The three most affected SAARC countries are India, Bangladesh and Pakistan, respectively, which are also the densely populated nations of SAARC (see Appendix Table A.1). It is evident that the death toll was highest in India followed by Pakistan and Bangladesh. Bhutan on the other hand records only one COVID-19 related death till date and has least number of COVID-19 confirmed cases. On contrary, it is observed that the infection per capita was highest in Maldives, followed by Nepal and India. This is possibly because Maldives has the highest population density with 1,454 people living per square km, which might have resulted in higher transmission rates of infections. It was found that CFR is highest in Afghanistan, being 4.36% deaths of the confirmed cases. The reason of high fatality may be limited health resources and poor health knowledge. Poverty is another issue contributing to worsening the situation of COVID-19 in a war-torn Afghanistan, which is in line other studies, such as, Husseini and Kamil (2020), Sultana and Reza (2020), among others.

The findings of impact of country specific factors on CFR suggests that countries with higher diabetes prevalence and cardiovascular death rates are associated with higher CFR. Also, countries with higher cardiovascular death rate have higher number active cases. Complications are more common in patients with cardiac complication and diabetes with higher mortality than those without it. The results are supported by various studies such as Yang *et al*. (2020) and Zheng *et al*. (2020). Another finding of this analysis is that countries with higher life expectancy have higher CFR. Reports by Onder *et al*. (2020) and Hussain *et al*. (2020) showed similar results that older patients with chronic diseases were at higher risk for severe COVID-19 related mortality.

Health infrastructure is an important factor that affects cases and fatalities. It is found that higher number of hospital beds per ten thousand people was associated with lower CFR and reduced active cases each day. Hospital beds are crucial during an outbreak such as COVID-19 to assess the health facility, as critical cases need medical care in hospital settings for a longer time compared to ordinary patients, thus reducing the active cases and mortality rates. The result is line with the findings of Khan *et al*. (2020) and Blondel and Vranceanu (2020) reporting that COVID-19 fatalities are lower in countries with significant resources dedicated to health care such as hospital beds. Further we have also found that the countries which spend higher percentage of GDP on healthcare witnessed higher number of daily reported cases and CFR. This means that nations that spent more percentage of GDP on healthcare were not insulated from COVID-19 related deaths. This trend was also seen among the wealthy nations such as North America and Europe. There are a few possible explanations for this result between healthcare investment and CFR related to COVID-19 among SAARC nations. Higher healthcare expenditure (% of GDP) was not associated with higher GDP per capita (see Appendix Table A.1). For example, Afghanistan has the lowest GDP per capita but spends more that 10% of GDP on healthcare, which also reports highest CFR. With higher underlying disease burden and higher population, these nations have much lower number of hospital beds and advanced equipment per population, and fewer medical staff to respond to

this unprecedented threat from the pandemic. Similar results have been reported by Sorci *et al*. (2020) and Khan *et al*. (2020).

The findings shows that countries with higher TB prevalence have more reported cases per day. SAARC has 37% of the global burden of TB [STAC (2018)]. It may be the case that chronic respiratory diseases such as active TB could lead to increase in susceptibility to the COVID virus in this region. Other studies have also arrived on similar results [Liu, Bi *et al*. (2020), Maciel *et al*.(2020)].

Our analysis also suggests that countries with higher population density are associated with increased daily cases and active cases. Higher population density may potentially facilitate interactions between susceptible and infectious individuals, which sustains continued transmission and spread of COVID-19. This has been observed in case of Maldives which has the highest population density and has the highest infection per capita. Higher age-dependency ratio is associated with have higher number of daily cases and lower active cases. Younger individuals tend to have a higher proportion of asymptomatic or mild symptoms which are less likely to be detected in testing [Cortis (2020)]. On the contrary, elderly family members with underlying comorbidities are more susceptible to the disease [Liu *et al*. (2020)]. Further, demography of Asia has a lower proportion of elderly individuals than Western nations, with about 85% of the population in India, Pakistan and Bangladesh is younger than 50 years [Sultana and Reza (2020), Gupta and Misra (2020)]. Younger population has shorter disease course than elderly, hence lower active cases [Yang, Hung *et al*. (2020)].

The results also show that the countries with higher proportion of employed people earning below poverty line are associated with increased number of active cases but witnessed lower CFR. About 33.4% of the population in South Asia is living on less than US$1.99 per day income [Sultana and Reza (2020)]. This large population needs to go out to earn living, which increases the chances of infection spread. Also, poor housing facilities and overcrowded accommodation with limited access to personal outdoor space reduces compliance with social distancing thus increasing the overall active cases. The potential reason for lower CFR might include low testing and poor quality of data [Sannigrahi *et al.* (2020), Schellekens and Sourrouille (2020)]. As the huge proportion of this population is underprivileged, illiterate with poor health knowledge and have poor access to healthcare services due to limited income may contribute to the reason behind poor death records.

It is found that even countries that were in relatively more prepared condition according to the GHS index witnessed higher number of daily cases. This may point to the issue that health security is essentially weak in the SAARC nations. The average GHS index of SAARC nations is 36.55 and the global average is just 40.2 [Index Project Team (2019)]. The results also show that countries with higher handwashing facilities have higher reported active cases. The potential reason for such relationship is that high population density in SAARC nations makes it very difficult, if not impossible, to follow basic handwashing, hygiene and physical distancing practices advised during the COVID-19 outbreak, increasing transmission risks, and leading to increased precariousness in living conditions.

Another important finding is that the countries with higher stringency index witnessed increase in daily reported cases. This suggests that the stringent measures did not insulate the nations from spread of infection. Imposing stringent measures is very resource intensive which requires widespread testing and scrupulous contact tracing. The weak healthcare system with

low testing rate in SAARC nations coupled with other factors such as, economically unprivileged population and high population density made maintaining social distance and lockdown challenging in these countries [Niazi *et al.* (2020)].

There are important limitations of the study. First, this study does not consider post COVID-19 factors such as increase in isolation camps, hospital beds, handwashing facilities and number of tests performed in each country. There was missing data for stringency index (Maldives) and handwashing facilities (Sri Lanka). This may have introduced important unintended bias. Missing values were not treated. Various other assumptions are also considered in the study. The basic definition of CFR is utilized for global comparison. The other popular definition, such as, the ratio of number of deaths to the sum of recovered and death cases was not used. Secondly, the asymptomatic COVID-19 population is not considered in the study. Thirdly, the population is assumed to be constant, i.e., it is closed for birth, death and migration. Also, the basic definition of population density is used that is the number of people living per square kilometre. This definition in denominator includes non-habitable lands where no or very little population resides. These assumptions might have led to underestimation of the results.

5. Conclusion

The COVID-19 pandemic has had a significant impact all over the world. During this time, a high number of deaths, public stress and economic damage was witnessed. This study addresses the association of various socio-economic and demographic factors with pandemic related health outcomes in the countries of the SAARC region. The results reveal that diseases, such as tuberculosis, cardiovascular diseases and diabetes, are related with increased mortality and national caseloads. Higher life expectancy is associated with increased mortality. Healthcare infrastructure such as higher number of hospital beds are associated with reduced active cases and mortality. Countries with higher GHS index witnessed higher number of caseloads. Increasing active cases and daily reported caseloads have a positive association with high population density. The findings from the data also suggest that during the later phase of the study period, socioeconomic factors such as, health expenditure (% of GDP), proportion of employed people earning below poverty line, hospital bed, age dependency ratio became more prominent in describing the path of pandemic.

There are many challenges before the SAARC nations, especially in the health sector. Due to this pandemic, healthcare has become a central point of economic and social well-being of all, even more so than before. It has made us realise how important it is to work on all dimensions jointly to save the mankind's present and future. This study will be helpful for evaluation of public health policies in SAARC countries.

Acknowledgement

The authors are very grateful to the reviewers and the editors for their invaluable comments and suggestions, which have helped us to improve the paper substantially.

References

Agresti, A. (2003). *Categorical Data Analysis* (Vol. 482). John Wiley & Sons. (ISBN: 0-471-36093-7).

Allison, P. D. (2014, March). Measures of fit for logistic regression. In *Proceedings of the SAS Global Forum* 2014 *Conference* (pp. 1-13).

Asian Development Bank (ADB). (2020). *Basic Statistics* 2020. Accessed on: 1st September 2020. Website URL: https://www.adb.org/publications/basic-statistics-2020

Augustine P., VID, C. Covid-19 (2020). A New Hope for Global Multilateralism. Website URL:http://ris.org.in/newsletter/diary/2020/Covid%2019%20II/pdf/AUGUSTINE%20PETER.pdf

Awasthi, S. (May,2020). *India and SAARC Combating COVID Pandemic*. BIPSS commentary.

Blondel, S. and Vranceanu, R. (2020). COVID-19 mortality and health expenditures across European countries: the positive correlation puzzle. Available at: SSRN 3679972.

Chaudhry, R., Dranitsaris, G., Mubashir, T., Bartoszko, J. and Riazi, S. (2020). A country level analysis measuring the impact of government actions, country preparedness and socioeconomic factors on COVID-19 mortality and related health outcomes. *EClinical Medicine*, 25, 100464.

Cortis, D. (2020). On determining the age distribution of COVID-19 pandemic. *Frontiers in Public Health*, **8**, 202.

COVID, C. (2019). *Dashboard by the Center for Systems Science and Engineering* (*CSSE*) at Johns Hopkins University (JHU), 2020.

Croissant, Y. and Millo, G. (2008). Panel data econometrics in R: The plm package. Journal of statistical software, **27(2)**, 1-43.

Cucinotta, D. and Vanelli, M. (2020). WHO Declares COVID-19 a Pandemic. Acta bio-medica: Atenei Parmensis, 91(1), 157–160. DOI: https://doi.org/10.23750/abm.v91i1.9397

Deo, V., Chetiya, A. R., Deka, B. and Grover, G. (2020). Forecasting transmission dynamics of COVID-19 in India under containment measures - A time-dependent state-space SIR approach. *Statistics and Applications*, **18**(**1**), 157-180.

Global Health Observatory, WHO. (2020). Cholera case fatality ratio (%). Accessed on: 30th September 2020. Website URL: https://www.who.int/gho/epidemic_diseases/cholera/situation_trends_case_fatality_ratio/en/

Global Health Security (GHS) Index Project Team. (2019). GHS index: building collective action and accountability [Online]. Website URL: https://www.ghsindex.org/wp-content/uploads/2020/04/2019-Global-Health-Security-Index.pdf

Gupta, R. and Misra, A. (2020). COVID19 in South Asians/Asian Indians: Heterogeneity of data and implications for pathophysiology and research. *Diabetes Research and Clinical Practice*, **165**, 108267.

Hale, T., Webster, S., Petherick, A., Phillips, T. and Kira, B. (2020). Oxford covid-19 government response tracker. Blavatnik School of Government, **25**.

Hussain, A., Bhowmik, B. and do Vale Moreira, N. C. (2020). COVID-19 and diabetes: Knowledge in progress. Diabetes Research and Clinical Practice, 108142. DOI: https://doi.org/10.1016/j.diabres.2020.108142

Husseini, A. A. and Kamil, A. A. (2020). Estimating COVID-19 Dynamics in Afghanistan. *Erciyes Med J*, **42**(**4**), 468-73.

Indicators, S. D. G. Metadata repository. United Nations. The Sustainable Development Goal indicators. (2020). Website URL: https://unstats.un.org/sdgs/metadata/files/Metadata-01-01-01b.pdf

Jayatilleke, A. U., Dayarathne, S., de Silva, P., Siribaddana, P., Abeygunawardana, R. A. B., Nieveras, O., de Silva, N. and de Silva, J. (2020). COVID-19 case forecasting model for Sri Lanka based on Stringency Index. medRxiv. DOI: https://doi.org/10.1101/2020.05.20.20103887

Khan, J. R., Awan, N., Islam, M. and Muurlink, O. (2020). Healthcare Capacity, Health Expenditure, and Civil Society as Predictors of COVID-19 Case Fatalities: A Global Analysis. Frontiers in Public Health, **8**, 347.

Liu, K., Chen, Y., Lin, R. and Han, K. (2020). Clinical features of COVID-19 in elderly patients: A comparison with young and middle-aged patients. *Journal of Infection*, **80**(**6**), e14-e18.

Liu, Y., Bi, L., Chen, Y., Wang, Y., Fleming, J., Yu, Y., Gu, Y., Liu,C., Fan, L., Wang,X. and Cheng, M. (2020). Active or latent tuberculosis increases susceptibility to COVID-19 and disease severity. *Medrxiv*. DOI: https://doi.org/10.1101/2020.03.10.20033795

Ma, C., Su, S., Wang, J., Wei, L., Du, L. and Jiang, S. (2020). From SARS-CoV to SARS-CoV-2: safety and broad-spectrum are important for coronavirus vaccine development. *Microbes and Infection*, **22**(**6-7**), 245–253. DOI: https://doi.org/10.1016/j.micinf.2020.05.004

Maciel, E. L. N., Gonçalves Júnior, E. and Dalcolmo, M. M. P. (2020). Tuberculosis and coronavirus: what do we know? *Epidemiologia e Serviços de Saúde*, **29**, e2020128.

Menard, S. (Ed.). (2007). *Handbook of Longitudinal Research*: *Design, Measurement, and Analysis*. Elsevier. (ISBN: 978-0-12-370481-8).

Niazi, A., Kifayat, S., Javed, N. and Khan, M. S. (2020). Assessing the causes for a relatively lower caseload of Covid-19 in South Asia. DOI: https://doi.org/10.31219/osf.io/sf59y

Onder, G., Rezza, G. and Brusaferro, S. (2020). Case-fatality rate and characteristics of patients dying in relation to COVID-19 in Italy. *Jama*, **323**(**18**), 1775-1776.

Rose, C. E., Martin, S. W., Wannemuehler, K. A. and Plikaytis, B. D. (2006). On the use of zero-inflated and hurdle models for modeling vaccine adverse event count data. Journal of biopharmaceutical statistics, **16**(**4**), 463-481.

Roser, M., Ritchie, H., Ortiz-Ospina, E. and Hasell, J. (2020). Coronavirus pandemic (COVID-19). Our world in data. Website URL: https://ourworldindata.org/coronavirus?fbclid=IwAR10byAszhXJ7hAN6VITOfYv8llw5BHhWfa-BsFKW2D_3TqWvg9uv7FnoXY

SAARC Tuberculosis and HIV/AIDS Centre (STAC) (2018). SAARC Epidemiological Response On Tuberculosis, SAARC Tuberculosis & HIV/AIDS Centre (STAC), (www.saarctb.org)

Sannigrahi, S., Pilla, F., Basu, B., Basu, A. S. and Molter, A. (2020). Examining the association between socio-demographic composition and COVID-19 fatalities in the European region using spatial regression approach. *Sustainable Cities and Society*, **62**, p.102418.

Schellekens, P. and Sourrouille, D. M. (2020). Covid-19 mortality in rich and poor countries: a tale of two pandemics? . *World Bank Policy Research Working Paper*, (9260). Available at SSRN: https://ssrn.com/abstract=3614141

Shohan, M.U.S., Alam, A.R.U., Rakhi, N.N., Kabir, M., Siam, M.K.S., Hasan, M.M., Sheriff, R., Karim, M.A., Rahman, A., Moni, M.A. and Arif, M. (2020). Onset, Transmission, Impact, and Management of COVID-19 Epidemic at Early Stage in SAARC Countries. Authorea Preprints. DOI: 10.22541/au.159775079.90952648

Sorci, G., Faivre, B. and Morand, S. (2020). Explaining among-country variation in COVID-19 case fatality rate. *Scientific Reports*, **10**(**1**), 1-11.

Sultana, F. and Reza, H. M. (2020). Are SAARC countries prepared to combat COVID-19 to save young, working-age population? *AIMS Public Health*, **7**(**3**), 440.

WHO (2020). WHO Coronavirus Disease (COVID-19) Dashboard [Online]. Website URL: https://covid19.who.int/?gclid=CjwKCAjwnef6BRAgEiwAgv8mQezLbX7QG5rMuuWdkPVAjZlngItN69DBphoPXo4L0mn17SlE0NUqChoClxMQAvD_BwE

World Bank. (2019). Age dependency ratio (% of working-age population). Website URL: https://data.worldbank.org/indicator/sp.pop.dpnd

World Health Organization. (2020). Modes of transmission of virus causing COVID-19: implications for IPC precaution recommendations: scientific brief, 27 March 2020 (No. WHO/2019-nCoV/Sci_Brief/Transmission_modes/2020.1). World Health Organization.

Yang, M. C., Hung, P. P., Wu, Y. K., Peng, M. Y., Chao, Y. C. and Su, W. L. (2020). A three-generation family cluster with COVID-19 infection: should quarantine be prolonged?. Public Health, 185, 31-33.

Yang, X., Yu, Y., Xu, J., Shu, H., Xia, J., Liu, H., Wu, Y., Zang, L., Yu, Z., Fang, M., Yu, T., Wang, Y., Pan, S., Zou, X., Yuan, S. and Shang, Y. (2020). Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *The Lancet Respiratory Medicine*, ***8*(5)**, 475-481.

Zaidi, A. K., Awasthi, S. and de Silva, H. J. (2004). Burden of infectious diseases in South Asia. *BMJ*, **328(7443)**, 811-815.

Zheng, Y. Y., Ma, Y. T., Zhang, J. Y. and Xie, X. (2020). COVID-19 and the cardiovascular system. *Nature Reviews Cardiology*, **17**(**5**), 259-260.

Appendix

Table A.1: Socioeconomic and demographic factors of SAARC nations

Factors	Afghanistan	Bangladesh	Bhutan	India	Maldives	Nepal	Pakistan	Sri Lanka
Population	38928341	164689383	771612	1380004385	540542	29136808	220892331	21413250
Population density	54.42	1265.04	21.19	450.42	1454.43	204.43	255.57	341.96
GDP per capita	1803.99	3523.98	8708.60	6426.67	15183.62	2442.80	5034.71	11669.08
Cardiovascular death rate	597.03	298.00	217.07	282.28	164.91	260.80	423.03	197.09
Diabetes prevalence	9.59	8.38	9.75	10.39	9.19	7.26	8.35	10.68
Handwashing facilities	37.75	34.81	79.81	59.55	95.80	47.78	59.61	n.a
Life expectancy	64.83	72.59	71.78	69.66	78.92	70.78	67.27	76.98
Health expenditure (% of GDP)	10.20	2.40	3.50	3.70	10.60	6.30	2.80	3.90
Employed persons below poverty line	40.10	9.20	1.30	10.70	1.70	6.10	2.30	0.30
Hospital beds (per 10,000)	3.90	7.95	17.40	5.30	43.00	3.00	6.30	41.50
GHS	32.30	35.00	40.30	46.50	33.80	35.10	35.50	33.90
TB prevalence	189	221	149	199	33	151	265	64
Age dependency ratio	84	49	47	50	31	57	66	53

Source: Roser et al. (2020), Asian Development Bank (ADB) (2020), GHS Index Project Team (2019)

Table A.2: COVID-19 situation as on 31st January 2021

Country	Confirmed	Deaths	Recovered	Active
Afghanistan	55023	2400	47679	4944
Bangladesh	535139	8127	479744	47268
Bhutan	859	1	814	44
India	10757610	154392	10434983	168235
Maldives	15841	52	14139	1650
Nepal	270959	2029	266336	2594
Pakistan	546428	11683	501252	33493
Sri Lanka	64157	316	57159	6682

Source: Our World in Data (Roser *et al.*(2020))

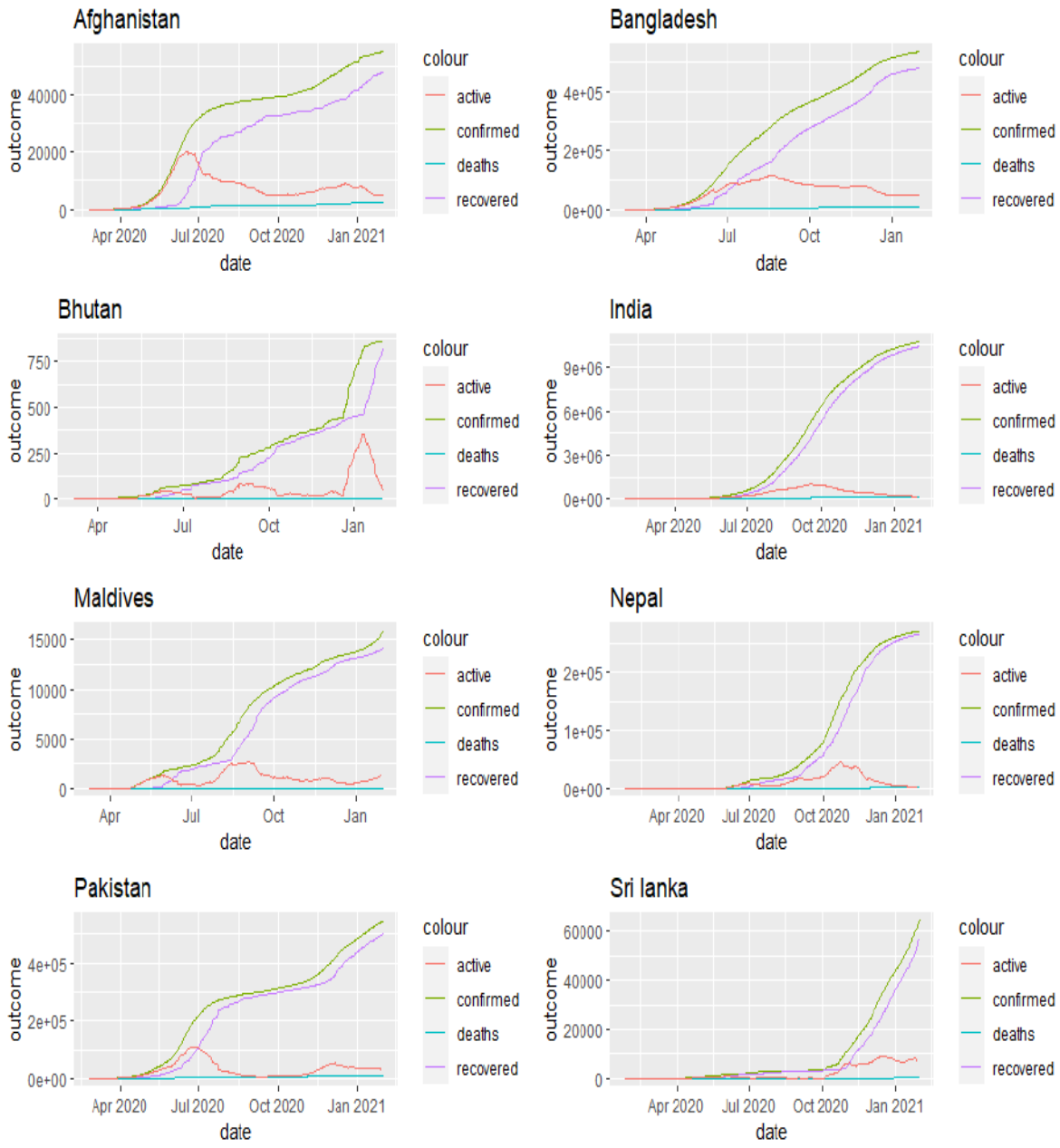


Figure A.1: COVID-19 related outcomes in SAARC nations

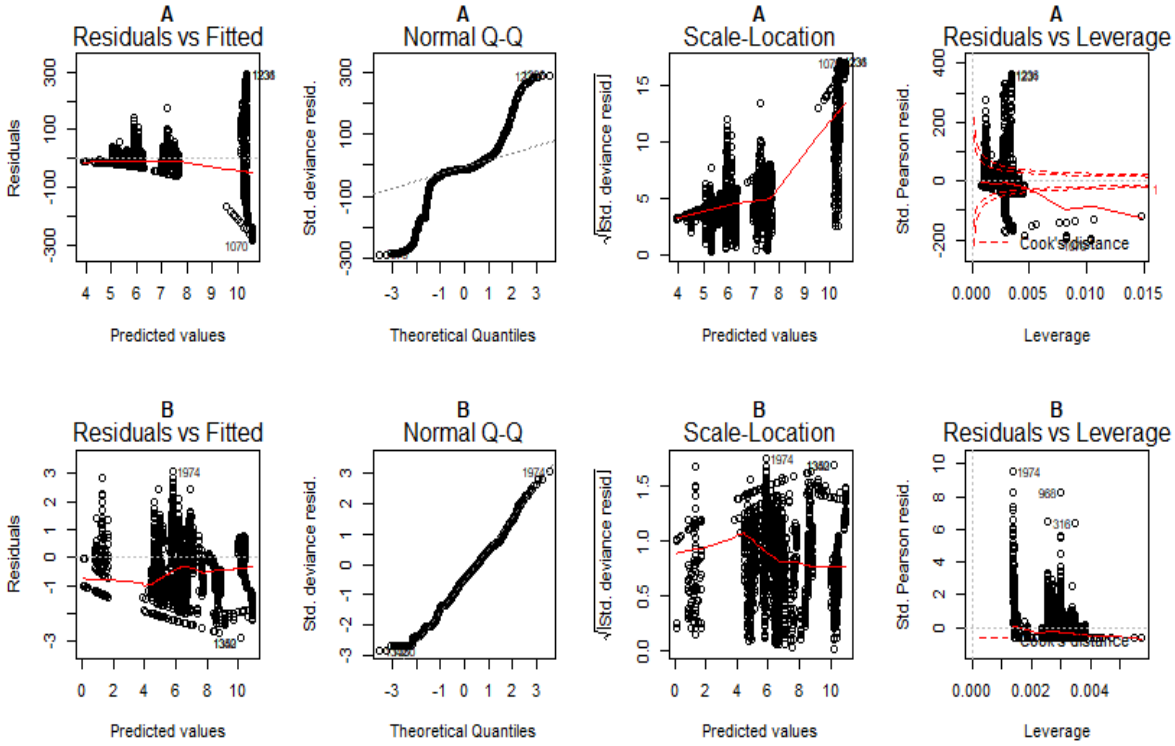


Figure A.2: Poisson regression plot (A) and Negative Binomial regression plot (B) for 'Daily cases'

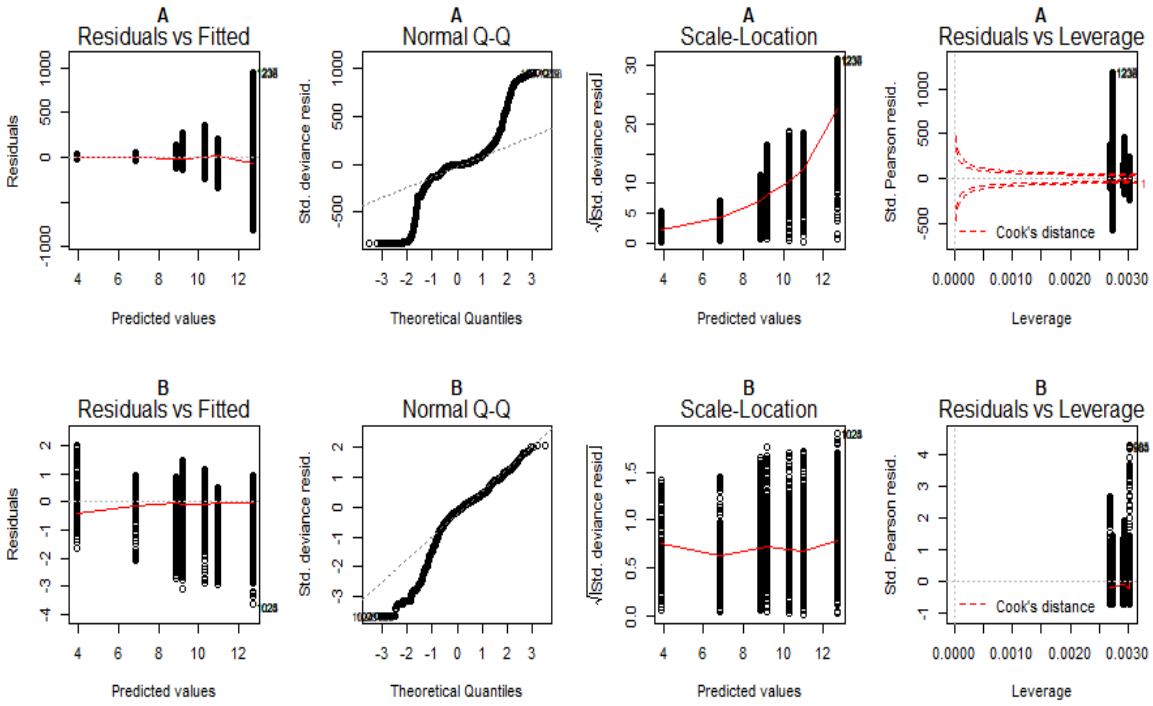


Figure A.3: Poisson regression plot (A) and Negative binomial regression plot (B) for 'Active cases'

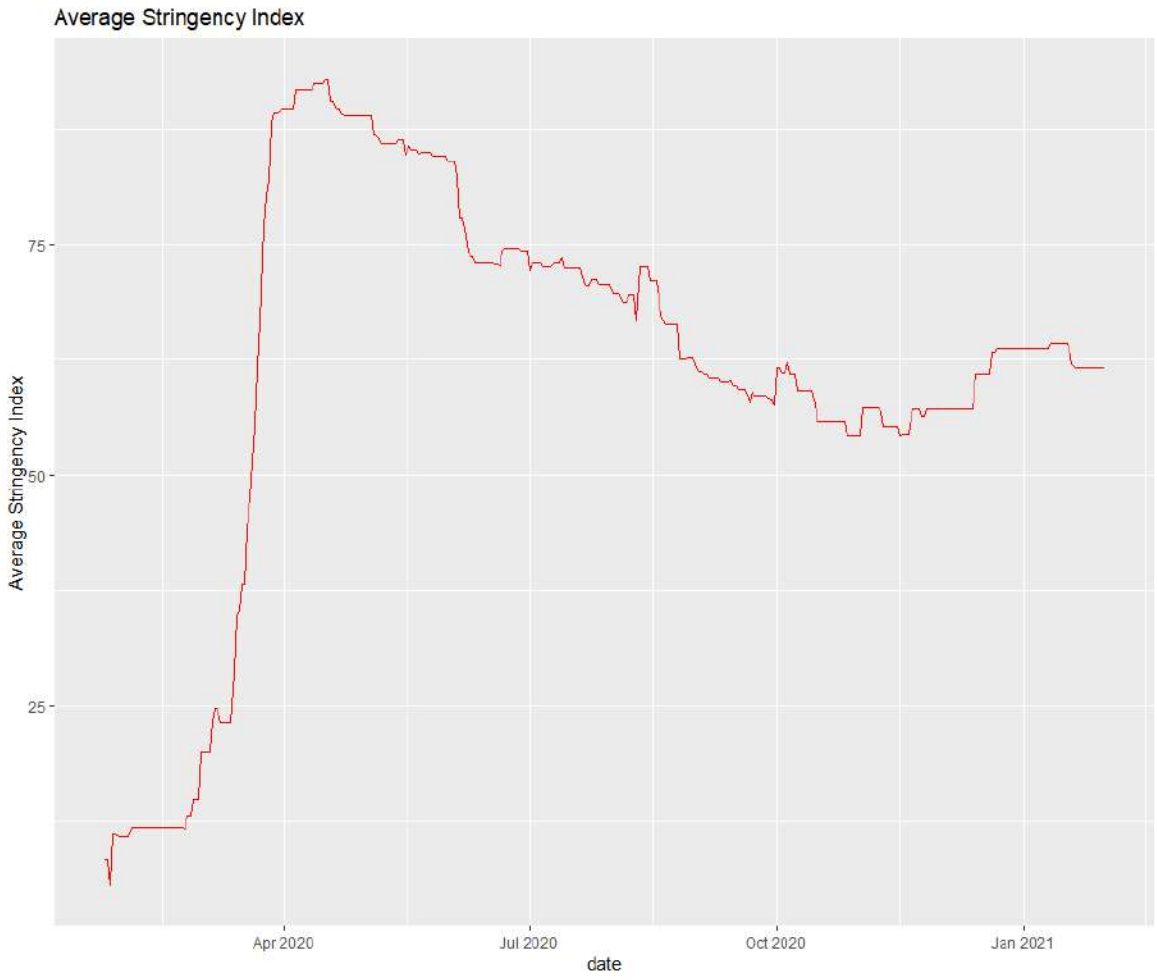


Figure A.4: Average stringency index per day among SAARC nations

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 387-399

Experimental Designs for Alley Cropping to Estimate Shrub × Grass Interaction

Murari Singh
International Center for Agricultural Research in the Dry Areas (ICARDA), Amman, Jordan
and
Department of Mathematics and Statistics, Concordia University, Montreal, Canada

Received: 17 July 2020; Revised: 15 February 2021; Accepted: 13 March 2021

Abstract

An alley cropping comprises rows of perennial shrubs/trees bordering the alleys of grasses/crops. An appropriately chosen alley cropping provides improvement in feeds for small ruminants, food for human consumption, and contributes to economic security and environmental sustainability. Several experimental designs and statistical models are presented. The experimental/environment designs considered are the complete block with or without split-plot frames for the self-borders and partial diallel borders in shrubs and alley experimental units. The treatment designs include a factorial structure of shrub-borders and grasses. The linear models consisting of shrubs effects, grasses effects and their interaction with different structures are considered with parameters defined for shrub × grass interaction. A statistical analysis of the alley- responses will be illustrated with a simulated dataset.

Key words: Alley cropping; Shrub and grass effects and interaction; Self and diallel designs; Blocks; Split plots.

1. Introduction

An alley cropping, an agroforestry system, comprises rows of perennial shrubs or trees bordering the alleys of grasses/crops, is a low input system for forage and food production and serves as a mechanism for sustainable agriculture. With suitable choice of crop, shrub or tree species in the system it supports diverse needs of human and other domestic animals, arrest the land degradation and soil erosion, and plays a major role in mitigating climate change. Alley cropping manages the soil nutrients more effectively between the species, *e.g*., perennial trees/shrubs and annual crops, and different layers of soil depth. A wide range of references are available on various types of crop production systems including alley cropping (Solaimalai *et al*., 2005; AFNTA 1992a, b). Rangeland and forage development studies aim at evaluation of interference of shrubs (*e.g*., saltbush Atriplex) with the grasses or fodder/forage crops (*e.g.,* vetch/barley).

Experimental designs and data analysis for evaluating shrub × grass interaction are presented here. These designs can also be used to estimate main effects and interaction of crops involved in inter-cropping systems. Section 2 presents the construction of experimental

Corresonding Author: Murari Singh
Email: mandrsingh2010@gmail.com

designs and proposed linear models for estimation of effects and interactions. For one of the experimental designs discussed here, the proposed method of estimation has been illustrated with simulated data.

2. Experimental Designs and Models for Statistical Analysis

Consider a set of s shrubs denoted by $S_1, \ldots, S_i, \ldots, S_s$ for planting as the borders and a set of g grasses/crops $G_1, \ldots, G_j, \ldots, G_g$ for the alleys. The following frames of experimental units, or shrub–grass plots will be considered. Experimental units receive: (1) combinations of shrubs and grasses, or (2) shrubs with long borders and all grasses in smaller alleys within these borders. The following two treatment designs, (1) self-borders and grasses combination, and (2) diallel-borders and grasses combinations, can be implemented with any one of the above two frames of the experimental units. The resulting designs may or may not share borders between two alleys. In case they do, search for appropriate covariance structures for grass plot errors would be needed. Examples of such designs are given in the following schemas along with models for data analysis.

3. Designs for Non-Shared Borders

Consider the case where the borders are not shared between the alleys, *i.e.*, same shrub does not affect the grasses on its opposite sides of alleys.

a. Self- borders

In the self-border situation is defined here as the one where the same shrubs serve as the border of a grass plot on its both sides.

Design 1. Self-borders of shrubs and grasses combinations in Randomized Complete Block (RCB) design

Method of construction: Get all the possible combinations of shrubs ($S_1, \ldots, S_i, \ldots, S_s$) and grasses ($G_1, \ldots, G_j, \ldots, G_g$). For a combination ($S_iG_j$), the grass G_j will have shrub S_i on both (left and right) borders. These sg combinations are randomized independently within each of the r complete blocks. Figure 1 shows an example of randomized plan for $s = 4$ shrubs ($S_1, \ldots, S_4$), $g = 3$ grasses (G_1, G_2, G_3), for one replicate.

Replicate	1											
Left border	S_1	S_1	S_2	S_2	S_1	S_3	S_4	S_4	S_3	S_2	S_4	S_3
Alley	G_2	G_3	G_2	G_3	G_1	G_3	G_2	G_1	G_2	G_1	G_3	G_1
Right border	S_1	S_1	S_2	S_2	S_1	S_3	S_4	S_4	S_3	S_2	S_4	S_3
Plots	101	102	103	104	105	106	107	108	109	110	111	112

Figure 1: Schema of a randomized plan for 4 shrubs (S_1, S_2, S_3, S_4), 3 grasses (G_1, G_2, G_3), self-borders, factorial in RCB design, one replicate shown.

Statistical model for response of grasses (under Design 1)

Let $y_{i,jj,l}$ = response from the alley under grass G_i or i, self-borders (left, right): (S_j, S_j) or jj, block/replicate l, μ = general mean; β_l = Effect of block l; γ_i = effect of grass i; ψ_j =

effect of borders, jj , under shrub j from both sides; δ_{ij} = interaction between grass i and shrub borders jj ; $i = 1,\ldots,\mathrm{g}$; $j = 1,\ldots,s$; and $l = 1,\ldots,r$.

The following response model can be assumed:

Response = general mean + block effect + grass effect + shrub-effect + shrub × grass interaction + Error, or,

$$y_{i,jj,l} = \mu + \beta_l + \gamma_i + \psi_j + \delta_{ij} + \varepsilon_{i,jj,l}$$

where independently distributed errors $\varepsilon_{i,jj,l} \sim N(0,\sigma^2)$.

For generating this class of experimental design and carrying out data analysis, modify the "Randomize" directive in the Genstat software (VSN Inc., 2015) codes given in the Appendix.

Design 2. Self-borders of shrubs in main plots in RCB design and grasses in sub-plots.

Method of construction: Get the each of shrubs (S_1, …, S_i, …, S_s) as both the borders long enough to accommodate the plots of all the grasses (G_1, …, G_j, …, G_g). Randomize the shrubs within each block. Randomize the grasses within each shrub border. In this way shrubs form main-plots and grasses subplot within each of the r complete blocks. Figure 2 shows an example of randomized plan for s = 4 shrubs (S_1, S_2, S_3, S_4) and g = 3 grasses (G_1, G_2, G_3), for one replicate.

Replicate	1											
Left border	S_2	S_2	S_2	S_3	S_3	S_3	S_1	S_1	S_1	S_4	S_4	S_4
Alley	G_1	G_3	G_2	G_2	G_3	G_1	G_1	G_2	G_3	G_1	G_2	G_3
Right border	S_2	S_2	S_2	S_3	S_3	S_3	S_1	S_1	S_1	S_4	S_4	S_4
Plots	101	102	103	104	105	106	107	108	109	110	111	112

Figure 2: Schema for a randomized plan for 4 shrubs (S_1, S_2, S_3, S_4), 3 grasses (G_1, G_2, G_3), self-borders, split-plot (Shrub-borders main plot) in RCB design, one replicate.

Statistical model for response of grasses (under Design 2)

Response = general mean + block effect + shrub-effect + Error (a) [Block × Shrub interaction] + grass effect + Shrub × grass interaction + Error(b), or,

$$y_{i,jj,l} = \mu + \beta_l + \psi_j + (\beta\psi)_{jl}[= \mathrm{E}\,rror(a)] \;\; + \gamma_i + \delta_{ij} + \varepsilon_{i,jj,l}[= Error(b)]$$

For generating this class of experimental design and carrying out data analysis, modify the "Randomize" directive in the Genstat software codes given in the Appendix.

b. Diallel- borders:

Different shrubs on the borders will be used in the following two designs.

Design 3. Combinations of shrub diallel-borders and grasses in RCB design

Method of construction. We create borders of shrubs (S_1, …, S_i, …, S_s), by selecting them from a diallel crosses plan in s lines, say, $S_iS_{i'}$. It may be noted that there are no genetic crosses between two shrubs being made, but the combinations of lines that would have been in a cross are used for the borders. Such combinations of shrubs will be called diallel-borders. The all possible cominations of the diallel- border and grasses are randomized within each of the r blocks. Figure 3 shows an example of randomized plan for $s = 4$ shrubs (S_1, S_2, S_3, S_4), with dillel-borders (S_1S_3, S_3S_4, S_4S_2, S_2S_1), and $g = 3$ grasses (G_1, G_2, G_3), for one replicate.

Replicate	1											
Left border	S_1	S_4	S_2	S_2	S_1	S_3	S_4	S_1	S_3	S_2	S_3	S_4
Alley	G_2	G_1	G_1	G_3	G_3	G_3	G_2	G_1	G_1	G_2	G_2	G_3
Right border	S_3	S_2	S_1	S_1	S_3	S_4	S_2	S_3	S_4	S_1	S_4	S_2
Plots	101	102	103	104	105	106	107	108	109	110	111	112

Figure 3: Schema for a randomized plan for diallel-borders in 4 shrubs, and 3 grasses as factorial combinations in RCB design, one replicate.

In case of diallel-borders, the number of borders (shrub pairs) p say, may not necessarily be equal to s , the number of shrubs. For generating this class of experimental designs based on diallel boders, we may use the partial crosses designs presented in Curnow and Kempthorne (1961), Curnow (1963), Arya (1983), Singh and Hinkelmann (1990) among other papers, and also reviewed in Singh *et al.* (2012). These designs are constructed for estimation of general combining ability (gca) effects while specific combining ability (sca) effects are assumed absent or can be ignored. In case of the complete diallel crosses, sca effects are also estimable.

Statistical model for response of grasses (under Design 3)

Let $y_{i,jk,l}$= response from the alley under grass i, diallel-borders (left, right): (S_j, S_k) or jk (shrub j left border and shrub k on the right) and block/replicate l

A statistical model for the response is

$$y_{i,jk,l} = \mu + \beta_l + \gamma_i + \psi_j + \psi_k + \psi_{jk} + \delta_{ij} + \delta_{ik} + \delta_{ijk} + \varepsilon_{i,jk,l}$$

In the above model, the parameters ψ_j in the alley cropping design is the general effect of shrub S_j (irrespective of border direction) on the grasses (gesg) and is equivalent to the gca in the case of partial dial crosses. The ψ_{jk}, is the specific effect of the shrub borders (S_j, S_k) on the grasses (sesg) and would be equivalent to the sca in the diallel crosses situation. The quantity δ_{ij} is the interaction between shrub effect ψ_j and grass effect γ_i and may be termed as grass-specific general effect of shrub S_j (irrespective of border direction) on the grass (gs-ggseg), and δ_{ijk} is grass-specific specific effect of the shrub borders (S_j, S_k) on the grass (gs-sesg). Errors $\varepsilon_{i,jk,l} \sim N(0, \sigma^2)$.

There may be situations where the following assumption may apply.

Assumption: sesg ψ_{jk} and gs-sesg δ_{ijk} may be absent or negligible

In this case the model reduces to

$$y_{i,jk,l} = \mu + \beta_l + \gamma_i + \psi_j + \psi_k + \delta_{ij} + \delta_{ik} + \varepsilon_{i,jk,l}$$

Further, ψ_j's under the designs 1 and 2 (self-borders) would be different from those under the diallel borders. However, in case ψ_{jk} (specific border combination effects) are absent, then ψ_j's under Designs 1 and 2 would be twice of those under Design 3 and Design 4.

Design 4. Diallel-borders in main plots in RCB design and grasses in sub-plots

Method of construction: Get the each of diallel combinations of shrubs ($S_iS_{i'}$) as the borders long enough to accommodate the plots of all the grasses ($G_1, \ldots, G_s$). Randomize theses diallel borders shrubs within each block. Randomize the grasses within each diallel-borders of the shrubs. In this way pair of shrubs (diallel-borders) form main-plots and grasses subplot within each of the r complete blocks. Figure 4 shows an example of randomized plan for s = 4 shrubs (S_1, S_2, S_3, S_4) with dillel-borders ($S_1S_3, S_3S_2, S_2S_4, S_4S_1$) in main-plots, and g = 3 grasses (G_1, G_2, G_3), for one replicate.

Replicate	1											
Left border	S_2	S_2	S_2	S_4	S_4	S_4	S_1	S_1	S_1	S_3	S_3	S_3
Alley	G_2	G_3	G_1	G_3	G_1	G_2	G_3	G_2	G_1	G_2	G_3	G_1
Right border	S_4	S_4	S_4	S_1	S_1	S_1	S_3	S_3	S_3	S_2	S_2	S_2
Plots	101	102	103	104	105	106	107	108	109	110	111	112

Figure 4: Schema for a randomized plan for 4 shrubs, 3 grasses, diallel-borders, split-plot (Shrub-borders main plot) in RCBD.

Statistical model for response of grasses (under Design 4)

$$y_{i,jk,l} = \mu + \beta_l + \psi_j + \psi_k + \psi_{jk} + (\beta\psi)_{jk,l}[= \mathrm{E}\,rror(a)] + \gamma_i + \delta_{ij} + \delta_{ik} + \delta_{ijk} + \varepsilon_{i,jk,l}[= Error(b)]$$

Assumption: sesg ψ_{jk} and gs-sesg δ_{ijk} may be absent or negligible

$$y_{i,jk,l} = \mu + \beta_l + \psi_j + \psi_k + (\beta\psi)_{jk,l}[= \mathrm{E}\,rror(a)] + \gamma_i + \delta_{ij} + \delta_{ik} + \varepsilon_{i,jk,l}[= Error(b)]$$

Design for diallel boders as discussed in Design 3 can be used for conducting the trial in split-plots with diallel-borders in mainplots and grasses in sub-plots. The codes for generating the Design 4 are given in the Appendix.

Estimation of the effects and interactions

A practical approach would be to estimate the response of the combinations of shrub-borders and gasses with adjustment for block differences, covariates for slope and fertility trend in the alleys, spatial error structures. Let the adjusted mean for the treatment combination: grass

i and diallel-border (left, right) (j,k) be denoted by $\bar{y}_{i,jk}$. In vector notation, we can use $\bar{y} = (\bar{y}_{1,12}, \bar{y}_{1,13}, \bar{y}_{1,1s,} ..., \bar{y}_{g,s-1s})'$. One may use all pairs of shrubs (S_j, S_k), equivalent to (S_k, S_j) as borders, but limited resouces may lead to the choice of partial diallel-borders. Based on a simple cyclic structure in shrubs may give a set of diallel-borders as: $(S_1, S_2), (S_2, S_3), ..., (S_s, S_1)$, which could be chosen for all the replicates, or even a better spread could be carried over the replication by using a different spacing between the shrub numbers, *e.g.*, $(S_1, S_3), (S_3, S_5), ..., (S_\_, S_\_)$ in replicate 2, where the subscripts "_" stand for appropriate shrub numbers, *etc*. Let the estimated variance covariance of vector $\bar{y}$ be denoted by $\hat{\Sigma}$. For the full factorial of border and alley treatment factors in an RCB design with r replicates and estimated residual mean square $\hat{\sigma}^2$, $\hat{\Sigma} = (\hat{\sigma}^2 / r)I$. Let the grass effects, shrub effects and their interaction be represented in vector form respectively as:

$$\gamma = (\gamma_1, ..., \gamma_g)', \ \psi = (\psi_1, ..., \psi_s)' \text{ and } \delta = (\delta_{11}, \delta_{12} ..., \delta_{1s}, ..., \delta_{g1}, \delta_{g2}, ..., \delta_{gs})'.$$

Let the interaction between grass and border combinations (not the shrubs) be denoted by $\phi = (\phi_{11}, \phi_{12} ..., \phi_{1p}, ..., \phi_{g1}, \phi_{g2}, ..., \phi_{gp})'$.

Thus $\phi_{im} = \delta_{ij} + \delta_{ik}$ where m stands for the border comprising of the shrubs S_j and S_k; $m = 1, ..., p$.

A model for estimation of γ, ψ and ϕ may be wriiten as

$$\bar{y} = \mu J + X_1 \gamma + X_2 \psi + X_3 \phi + \bar{\varepsilon}$$

where μ is general mean, J a vector of 1s and length of $\bar{y}$, and vector of mean errors with $\bar{\varepsilon} \sim \text{MVN}(0, \hat{\Sigma})$.

Conditions on the vectors of effects are: $\gamma' J = 0$, $\psi' J = 0$ and more than one conditions on the interaction vector: $(I_p \otimes J_g')\phi = 0_{p,1}$ and $(J_p' \otimes I_g)\phi = 0_{1,g}$.

The estimation can have one of the several approaches, particularly in case of orthogonal structure between grasses and diallel-borders.

Approach 1: One can estimate grasses and borders effects and interaction using ANOVA directives. The border effects overall the grasses or for individual grasses data can be modelled by fitting columns of X_2 (no intercept) to estimate ψ s and δ s respectively.

Approach 2: Another could be based on matrices but still using the ANOVA estimates of border effects with variance-covariance matrix or ignoring the covariances. This may be completed in the following two stages:

Stage 1: Estimate γ gamma and ψ, we can fit the general model, ignoring ϕs and fitting a reduced model for $\bar{y} \sim MVN(X\beta, \hat{\Sigma})$, where $X = [J : X_1 : X_2]$ of order $(p, 1+g+s)$ and $\beta = (\mu, \gamma', \psi')'$.

Using Rao (1973), $\hat{\beta} = (\hat{\mu}, \hat{\gamma}', \hat{\psi}')' = S^{-1}Q$ where,

$S = X'\hat{\Sigma}^{-1}X$ and $Q = X'\hat{\Sigma}^{-1}\bar{y}$, assuming that the design keeps matrix S non-singular, otherwise replace S^{-1} by its Moore-Penrose psuedoinverse denoted by S^{+}.

Estimated variance-covariance matrix of $\hat{\beta}$ is $D(\hat{\beta}) = S^{-1}$.

Borders × grass interaction vector ϕ can be estimated as the residual vector

$$\hat{\phi} = \bar{y} - X\hat{\beta} \text{ with } \mathrm{D}(\hat{\phi}) = \Sigma - XS^{-1}X' = \Sigma^{*}, \text{ say.}$$

Actually, the variance-covariance matrix $D(\hat{\phi})$ may be available along with $\hat{\phi}$ while using ANOVA in any software, *e.g.*, Genstat (VSN Inc. 2015).

Stage 2: Next step would be to partition $\hat{\phi}$ into δ's estimates as follows. Obtain a matrix Z with its column number $i_j = j + (i-1)s$ obtained by element-wise multiplication of i th column of X_1 and j th column of X_2, *i.e.* Schur multiplication of all possible cross combinations between columns of X_1 and X_2. The order of Z is $p \times gs$. We can obtain gs parameters in δ s by solving the equation:

$$\hat{\phi} = Z\delta, \text{ where } D(\hat{\phi}) = \Sigma^{*}$$

to obtain

$$\hat{\delta} = (Z'\Sigma^{*+}Z)^{+}Z'\Sigma^{*+}\hat{\phi} \text{ and } D(\hat{\delta}) = (Z'\Sigma^{*+}Z)^{+}$$

where for a matrix A, A^{+} denotes its Moore-Penrose psuedoinverse.

Optimal design: Optimality and efficiency of the design can be studied in terms of the respective covariance matrices for $\hat{\gamma}$ gamma, $\hat{\psi}$ s and $\hat{\delta}$'s. There could be alternative options to estimate the effects using a software. Genstat codes are given on the set of data generated for illustration in the following section.

4. Shared Borders

Design 5. Sharing of borders between the alleys would lead to a resource saving design. However, data analysis may be based on a relatively more complex model due to the feature that the same shrub may affect grasses on its opposite sides of alleys. Self-borders or diallel-boders can be used. Due to sharing of the same border between the alleys the randomization of the shrubs as borders would become quite restricted.

Method of construction: The construction can be easily explained using an example, which requires that one has a partial diallel design. Let the partial diallel considered for the borders be written as S_1S_3, S_3S_2, S_2S_4, S_4S_3, S_3S_1…, while the grasses (G_1, G_2, G_3) are in alleys, that is, the sub-plots. We will have much restricted randomization as these partial diallels in shrubs now follow a sequence where the the shrub on right of a diallel is same the left of the diallel that follows in the sequence. Hence the randomization can be done within the set of all the shrubs. In the present case, the shrubs pair (S_1, S_3) from the partial diallel, one border, say left-border is S_1 while the right-border is S_3. This right-border S_3 is same as the left border in S_3S_2, therefore S_3 is shared border. Similarly, S_2 in (S_3S_2) is shared border with (S_2S_4), and so on (Figure 5).

Left border	S_1	S_1	S_1
Alley	G_1	G_3	G_2
Shared border	S_3	S_3	S_3
Alley	G_2	G_3	G_1
Shared border	S_2	S_2	S_2
Alley	G_1	G_2	G_3
Shared border	S_4	S_4	S_4
Alley	G_3	G_1	G_2
Shared border	S_3	S_3	S_3
Alley	G_2	G_3	G_1
Shared border	S_1	S_1	S_1
Alley	G_3	G_2	G_1
Shared border		.	
Alley		.	

Figure 5. Schema for a randomized plan for 4 shrubs (S_1, S_2, S_3, S_4) using the partial diallel (S_1S_3, S_3S_2, S_2S_4, S_4S_3, S_3S_1... more) and 3 grasses with shared diallel-borders in a split-plot (Shrub-borders main plot) in RCB design.

Statistical model for response of grasses (under Design 5)

In this case correlated responses may be assumed, and covariance modelling would a worthy exercise to induct in the analysis. Model:

$$y_{i,jk,l} = \mu + \beta_l + \psi_j + \psi_k + (\beta\psi)_{jk,l}[= \mathrm{E}\,rror(a)] \quad + \gamma_i + \delta_{ij} + \delta_{ik} + \varepsilon_{i,jk,l}[= Error(b)]$$

Correlated model structures:

$\mathrm{Cov}((\beta\psi)_{jk,l}, (\beta\psi)_{km,l})$ and $Cov(\varepsilon_{i,jk,l}, \varepsilon_{i,km,l})$ may need to be simplified using a criterion such as Akaike Information Criterion (AIC) (Akaike, 1974). The selected covariance structure(s) can then be used for estimation of the effects and interaction.

5. An Illustration

A dataset was generated for experimental design situation, Design 4 using the following set of values of effects taken for random generation of data (Table 1).

Table 1: **Values of parameters to model the response from Design 4**

General mean: $\mu = 5$

Coefficient of variation based on main-plot error = 10%

Coefficient of variation based on subplot error = 15%

Block effects: β_l (l=1…3) = –1.0, –0.5, 0.0

Grasses effects: γ_i (i=1…3) = –2, –1, 3

Shrubs effects: ψ_j (j) = –1., –.5, 1., 0.5, 0.0

Interactions δ_{ij}	Shrubs				
Grasses	S_1	S_2	S_3	S_4	S_5
G_1	0.2	–0.4	–0.2	0.0	0.4
G_2	–0.3	0.2	0.4	0.1	–0.4
G_3	0.1	0.2	–0.2	-0.1	0.0

Genstat codes for generating a randomized plan and data analysis (Design 4) are given in Appendix. The statistical analysis was repeated 100 randomly generated data sets. Table 2 gives average, over the simulation runs, of each of the effects and interactions parameters set in Table 1. It may be observed that, the displayed averages of gesg (ψ_j s general effects of shrubs on the grasses) and δ_{ij} is the interaction (or gs-gseg, the grass-specific general effect of shrub S_j on the grass G_i) are very close to the values of the respective parameters. GenStat codes for construction and analysis of data using the other experimental designs are given in Singh (2017).

Table 2: Mean of 100 simulations of estimates of shrub effects and interaction with grasses

A. Shrub Effects		
Shrub S_j	True value (ψ_j)	Average of 100 simulations
S_1	–1.0	–0.997
S_2	–0.5	–0.518
S_3	1.0	1.068
S_4	0.5	0.478
S_5	0.0	–0.031
SE		±0.325

B. Shrub × Grass Interaction			
Grass i	Shrub S_j	True value (δ_{ij})	Average of 100 simulations
1	S_1	0.2	0.230
	S_2	–0.4	–0.447
	S_3	–0.2	–0.200
	S_4	0	–0.018
	S_5	0.4	0.435
2	S_1	–0.3	–0.344
	S_2	0.2	0.248
	S_3	0.4	0.372
	S_4	0.1	0.122
	S_5	–0.4	–0.399
3	S_1	0.1	0.114
	S_2	0.2	0.199
	S_3	–0.2	–0.172
	S_4	–0.1	–0.104
	S_5	0.0	–0.037
SE			±0.455

SE = Estimated standard error

6. Conclusions

Alley cropping with shrubs as borders or hedges and crops/grasses in the alleys are often agroforestry practices for sustainable crop production. The experimental designs and statistical models for data analyses are discussed for commonly occurring situations. These designs are recommended for conducting alley cropping trials. Once the real data become available, the steps presented here may be used for analysis. These designs and the approach of analysis can also be adapted for examining interactions or interference in intercropping experiments, which would need further extension to analyze two or more correlated responses on the component crops.

Acknowledgements

I acknowledge Dr. Mounir Louhaichi and Ms. Sawsan Hassan, Range Ecology and Management Research Team, SIRPS, ICARDA, Amman for helpful discussions on alley cropping experiments and motivation to complete this work. I would also like to tank the anonymous reviewer and the editor for teir helpful suggestions.

References

AFNTA [Alley Farming Network for Tropical Africa] (1992a). *Alley Farming Training Manual, Vol 1: Core Course in Alley Farming*. Editors: B.R. Tripathi and Paul J. Psychas. International Institute of Tropical Agriculture, Ibadan. 196 pp.

AFNTA [Alley Farming Network for Tropical Africa] (1992b). *Alley Farming Training Manual, Vol 2: Source Book for Alley Farming Research*. Editors: B.R. Tripathi and Paul J. Psychas. International Institute of Tropical Agriculture, Ibadan. 226 pp.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control,* **19**(**6**), 716-723, doi:10.1109/TAC.1974.1100705.

Arya, A. S. (1983). Circulant plans for partial diallel crosses. *Biometrics*, **39**, 43-52.

Curnow, R. N. (1963). Sampling the diallel cross. *Biometrics*, **19**, 287-306.

Kempthorne, O. and Curnow, R. N. (1961). The partial diallel crosses. *Biometrics*, **17**, 229-250.

Rao, C. R. (1973). *Linear Statistical Inference and its Applications*. 2nd Edn. John Wiley & Sons, New York.

Singh, M. (2017). Experimental designs for alley cropping to estimate shrub × grass interaction. *Biometrics and Statistics Technical Report* No. 1(2017). BSS/DDG-R, ICARDA, Amman. 48pp. https://repo.mel.cgiar.org/handle/20.500.11766/5672 (Accessed 19 April 2020).

Singh, M., Gupta, S. and Parsad, R. (2012). Genetic crosses experiments. Pages 1-71. In *Design and Analysis of Experiments*, Volume 3: *Special Designs and Applications*, John Wiley & Sons (Editor: K. Hinkelmann).

Solaimalai, A., Muralidaran C. and Subburamu, K. (2005). Alley cropping in rainfed agroecosystem - a review. *Agricultural Reviews,* **26**(**3**), 157-172.

Appendix

a. GenStat Codes for generation of Design 4

Let NGrass, NShrub and NRep be the number of grasses, shrubs and replications respectively. The GenStat codes:

```
"Generate Design 4..... Diallel-borders in main-plots of a split plot design"
Scal NRep, NGrass, NBorder, NPlots; 3, 3, 5,*
Calc NPlots=NRep*NGrass*NBorder

Unit[NPlots]
Factor[Levels=NRep] Rep  : &[Levels=NGrass] Grass : &[Levels=NBorder; labels=!t('S5S4',
'S4S1', 'S1S2','S2S3', 'S3S5')]Border
Generate Rep,  Border, Grass

Randomize[Block=Rep/Border/Grass; Seed=130572] Border, Grass
Prin Rep,  Border, Grass
```

b. GenStat Codes for analysis of data from Design 4

Let Yield be the vector of plot yields. The codes for statistical analysis of variance and estimation of means, effects and interaction are:

```
"..... Analysis part....."
"Genstat codes for estimating the effects and interactions using ANOVA and linear model fitting directives
..... Analysis part....."
"Stage 1: Estimate borders [shrub combinations] and grasses effect and their interactions"
```

```
Dele[rede=y] Borders0, LeftBorder0, RightBorder0,Grass0,Sh[1...NShr], yMeans, Weight, tbEff,
yEff,tbMn, tbSeMn
            Block Rep/Border/Grass
            Treat Border*Grass
            Anova[print=a,%cv, eff, mean;pse=m;fpro=y]Yield
"Get the Grasses effects from above or below: Gammas"
            Akeep  Grass; means=tbMn ; SEmeans=tbSe; Effect=tbEff
Prin tbMn, tbEff, tbSe

 "Stage 2:.....To get direct effect of Shrubs (saai s) on grasses"

            Akeep  Border; means=tbMn ; SEmeans=tbSe; Effect=tbEff
" Get error mean squre, weight of means "
Vtable Table=tbMn, tbSe, tbEff; Vari=yMeans, ySeMn, yEff; Class=!P(Borders0)
Vari [nval=NShr]Sh[1...NShr], Weight
Calc Weight=1/ySeMn**2
"Decode Borders into left and right border shrubs: S5S4, S4S1, S1S2,S2S3, S3S5"
Vari[Values=5,4,1,2,3]LeftB
Vari[Values=4,1,2,3,5]RightB
For i=1...NShr;dd=Sh[1...NShr]
  Calc dd=(LeftB.eq.i.or.RightB.eq.i)
Endf
"Print Borders0, LeftB, RightB,Sh[1...NShr], yMeans, ySeMn,yEff, Weight; field=6"

" Regression Model/Fit to estimate Shrub direct effects: saai s "
Print '   ***** Saai s and their standard errors for shrubs *****'
Model[Weight=Weight; disp=1] yEff
Terms [Full=Y] Sh[1...NShr]
Fit[Prin=m,s,e; cons=o; fpro=y; tpro=y] Sh[1...NShr]

"Estimate Shrub X Grass interaction delta s"
Dele[rede=y] GrassBorders0, GrassLeftBorder0, GrassRightBorder0,Grass0,Sh[1...NShr], yMeans,
Weight, tbMn, tbSe,tbEff, yEff
            Akeep  Grass.Border; means=tbMn ; SEmeans=tbSe; Effects=tbEff
" Get error mean squre, weight of means "
Vtable Table=tbMn, tbSe, tbEff; Vari=yMeans, ySeMn, yEff; Class=!P(Grass0,Borders0)

Scal NGrassXNShrub : Calc NGrassXNShrub=NShr*NGrass
Scal NGrassXNBorder : Calc NGrassXNBorder=NBorder*NGrass
Vari [nval=NGrassXNBorder]Sh[1...NShr], Grs[1...NGrass], Weight
Calc Weight=1/ySeMn**2
" Decode Borders into left and right border shrubs: S5S4, S4S1, S1S2,S2S3, S3S5 for each grass
Borders0  Boders and grasses:
  G1          G1          G1 G1      G1          / G2          G2          G2          G2  G2      / G3
              G3  G3      G3          G3
 S5S4, S4S1, S1S2,S2S3, S3S5   S5S4, S4S1, S1S2,S2S3, S3S5   S5S4, S4S1, S1S2,S2S3, S3S5
"
Vari[Values=(5,4,1,2,3)3]LeftBG
Vari[Values=(4,1,2,3,5)3]RightBG
For i=1...NShr;dd=Sh[1...NShr]
  Calc dd=(LeftBG.eq.i.or.RightBG.eq.i)
Endf
For i=1...NGrass; dd=Grs[1...NGrass]
  Calc dd=(Grass0==i)
```

```
Endf
Print Grass0, Borders0, LeftBG, RightBG,Sh[1...NShr], Grs[1...NGrass],yMeans, ySeMn,yEff,
Weight ; field=6
" Shrubs x Grass interaction: deltas and SE for each grass"
For i=1...NGrass
Print '   ***** Deltas and their standard errors for Grass = ', i, '  *****'
Rest Sh[1...NShr], yEff ; Grass0==i
Model[Weight=Weight; disp=1] yEff
Terms [Full=Y] Sh[1...NShr]
Fit[Prin=m,s,e; cons=o; fpro=y; tpro=y] Sh[1...NShr]
Rest Sh[1...NShr], yEff
Endf
 STOP
```

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 401–415

Second Order Asymptotics of a Fine-Tuned Purely Sequential Procedure for the Generalized Partition Procedure

Tumulesh K. S. Solanky
Department of Mathematics
University of New Orleans, New Orleans, Louisiana, USA

Received: 24 January 2021; Revised: 01 March 2021; Accepted: 17 March 2021

Abstract

In the area of selection and ranking, partitioning of treatments by comparing them to a control treatment is an important statistical problem. For over seventy years this problem has been investigated by a number of researchers via various statistical designs to specify the partitioning criteria and optimal strategies for data collection. Tong (1969) had proposed a design which had generalized many formulations known at that time. Relying upon Bechhofer's (1954) indifference-zone formulation, Tong (1969) had designated the region between the boundaries for "good" and "bad" treatments as the indifference-zone. Since then the formulation presented in Tong (1969) has been adopted by a number of researchers to study various aspects of the partition problem. However, in "Tong's formulation" the dual role the indifference zone plays, in defining the preference-zone and boundaries for "good" and "bad" treatments, could potentially make the formulation impractical when there is a large difference between the "good" and "bad" treatments. In Solanky and Jhou (2015), a generalization of the "Tongs formulation" was introduced so that the treatments which fall between the "good" and "bad" treatments can be partitioned as a separate identifiable group by introducing two indifference-zones. It was also shown that the formulation design in Tong (1969) is a special case of Solanky and Jhou's (2015) formulation. However, the second-order expansions of the probability of correct partition given in Solanky and Zhou (2015) does not make it clear how close one really gets to the target probability requirement. In this paper, we have proposed a fine-tuned purely sequential procedure which is asymptotically unbiased and guarantees the probability requirement by taking a few additional samples along the lines of Mukhopadhyay and Datta (1995). The "first-order" and "second order" asymptotics of the fine-tuned procedure are derived and it is shown that the second-order expansion of the stopping time has the same order of the remainder as that for the original procedure in Solanky and Zhou (2015). The performance of the proposed fine-tuned procedure is studied via Monte Carlo simulations.

Key words: Fine-tuned purely sequential procedure; Control population; Indifference zone; Probability of correct decision; Normal distribution; Simulations.

AMS Subject Classifications: 62F07, 62L10

Corresponding Author: Tumulesh K. S. Solanky
Email: tsolanky@uno.edu

1. Introduction

The problem of comparisons with a control has intrigued the researchers for the last seventy plus years. It has been investigated under different designs and sampling methodologies. Among the oldest investigations available in the statistical literature, Roessler (1946), Paulson (1952), and Dunnett (1955) are three studies which had introduced the need for statistical methodologies to compare treatments with a control treatment and had formulated designs for such a comparison.

For a somewhat related problem of selecting or isolating the best population, two pioneering papers in 1950s had presented contrasting formulations. The first formulation was in Bechhofer (1954) which introduced the idea of indifference-zone for selecting the best normal population from a group of several normally distributed populations. This formulation, known as "indifference-zone" formulation, had the property of selecting the best population with the pre-determined probability specified by the practitioner for the populations that are in the preference-zone. The region outside of the indifference-zone is referred to as the *preference-zone*. From a practical point of view, the indifference zone was envisioned to be small enough so that the experimenter could be easily be *indifferent* to the treatments that fall inside this region. The other pioneering formulation, introduced in Gupta (1956), did not restrict the selection from the preference-zone but rather the selection was carried out from the entire parameter space. The formulation in Gupta (1956), known as "subset-selection formulation", is designed to select a subset of random size which includes the best treatment with the pre-determined probability of correct selection. Since then, the literature in this area has grown enormously extending Bechhofer's (1954) indifference-zone formulation and Gupta's subset-selection formulation, to solve problems related to selecting the best treatment and by defining the "best-treatment" in many ways to meet the goals of the study. This area of research is broadly known as the area of *selection and ranking* in the statistical literature. Along side with the research in the *selection and ranking* to select the best population, another research problem has also been quite active which is concerned with comparisons of treatments with a specific treatment of choice. What made this research area different from selecting the best treatment is the experimenters requirement that the the population to be selected must be some "specified amount better" than some other treatment typically referred to as a *control* or *standard*. This area of research is typically known as the problem of "comparisons with a control" or the "partition problem" in statistical literature.

1.1. Tong's formulation of partition problem

We will start by presenting the partition problem formulation introduced in Tong (1969) for the populations that follow a normal distribution. The formulation starts by specifying the "good" and "bad" populations based on the input from experts in the area of the application. The region that falls in between these two boundaries is next defined in Tong (1969) as the "indifference zone".

The concept of indifference-zone was introduced in Bechhofer (1954) for selecting the best normal population in order to create a spacing between the best and the rest of the treatments under consideration with the underlying requirement that the experimenter would be "indifferent" to the treatments falling inside the indifference-zone. In other words, any population that is inside the indifference-zone cannot be miss-classified.

This dual role the indifference zone, in specifying the "good" and "bad" populations boundaries and also the "indifference zone" itself, could potentially make the formulation impractical in cases when there is a large difference between "good" and "bad" populations. We will revisit this issue a bit later in this paper.

Let $\pi_0, \pi_1, \cdots, \pi_k$ denote the $(k+1)$ independent and normally distributed populations with respective means $\mu_i,\ i = 0,\ 1, \cdots, k$, and common variance σ^2. Assume that all the parameters $\mu_i,\ i = 0,\ 1, \cdots, k$, and variance σ^2 are unknown. We will denote the population π_0 as the control population with which remaining k populations will be compared.

Based on the guidelines of "good" and "bad" treatments from the practitioner in the area, Tong (1969) defined two appropriate constants δ_1 and δ_2, with the requirement that $\delta_1 < \delta_2$, to split the parameter space Ω into three sets following Bechhofer's (1954) indifference-zone formulation, as follows

$$\begin{array}{rcl} \Omega_{BT} & = & \{\pi_i : \mu_i \leq \mu_0 + \delta_1,\ i = 1, \cdots, k\}, \\ \Omega_{IT} & = & \{\pi_i : \mu_0 + \delta_1 < \mu_i < \mu_0 + \delta_2,\ i = 1, \cdots, k\}, \\ \Omega_{GT} & = & \{\pi_i : \mu_i \geq \mu_0 + \delta_2,\ i = 1, \cdots, k\}. \end{array} \tag{1}$$

Without loss of generality, the set Ω_{GT} will be used to denote "good" populations and the set Ω_{BT} the "bad" populations. From the applications point of view, the values of constants δ_1 and δ_2 are determined based on input from the experts to specify which population should be classified as "good" population compared to control population and which as a "bad" population compared to control population. The partition problem formulated in Tong (1969) was designed to correctly partition the populations which belonged to only Ω_{BT} and Ω_{GT}. Whereas, since the set expressed as Ω_{IT} was the "indifference-zone" set, the mathematical formulation was indifferent to the populations which are in this set. In other words, the populations that are in Ω_{IT} could be partitioned in Ω_{BT} or Ω_{GT} without any penalty. The decision rule proposed in Tong (1969) was to partition or classify the set Ω of the k populations into non-overlapping subsets S_{BT} and S_{GT} of Ω, such that, $\Omega_{GT} \subseteq S_{GT}$ and $\Omega_{BT} \subseteq S_{BT}$. A decision rule which classifies all populations in Ω_{BT} or Ω_{GT} correctly was termed as a "correct decision (CD)". The design constants used in the partition rule used in Tong (1969) are defined below

$$\begin{array}{l} d_1 = \frac{\delta_1+\delta_2}{2}, \quad a_1 = \frac{\delta_2-\delta_1}{2}, \quad \lambda = \frac{\sigma}{a_1}, \\ m_1 = \begin{cases} \frac{k}{2} & \text{for even values of } k, \\ \frac{k+1}{2} & \text{for odd values of } k. \end{cases} \end{array} \tag{2}$$

Mathematically, for any pre-specified probability P^*, where $\frac{1}{2^k} < P^* < 1$, Tong (1969) presented a decision rule $\wp_T$ to obtain sets S_{BT} and S_{GT}, which satisfy the probability requirement

$$P\{CD|\boldsymbol{\mu}, \sigma^2, \wp\} \geq P^* \qquad \forall\ \boldsymbol{\mu} \in \boldsymbol{R}^{k+1},\ \sigma \in \boldsymbol{R}^+, \tag{3}$$

where $\boldsymbol{\mu} = (\mu_0, \mu_1, \cdots, \mu_k)$.

Adopting a sampling design to determine the sample size of N observations from all the

k+1 populations and control, the decision rule to partition the k populations in Ω presented in Tong (1969) was the following:

$$\begin{array}{rcl} S_{BT} & = & \{\pi_i : \bar{X}_{iN} - \bar{X}_{0N} \leq \ d_1, \ i = 1, \cdots, k\}, \\ S_{GT} & = & \{\pi_i : \bar{X}_{iN} - \bar{X}_{0N} \geq \ d_1, \ i = 1, \cdots, k\}, \end{array} \tag{4}$$

where d_1 comes from (2).

For "partition problem" as outlined above, Tong (1969) derived a fixed sample size solution for the case when σ^2 is known by determining the optimal sample size needed to meet the probability condition specified in (3). Tong (1969) also considered the case when the common variance σ^2 unknown by designing a two-stage stopping rule and a purely-sequential in order to meet the probability requirement (3). Tong's (1969) formulation was utilized by Datta and Mukhopadhyay (1998) to construct a fine-tuned purely sequential procedure, an accelerated sequential procedure and a three-stage procedure focusing on the second-order asymptotics for each procedure. Also following Tong's (1969) formulation, with the additional goal of reducing the sampling from populations which can be partitioned based on smaller sample sizes due to being significantly better or worse than the control population, Solanky (2001) constructed an elimination type purely-sequential stopping rule. Using triangular boundaries, Solanky (2001) was designed to reduce the sampling cost by not following "vector-at-a-time" sampling methodology and instead it presented a sampling design that eliminated and partitioned the non-contending populations early during the sampling process. The operational inconvenience of purely-sequential stopping rules was also the focus of Solanky (2006) which designed a two-stage sampling rule which had the desirable property of eliminating "too inferior" or "too superior" populations based on samples collecting during the first stage by tactfully implementing Gupta's subset selection for screening of such populations. In the stage two of sampling in Solanky (2006), only the competing treatments which were shortlisted in stage one were sampled from and partition was implemented using Bechhofer's indifference zone approach. Also relying on the Tong's (1969) formulation, Solanky and Wu (2004) had constructed an "unbalanced sampling design" which allows an experimenter to collect a bigger sample size from the control population while reducing the sample size from the non-control populations from which "vector-at-a-time" sampling methodology is adopted. For references on partition problem for Binomial treatments the reader is referred to Buzaianu (2019).

1.2. Generalization of the partition problem

The formulation presented above due to Tong (1969) was constructed under the Bechhofer's(1954) "indifference-zone formulation" to partition the k populations as either a 'Good" treatments or a "Bad" treatments with respect to the control population. The methodology of Tong (1969) was formulated to partition the populations which may fall under the indifference-zone as either "Good" populations or "Bad" populations without any penalty on the probability requirement specified in (3). This requirement of Tong's (1969) formulation would be intuitively serve the experimenter well as long as the distance between the "Good" populations and "Bad" populations is is not large so that one could be "indifferent" to how the populations which fall inside it are partitioned. However, when there is large gap between the "Good" populations and "Bad" populations boundaries then this could result in misleading results and would also violate the concept of minimum distance

worth measuring as represented by the indifference-zone. One is referred to Solanky and Zhou (2015) for illustrations and further insights.

In Solanky and Jhou (2015), the partition problem was formulated in such a fashion that does not require the region that falls between the "Good" populations and "Bad" populations boundaries to be designated as indifference zone. Instead, by utilizing the creation of two indifference-zones, the generalized formation was able to partition the treatments which fall between the "Good" populations and "Bad" populations as a group by itself. The two indifference-zones thus created were based on the fundamentals of "indifference-zone" as intended in Bechhofer (1954) and they could be as small as the experimenter desires and also met the criteria to be minimum distance worth measuring. Put differently, the Solanky and Jhou (2015) generalized formulation had the capability of creating indifference-zones independently of the boundaries of "Good" populations and "Bad" populations. Next, we represent the mathematical details of the generalized formulation of Solanky and Jhou (2015).

Based on the input from experts in the area, the statistical design would start by selecting two design constants δ_1 and δ_4, $\delta_1 < \delta_4$, to define the "Good" populations and "Bad" populations compared to the control population. Next, based on experts understanding of how much distance is worth detecting or the "threshold", one would quantify that information by the design constants δ_2 and δ_3 so that $\delta_2 - \delta_1 (> 0)$ and $\delta_4 - \delta_3 (> 0)$ are the spacing's which will be used to construct two 'indifference-zones". Next, one would define the following sets to split the entire parameter space Ω along the lines of Bechhofer's (1954) "indifference-zone formulation"

$$\begin{array}{rcl} \Omega_B & = & \{\pi_i : \mu_i \leq \mu_0 + \delta_1,\ i = 1, \cdots, k\}, \\ \Omega_{I_1} & = & \{\pi_i : \mu_0 + \delta_1 < \mu_i \leq \mu_0 + \delta_2,\ i = 1, \cdots, k\}, \\ \Omega_M & = & \{\pi_i : \mu_0 + \delta_2 < \mu_i \leq \mu_0 + \delta_3,\ i = 1, \cdots, k\}, \\ \Omega_{I_2} & = & \{\pi_i : \mu_0 + \delta_3 < \mu_i \leq \mu_0 + \delta_4,\ i = 1, \cdots, k\}, \\ \Omega_G & = & \{\pi_i : \mu_i > \mu_0 + \delta_4,\ i = 1, \cdots, k\}. \end{array} \tag{5}$$

One may note that the sets Ω_{I_1} and Ω_{I_2} will serve as the two "indifference-zones". It is also apparent that the sizes of these two "indifference-zones" does not depend on the what a practitioner may have picked to specify the "Good populations" and "Bad populations" compared to the control population. Next, following Solanky and Jhou (2015), we define the constants

$$d_1 = (\delta_1 + \delta_2)/2, \qquad d_2 = (\delta_3 + \delta_4)/2, \qquad a_1 = (\delta_2 - \delta_1)/2, \qquad a_2 = (\delta_4 - \delta_3)/2. \tag{6}$$

We write X_{ij} to denote a random sample of size n from the population π_i where $j = 1, \cdots, n$ and $i = 0, 1, \cdots, k$. Writing $\bar{X}_i$ and S_i^2 for the usual sample mean and sample variance from the ith population π_i as $\bar{X}_i = \frac{\sum_{j=1}^{n} X_{ij}}{n}$, and $S_i^2 = \frac{\sum_{j=1}^{n} \left(X_{ij} - \bar{X}_i\right)^2}{n-1}$, we obtain the pooled estimator of the common variance σ^2 to be denoted as $U(n) = \frac{\sum_{i=0}^{k} S_i^2}{k+1}$. Next, as in Solanky and Jhou (2015), we propose a intuitively defined decision rule $\mathcal{P}_n$ based on the vector-at-a-time sampling strategy of size n from all the populations:

$$\begin{aligned} S_B &= \{\pi_i : \bar{X}_{in} - \bar{X}_{0n} \leq d_1, i = 1, \cdots, k\}, \\ S_M &= \{\pi_i : d_1 \leq \bar{X}_{in} - \bar{X}_{0n} \leq d_2, i = 1, \cdots, k\}, \\ S_G &= \{\pi_i : \bar{X}_{in} - \bar{X}_{0n} \geq d_2, i = 1, \cdots, k\}, \end{aligned} \tag{7}$$

where S_B is the set of populations that are classified as"Bad populations", S_M as "Medium populations", and S_G as"Good populations". As explained earlier, above decision rule does not change the definition of "Good" populations and "Bad" populations compared to the control population which are based on input from experts in the area.

Also, as it is customary for the selection and ranking problems, the decision rules and the probabilities of correct decision are generally derived under the worst possible parametric configurations, which are referred to as the "least favorable configurations (LFC)". And, when the parametric configurations become favorable, then the observed probabilities of correct decision would exceed the target probability values. Using symmetry argument, it is apparent that if Ω_{I_1} and Ω_{I_2} not equal in width then the LFC would become a function of constants δ_2 and δ_3, and hence LFC would cease to exist in a general form. Also, intuitively because there is no penalty for misidentifying the populations that fall inside the two "indifference-zones" Ω_{I_1} and Ω_{I_2}, under the LFC should not have any population that falls inside any of them. Utilizing the above information, in Solanky and Zhou (2015) the LFC was introduced taking advantage of the symmetry in the decision rule (7). We rewrite the spacing's using $a = \delta_4 - \delta_1$, the distance between the "Good" populations and "Bad" populations, and a constant ρ, $0 < \rho < \frac{1}{2}$, which expresses the size of the two "indifference-zones" as a function of a.

(1) $\delta_4 - \delta_3 = \delta_2 - \delta_1 = \rho a$, $0 < \rho < \frac{1}{2}$. Note that ρ is a design constant which is used to define the size of the indifference-zones;

(2) $r_2 + r_3 = \left[\frac{k}{2}\right] = k_1$, $r_1 + r_4 = k - k_1$, $r_2 = \left[\frac{k_1}{2}\right]$, $r_3 = k_1 - r_2$, $r_1 = \left[\frac{k-k_1}{2}\right]$, $r_4 = k - k_1 - r_1$, where r_1, r_2, r_3, and r_4 denotes the number of populations with the respective means: $\mu_0 + \delta_1, \mu_0 + \delta_2, \mu_0 + \delta_3$, and $\mu_0 + \delta_4$, where $[x]$ equals $\frac{x}{2}$ if x is even and $\frac{x+1}{2}$ if x is odd. We denote this parametric configuration as $\boldsymbol{\mu}^0(r_1, r_2, r_3, r_4)$.

As explained in Solanky and Zhou (2015), the condition (1) above forces the two indifference-zone's Ω_{I_1} and Ω_{I_2} to be of same size, whereas, the condition (2) adds symmetry to the problem and essentially allows equal number of populations at the four boundaries $\mu_0 + \delta_1, \mu_0 + \delta_2, \mu_0 + \delta_3$, and $\mu_0 + \delta_4$. Solanky and Zhou (2015) had shown that LFC is when $r_1 = r_2 = r_3 = r_4$. It is easy to note that as ρ becomes large and approaches $\frac{1}{2}$, the size of the Ω_{I_1} and Ω_{I_2} increases and the size of "Medium populations" which are classified as a separate group under generalized partition methodology becomes smaller. And, in this case the generalized partition rule would approach the formulation presented in Tong (1969). The constant ρ determines the "threshold" or "minimum distance worth detecting" along the lines of Bechhofer (1954). Whereas, in the formulation of Tong (1969) the "indifference zone" played the dual role of defining the "Good" and "Bad" populations and the role of the "threshold" or the "minimum distance worth detecting". As derived in Solanky and Jhou (2015), for the known common variance σ^2 case, the probability for the correct decision for (7) can be shown to be

$$P\Big[CD|\boldsymbol{\mu}^0(r_1,r_2,r_3,r_4),\sigma^2\Big] \geq P^* \tag{8}$$

if the sample size from the k populations and the control population at least $\frac{8b^2\sigma^2}{(\rho a)^2}$ $(=C)$ and the design constant $b = b(k, P^*)$ is obtained by solving:

$$\int_{-b}^{b}\int_{-b}^{b}...\int_{-b}^{b}(2\pi)^{-\frac{k_1}{2}}|\Sigma|^{-\frac{1}{2}}\exp(-\frac{1}{2}y'\Sigma^{-1}y)dy_1...dy_{\frac{k_1}{2}} = \frac{P^*+1}{2}. \tag{9}$$

In the statistical literature, $C = \frac{8b^2\sigma^2}{(\rho a)^2}$, is referred to as the fixed-sample solution to the partition problem. The values of design constant $b = b(k, P^*)$ are available in Solanky and Jhou (2015).

In section 2, we have proposed a fine-tuned purely sequential procedure that is asymptotically unbiased and guarantees the probability requirement by taking a few additional samples along the lines of Mukhopadhyay and Datta (1995) and Woodroofe (1991). The "first-order" and "second order" asymptotics of the proposed procedure are obtained and it is shown that the second-order expansion of the stopping time has the same order of the remainder as that for the original procedure in Solanky and Zhou (2015). The probability of the correct decision for the fine-tuned version of the purely-sequential procedure for the generalized partition problem will be shown to be at least P^* up to the second-order approximation and will not have additional terms.

2. Fine-Tuned Purely Sequential Procedure

Next, we propose a "fine-tuned purely sequential procedure" for the generalized partition problem described in (5). The procedure starts with initial sample size of $m_0\,(\geq 2)$ observations. Next, adopting the "vector-at-a-time" sampling design and it takes one observation at a time updating the statistic $U(n)$ after each observation until the first time the condition in the condition below is satisfied

$$N \equiv N\,(a) = \text{Inf}\left\{n \geq m_0 : n + \epsilon\ \geq \frac{8b^2U(n)}{(\rho a)^2}\right\}, \tag{10}$$

where the constant $\epsilon = \epsilon\,(k, P^*)$ is defined in (21) and $b = b(k, P^*)$ in (9). We first verify that the stopping rule (10) will terminate with probability one. Note that for given values of constants $\boldsymbol{\mu}$, σ^2, m_0, ρ and a, we can write

$$\begin{aligned} P\,(N<\infty) &= 1-\lim_{n\to\infty}P\,(N>n) \\ &\geqslant 1-\lim_{n\to\infty}P\left\{n+\epsilon<\frac{8b^2U(n)}{(\rho a)^2}\right\} \\ &= 1,\ \text{since } U(n)\xrightarrow{P}\sigma^2 \text{ as } n\to\infty. \end{aligned}$$

Hence, with probability one the proposed "fine-tuned purely sequential procedure" (10) will terminate. Based on the totality of all samples of size N, one computes the sample means $\bar{X}_{0N}, \bar{X}_{1N}, ..., \bar{X}_{kN}$ and implements the decision rule (7). Next, we will derive some theoretical

properties to establish the performance of the "fine-tuned purely sequential procedure" (10) asymptotically. It is easy to verify that as a becomes small, the sample size as determined by the stopping rule (10) becomes larger. We start with presenting first-order asymptotics "fine-tuned purely sequential procedure" (10).

Theorem 1: For N as defined by the"fine-tuned purely sequential procedure" (10), the following properties are satisfied :

(i) $\frac{N}{C} \xrightarrow{P} 1$ as $a \to 0$:

(ii) $E(\frac{N}{C}) \to 1$ as $a \to 0$:

(iii) $\frac{N-C}{C^{\frac{1}{2}}} \xrightarrow{L} N(0, \frac{2}{k+1})$ as $a \to 0$:

(iv) $\liminf_{a\to 0} P\left(CD|\mathcal{P}_N\right) > P^*$ under the LFC;

where $C = \frac{8b^2\sigma^2}{(\rho a)^2}$, $b = b(k, P^*)$ is defined in (9) and has been tabulated in Solanky and Zhou (2015), and the constant $\epsilon = \epsilon(k, P^*)$ is defined in (21).

Proof: We start the proof by noting that the N is well-defined and is a non-decreasing function of a. Next, along the lines of Theorem 2.4.1 of Mukhopdahyay and Solanky (1994) we use the Lemma 1 of Chow and Robbins (1965) to note $N \xrightarrow{P} \infty$ as $a \to 0$. Also, $U(N) \xrightarrow{P} \sigma^2$ and $U(N-1) \xrightarrow{P} \sigma^2$ as $a \to 0$. Next, inequality (2.4.3) from Mukhopadhyay and Solanky (1994) for the stopping rule (10) simplifies to

$$\frac{8b^2U(N)}{(\rho a)^2} - \epsilon \leq N \leq m_0 - \epsilon + \frac{8b^2U(N-1)^2}{(\rho a)^2}. \tag{11}$$

We can rewrite (11) as

$$\frac{8b^2U(N)}{(\rho a)^2C} - \frac{\epsilon}{C} \leq \frac{N}{C} \leq \frac{m_0 - \epsilon}{C} + \frac{8b^2U(N-1)^2}{(\rho a)^2C},$$

which simplifies to

$$\frac{U(N)}{\sigma^2} - (\frac{\epsilon\rho^2}{8b^2\sigma^2})a^2 \leq \frac{N}{C} \leq (\frac{(m_0-\epsilon)\rho^2}{8b^2\sigma^2})a^2 + \frac{U(N-1)}{\sigma^2}.$$

Above with $U(N) \xrightarrow{P} \sigma^2$ and $U(N-1) \xrightarrow{P} \sigma^2$ as $a \to 0$ completes the part (i) of the Theorem. Next, as in Mukhopadhyay and Solanky's (1994) Theorem 3.5.1, we rewrite the estimator $U(n)$ by using the "Helmert's orthogonal transformation" to obtain R_1', R_2', $\cdots$ which are i.i.d. $(k+1)^{-1}\sigma^2\chi^2_{k+1}$ to express $U(n) = \frac{1}{n-1}\sum_{i=1}^{n-1} R_i'$. Considering only one side of (11), and writing $R_* = \sup_{n\geq 2}\left\{\frac{1}{n-1}\sum\limits_{i=1}^{n-1} R_i'\right\}$, we have

$$N \leq m_0 - \epsilon + \frac{8b^2R_*}{(\rho a)^2}.$$

This can be rewritten as

$$\frac{N}{C} \leq \frac{m_0 - \epsilon}{C} + \frac{R_*}{\sigma^2}, \tag{12}$$

and when a is small enough so that $\frac{1}{C}$ become less than 1, the inequality (12) simplifies to $\frac{N}{C} \leq m_0 - \epsilon + \frac{R_*}{\sigma^2}$. By "Wiener's (1939) dominated ergodic theorem" we have $E(R_*) < \infty$. By applying the "dominated convergence theorem" and using Theorem's (i) result, have the proof for part (ii). The proof for part (iii) of the Theorem can be derived by comparing the "fine-tuned purely sequential procedure" (10) with (3.5.1) of Mukhopadhyay and Solanky (1994). The details are omitted for brevity.

Next, to prove the part (iv), note that without loss of generality we will assume that $\pi_i, i = 1, ..., r_1$ have means $\mu_0 + \delta_1$; $\pi_j, j = r_1 + 1, ..., r_1 + r_2$ have means $\mu_0 + \delta_2$; $\pi_l, l = r_1 + r_2 + 1, ..., r_1 + r_2 + r_3$ have means $\mu_0 + \delta_3$; and $\pi_m, m = r_1 + r_2 + r_3 + 1, ..., k$ have means $\mu_0 + \delta_4$. Next, the probability of correct decision can be written as

$$\begin{aligned} &P\Big(CD|\mathcal{P}_N; \boldsymbol{\mu}^0(r_1, r_2, r_3, r_4),\ \ \sigma^2\Big) \\ &\quad = P\Big[\bar{X}_i - \bar{X}_0 < d_1,\ d_1 < \bar{X}_j - \bar{X}_0 < d_2,\ d_1 < \bar{X}_l - \bar{X}_0 < d_2,\ \bar{X}_m - \bar{X}_0 > d_2, \\ &\quad 0 < i \leq r_1,\ \ r_1 < j \leq r_1 + r_2,\ \ r_1 + r_2 < l \leq r_1 + r_2 + r_3,\ \ r_1 + r_2 + r_3 < m \leq k\Big]. \end{aligned}$$

As in Solanky and Zhou (2015), the P(CD) can be tactfully written as probability expressions for the correct partition of the populations which are "Good" or "Bad populations" and probability expressions for the correct partition of the populations which are "Medium populations". Assuming that the practitioner considers the correct partition of these two groups of populations as equal, then the P(CD) expression would simplify to:

$$\begin{aligned} P(CD|\mathcal{P}_N) \ \ \geq \ \ (-1) + 2P\Big\{&d_1 < \bar{X}_j - \bar{X}_0 < d_2, r_1 + 1 \leq j \leq r_1 + r_2; \\ &d_1 < \bar{X}_l - \bar{X}_0 < d_2, r_1 + r_2 + 1 \leq l \leq r_1 + r_2 + r_3\Big\}. \end{aligned}$$

Above can be simplified as

$$\begin{aligned} P(CD|\mathcal{P}_N) \ \ \geq \ \ (-1) + 2P\Bigg\{&\frac{d_1 - \delta_2}{\sqrt{\frac{\sigma^2}{N}}} < \frac{\left(\bar{X}_j - \mu_j\right)}{\sqrt{\frac{\sigma^2}{N}}} - \frac{\left(\bar{X}_0 - \mu_0\right)}{\sqrt{\frac{\sigma^2}{N}}} < \frac{d_2 - \delta_2}{\sqrt{\frac{\sigma^2}{N}}}; \\ &r_1 + 1 \leq j \leq r_1 + r_2; \quad \frac{\rho a}{\sqrt{\frac{\sigma^2}{N}}} < \frac{\left(\bar{X}_l - \mu_l\right)}{\sqrt{\frac{\sigma^2}{N}}} \\ &-\frac{\left(\bar{X}_0 - \mu_0\right)}{\sqrt{\frac{\sigma^2}{N}}} < \frac{d_2 - \delta_3}{\sqrt{\frac{\sigma^2}{N}}},\ \ r_1 + r_2 + 1 \leq l \leq r_1 + r_2 + r_3\Bigg\}. \end{aligned}$$

Denoting, $b = \frac{\rho a/2}{\sqrt{\frac{\sigma^2}{N}}}$, and $c = \frac{(2a - 3\rho a)/2}{\sqrt{\frac{\sigma^2}{N}}}$, then the above can be rewritten as

$$\begin{aligned} P(CD|\mathcal{P}_N) \ \ \geq \ \ (-1) + 2P\{&-b < Z_j - Z_0 < c, r_1 + 1 \leq j \leq r_1 + r_2; \\ &-c < Z_l - Z_0 < b, r_1 + r_2 + 1 \leq l \leq r_1 + r_2 + r_3\}. \end{aligned}$$

In order to obtain a simpler closed form expression for $P(CD)$ note that for $\rho < \frac{1}{2}$, $b < c$, and hence $-b > -c$ giving the $P(CD)$ as

$$\begin{aligned} P(CD|\mathcal{P}_N) &\geq (-1) + 2P(-b < Z_j - Z_0 < b, \quad r_1 + 1 \leq j \leq r_1 + r_2, \quad -b < Z_l - Z_0 < b, \\ &\quad r_1 + r_2 + 1 \leq l \leq r_1 + r_2 + r_3) \\ &= (-1) + 2P(-b + Z_0 < Z_i < b + Z_0, \quad r_1 + 1 \leq i \leq r_1 + r_2 + r_3) \\ &= (-1) + 2E[\int_{-\infty}^{+\infty} \{\Phi(z+b) - \Phi(z-b)\}^{r_2+r_3} \phi(z) dz | Z_0 = z]. \end{aligned} \tag{13}$$

Rewriting the first result in the Theorem, we have $\sqrt{N}\frac{\rho a}{\sqrt{8\sigma}} \xrightarrow{P} b$ as $a \to 0$, and next using the "dominated convergence theorem" completes the proof of part (iv). □

Next, for the the "fine-tuned purely sequential procedure" (10) the "second-order expansions" are presented to asymptotically quantify the number of observations the procedure will sample compared to the optimal same size. We will also show that the proposed procedure is asymptotically unbiased version of the procedure in Solanky and Zhou (2015) in achieving the targeted value of probability requirement.

Theorem 2: The"fine-tuned purely sequential procedure" (10), have the following properties as $a \to 0$:

(i) $E(N) = C - \epsilon + \frac{\nu - 2}{k+1} + \circ(1)$, provided that $m_0 \geq 1 + \frac{2}{k+1}$;

(ii) $E(N^\omega) = C^\omega + [\omega(\frac{\nu-2}{k+1} - \epsilon) + \frac{1}{2}\omega(\omega-1)\frac{2}{k+1}]C^{\omega-1} + \circ(1)$, provided that (a) $m_0 \geq 1 + (3-\omega)\frac{2}{k+1}$ for $\omega \in (\infty, 2) - \{-1, 1\}$; (b) $m_0 \geq 1 + \frac{2}{k+1}$ for $\omega = 1$ and $\omega \geq 2$; (c) $m_0 \geq 1 + \frac{4}{k+1}$ for $\omega = -1$.

(iii) $P(CD|\mathcal{P}_N; \boldsymbol{\mu}^0(r_1, r_2, r_3, r_4), \ \sigma^2) > P^* + o(\frac{1}{C})$ provided that $m_0 > \frac{5}{k+1} + 1$ for the LFC;

where $C = \frac{8b^2\sigma^2}{(\rho a)^2}$, and the constant $\nu(k)$ is defined in equation (16).

Proof: First note that as in Woodroofe (1977) and Theorem 2.4.8 of Mukhopadhyay and Solanky (1994), the pooled estimator $U(n)$ of σ^2 can be written as sum of i.i.d. random variables as $U(n) = \frac{1}{n-1}\sum_{i=1}^{n-1} R_i'$, where R_1', R_2', $\cdots$ are i.i.d. $\frac{1}{k+1}\sigma^2\chi^2_{k+1}$ random variables. Let's write $R_i = (k+1)\sigma^{-2}R_i'$, with R_i being i.i.d. χ^2_{k+1}. Using this the purely sequential procedure could be rewritten as

$$N = \text{Inf}\left\{n \geq m_0 : n + \epsilon \ \geq \frac{8b^2}{(\rho a)^2}(n-1)^{-1}\frac{\sigma^2}{k+1}\sum_{i=1}^{n-1} R_i\right\}.$$

This can rewritten as

$$N = \text{Inf}\left\{n \geq m_0 : C^{-1}(n+\epsilon)(n-1)(k+1) \geq \sum_{i=1}^{n-1} R_i\right\}. \tag{14}$$

Next, following Mukhopadhyay and Solanky's (1994) Theorem 2.4.8 we express N as $T+1$, for the stopping rule T defined as

$$T = \text{Inf}\left\{ n \geq m-1 : C^{-1} n^2 (k+1) \left(1 + \frac{\epsilon+1}{n}\right) \geq \sum_{i=1}^{n} R_i \right\}. \tag{15}$$

The stopping rule T has been expressed in the general form of "equation (2.4.7) in Mukhopadhyay and Solanky (1994)" with the values: $L_0 = \epsilon + 1$; $\delta = 2$; $h^* = \frac{k+1}{C}$; $\theta = E(R_1) = k+1$; $r^2 = E(R_1^2) - \theta^2 = 2(k+1)$; $\beta^* = \frac{1}{\delta - 1}$ which simplifies to 1; the value of $n_0^* = (\frac{\theta}{h^*})^{\beta^*}$ simplifies to C; and the constant $p = \frac{\beta^{*2} r^2}{\theta^{-2}}$ becomes $\frac{2}{k+1}$. Also, note that with R_i being i.i.d. χ^2_{k+1}, we can express

$$P(R_i \leq r) = \int_0^r \frac{e^{\frac{-x}{2}} x^{\frac{k+1}{2}-1}}{2^{\frac{k+1}{2}} \Gamma(\frac{k+1}{2})}\, dx.$$

Since $e^{\frac{-x}{2}} \leq 1$, replacing $e^{\frac{-x}{2}}$ by 1 and carrying out the integral one will obtain

$$P(R_i \leq r) \leq \frac{2}{k+1} \frac{1}{2^{\frac{k+1}{2}} \Gamma(\frac{k+1}{2})} r^{\frac{k+1}{2}}.$$

Also, note that for the random variable R, we have

$P(R_i \leq r) \leq H r^h$, where $H = \frac{2}{k+1} \frac{1}{2^{\frac{k+1}{2}} \Gamma(\frac{k+1}{2})}$ and $h = (k+1)\big/2$. Next, using the nonlinear renewal theory from "Mukhopadhyay and Solanky (section 2.4.2)" we define the constant $\nu = \nu(k)$ as below:

$$\nu = \nu(k) = \frac{k+3}{2} - \sum_{n=1}^{\infty} \frac{1}{n} E\left[\max\left\{0, \chi^2_{(k+1)n} - 2(k+1)n\right\}\right]. \tag{16}$$

Next, following "equation (2.4.10) in Mukhopadhyay and Solanky (1994)" the constant $\eta = \frac{\beta^* \nu}{\theta} - L_0 \beta^* - \frac{\delta \beta^{*2} r^2}{2\theta^2}$ would become

$$\eta = (k+1)^{-1}(\nu - 2) - (\epsilon + 1). \tag{17}$$

Since $T = N - 1$ and therefore $E(T) = E(N) - 1$, next with η as defined above in (17), applying the "Theorem 2.4.8(v) of Mukhopadhyay and Solanky (1994)" would lead to

$$E(N) = C - \epsilon + \frac{\nu - 2}{k+1} + \circ(1),$$

provided that $m > 1 + 2(k+1)^{-1}$. This completes the proof of part (i).

The proof of the part (ii) follows by expressing the stopping variable T in the form of "Mukhopadhyay and Solanky (1994)'s equation (2.4.7)" and applying the "Theorem 2.4.8 of Mukhopadhyay and Solanky (1994)". The details are omitted for brevity. One may note that the part (i) is a special case of part (ii) when $\omega = 1$.

For part (iii), note that from (13) we can express the probability of correctly partitioning all k populations using the decision rule $\mathcal{P}_N$ under the LFC as

$$P(CD|\mathcal{P}_N) \;\geq\; (-1) + 2E[\int_{-\infty}^{\infty} \{\Phi(z+b) - \Phi(z-b)\}^{k_1} \phi(z)dz | Z_0 = z].$$

Next, we write $b = 2^{1/2}x$ and define a function $\beta(x)$ as

$$\beta(x) = \int_{-\infty}^{\infty} \left\{\Phi(z + 2^{1/2}x) - \Phi(z - 2^{1/2}x)\right\}^{k_1} \phi(z)dz.$$

It is easy to verify that

$$\beta'(x) = \int_{-\infty}^{\infty} k_1 2^{1/2} \left\{\Phi(z + 2^{1/2}x) - \Phi(z - 2^{1/2}x)\right\}^{k_1 - 1} (\phi(z + 2^{1/2}x) + \phi(z - 2^{1/2}x))\phi(z)dz$$

$$\begin{aligned}\beta''(x) \;=\; & \int_{-\infty}^{\infty} 2k_1(k_1-1) \left\{\Phi(2^{1/2}x + z) - \Phi(-2^{1/2}x + z)\right\}^{k_1-2} \\ & (\phi(2^{1/2}x + z) + \phi(-2^{1/2}x + z))^2 - 2k_1 \left\{\Phi(2^{1/2}x + z) - \Phi(-2^{1/2}x + z)\right\}^{k_1-1} \\ & \left((2^{1/2}x + z)\phi(2^{1/2}x + z) + (2^{1/2}x - z)\phi(-2^{1/2}x + z))\right) \phi(z)dz.\end{aligned}$$

Next, we write $g(x)$ for the function $\beta(bx^{\frac{1}{2}})$ and then we can express

$$\begin{aligned} g'(x) &= \frac{1}{2} b x^{-\frac{1}{2}} \beta'(bx^{\frac{1}{2}}) \\ g''(x) &= \frac{1}{4} b^2 x^{-1} \beta''(bx^{\frac{1}{2}}) - \frac{1}{4} b x^{-\frac{3}{2}} \beta'(bx^{\frac{1}{2}}) \end{aligned} \tag{18}$$

and

$$|g''(x)| \leq a_1 x^{-\frac{1}{2}} + a_2 x^{-1} + a_3 x^{-\frac{3}{2}},$$

where $a_1, \quad a_2 \quad a_3$ are positive constants. For the fine-tuned purely sequential procedure (10), one can verify that the distribution of N does not depend on the mean vector and $P(N < \infty) = 1$. Hence, by using "Theorem 3.2.1 of Mukhopadhyay and Solanky (1994)", we have

$$\inf_{a \to 0} P[CD|\mathcal{P}_N; \boldsymbol{\mu}^0(r_1, r_2, r_3, r_4), \;\; \sigma^2] > (-1) + 2E(g(\frac{N}{C})). \tag{19}$$

Next, taking a series expansion of the function $g(x)$ at $x = 1$ for random $Z \in (1, \frac{N}{C})$, we can write

$$g(x) = g(1) + g'(1)(x-1) + \frac{1}{2} g''(Z(x))(x-1)^2.$$

As shown above, we can express $|g''(x)| \leq \sum_{i=1}^{3} \frac{a_i}{x^{\alpha_i}}$, with $a_i > 0$ and $\alpha_1 > 0$, by applying "Lemma 3.5.1 of Mukhopadhyay and Solanky (1994)", we can show that for $m_0 > \frac{5}{k+1} + 1$, one will obtain

$$E(g(\frac{N}{C})) = g(1) + \frac{1}{C} g'(1) E(N - C) + \frac{1}{2C^2} E(g''(Z(x))(N - C)^2).$$

With $\alpha_1 = \frac{1}{2}, \alpha_2 = 1$, and $\alpha_3 = \frac{3}{2}$ and maximum of $(\alpha_1, \alpha_2, \alpha_3) = \frac{3}{2}$, the "Lemma 3.5.1 of Mukhopadhyay and Solanky (1994)" leads to

$$E(g(\frac{N}{C})) = g(1) + \frac{1}{C}g'(1)E(N-C) + \frac{1}{k+1}\frac{1}{C}g''(1) + o(\frac{1}{C}) \tag{20}$$

Following "Mukhopadhyay and Datta (1995)'s Theorem 2.1" with the function $g(.)$ defined as above, and ϵ as:

$$\epsilon = \epsilon\,(k, P^*) = (k+1)^{-1}[\nu - 2 + g''(1)(g'(1))^{-1}], \tag{21}$$

the $E(N)$ would simplify to

$$E(N) = C - (k+1)^{-1}[g''(1)(g'(1))^{-1}] + o(1), \text{ provided } m_0 > 1 + \frac{2}{k+1}. \tag{22}$$

Using the equations (20) and (22), one can obtain

$$E(g(\frac{N}{C})) = g(1) + o(\frac{1}{n^*}), \text{ provided } m_0 > 1 + \frac{5}{k+1}.$$

That is,

$$\inf_{a\to 0} P[CD|\mathcal{P}_N; \boldsymbol{\mu}^0(r_1, r_2, r_3, r_4),\ \ \sigma^2] > (-1) + 2g(1) + o(\frac{1}{n^*}), \text{ provided } m_0 > 1 + \frac{5}{k+1}.$$

Note that $g(1) = \frac{1}{2}(P^* + 1)$ to conclude part(iii) of the theorem. □

3. Simulation Study for the Fine-tuned Purely Sequential Procedure

In this section using the Monte Carlo simulation study, the "fine-tuned purely sequential procedure" (10) is replicated independently 5,000 times by picking different values of design constants to study how the asymptotic values provided in the Theorems 2.1 and 2.2 compare with the observed values when the procedure is simulated for small and moderate sample sizes. For the simulation results presented in the Table 1, we selected $k = 8$ and $P^* = 0.95$ and generated two independent normal populations at each of the four boundaries to generate data from LFC, as outlined in the Section 1. In order to obtain the value of constant b for given choice of k and P^*, we used the "Table 1 in Solanky and Zhou (2015)". For example, the value of constant b for $k = 8$ and $P^* = 0.95$ is 2.6959. By picking the values of the "optimal sample sizes" (=C) as 25, 50, 100, 200, 400, and 800 and taking $r = \frac{1}{3}$, we obtained the values of constant a which are reported in the Table 1. In a practical application, the values δ_1 and δ_4 would be chosen so as to reflect the "Good populations" and "Bad populations" based on the situation. Note that the difference between the "Good populations" and "Bad populations" is $a = \delta_4 - \delta_1$. In the Table 1, by picking $r = \frac{1}{3}$ we have divided the indifference-zone of Tong (1969) into three non-overlapping regions of equal size. The "middle" of these three regions would serve as the region which we will serve as Ω_M and the other two as indifference-zones Ω_{I_1} and Ω_{I_1} as defined in (5).

The procedure (10) starts with the initial sample size of $m_0 = 5$ observations from each of the 8 populations and the control population. Then, each additional observation is

Table 1: Performance of the Fine-Tuned Purely Sequential Procedure (10)

$(k = 8,\ P^* = 0.95,\ \rho = \frac{1}{3},\ \sigma = 1.0,\ m_0 = 5)$

C	ρa	a	$\bar{N}$	$std(\bar{N})$	$\bar{P}$	$std(\bar{P})$
25	1.5250	4.575	25.3246	0.0347	0.9710	0.0024
50	1.0784	3.2351	50.2660	0.0481	0.9720	0.0023
100	0.7625	2.2875	100.2612	0.0682	0.9776	0.0021
200	0.5392	1.6175	200.2736	0.0943	0.9746	0.0022
400	0.3813	1.1437	400.2050	0.1327	0.9772	0.0021
800	0.2696	0.8088	800.2242	0.1890	0.9750	0.0022

collected according to the following stopping rule (10). In the Table 1, the average value of the stopping time N and the proportion of the times all the 8 populations were partitioned correctly are reported as $\bar{N}$ and $\bar{P}$ respectively. The Table also reports the standard deviation of these two reported statistics as $std(\bar{N})$ and $std(\bar{P})$. For all the cases considered in the Table 1, the $\bar{P}$ is larger than the desired probability of correct decision which was selected to be 0.95. Also, the $\bar{N}$ matches the optimal sample size values $(=C)$ indicating that the "fine-tuned purely sequential procedure" (10) over-samples by less than half an observation on the average. The overall findings in the Table 1 confirm the theoretical results which were derived asymptotically in the Theorems 2.1 and 2.2 for the "fine-tuned purely sequential procedure" (10) are met even for small and moderate sample sizes.

Acknowledgements

It is my honor to receive an invitation to prepare this paper for the special issue of *Statistics and Applications* in honor of Professor Aloke Dey. I heartily thank the Guest Editors.

References

Bechhofer, R. E. (1954). A single-sample multiple decision procedure for ranking means of normal populations with known variances. *Annals of Mathematical Statistics*, **25**, 16-39.

Buzaianu, E. M. (2019). Selection among Bernoulli populations in comparison with a standard. *Sequential Analysis*, **38**, 184-198.

Datta, S. and Mukhopadhyay, N. (1998). Second-order asymptotics for multistage methodologies in partitioning a set of normal populations having a common unknown variance. *Statistics and Decisions*, **16**, 191-205.

Dunnett, C. W. (1955). A multiple comparison procedure for comparing several treatments with a control. *Journal of American Statistical Association*, **50**, 1096-1121.

Gupta, S. S. (1956). On a decision rule for a problem in ranking means. *Ph.D. dissertation, University of North Carolina*, Chapel Hill.

Paulson, E. (1952). On the comparison of several experimental categories with a control. *Annals of Mathematical Statistics*, **23**, 239-246.

Mukhopadhyay, N. and Datta, S. (1995). On fine-tuning a purely sequential procedure and associated second-order properties. *Sankhya: The Indian Journal of Statistics*, **57**, Series A, Pt.1: 100-117.

Mukhopadhyay, N. and Solanky, T. K. S. (1994). *Multistage Selection and Ranking Procedures*, New York: Dekker.

Roessler, E. B. (1946). Testing the significance of observations compared with a control. *Proceedings of American Society for Horticultural Science*, **47**, 249-251.

Solanky, T. K. S. (2001). A sequential procedure with elimination for partitioning a set of normal populations having a common unknown variance. *Sequential Analysis*, **20**, 279-292.

Solanky, T. K. S. (2006). A two-stage procedure with elimination for partitioning a set of normal populations with respect to a control. *Sequential Analysis*, **25**, 297-310.

Solanky, T. K. S. and Wu, Y. (2004). On unbalanced multistage methodologies for the partition problem, in *Proceeding of International Sri Lankan Statistical Conference: Version of Futuristic Methodologies*, B. Silva and N. Mukhopadhyay, eds., pp. 447-466, Sri Lanka: University of Peradeniya ; Melbourne: RMIT University.

Solanky, T. K. S. and Zhou, J. (2015). A generalization of the partition problem. *Sequential Analysis*, **34**, 483 – 503.

Tong, Y. L. (1969). On partitioning a set of normal populations by their locations with respect to a control. *Annals of Mathematical Statistics*, **40**, 1300-1324.

Woodroofe, M. (1977). Second order approximations for sequential point and interval estimation. *Annals of Statistics*, **5**, 984-995.

Woodroofe, M. (1991). The role of renewal theory in sequential analysis. *Chapter 6, Handbook of Sequential Analysis*, (eds. B.K. Ghosh and P.K. Sen), 145-167. Marcel Dekker, Inc., New York.

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 417–429

Unsaturated Log-Linear Model Selection for Categorical Data Analysis

Subir Ghosh[1] **and Arnab Chowdhury**[2]
[1]*Department of Statistics*
University of California, Riverside, California, USA
[2]*Department of Computational and Quantitative Medicine*
City of Hope National Medical Center, Duarte, California, USA

Received: 25 February 2021; Revised: 23 March 2021; Accepted: 25 March 2021

Abstract

The paper proposes a new metric SAVE for finding the best fitted unsaturated log-linear model to describe the categorical data in a contingency table with m categorical variables. Two kinds of extensions, standard and orthogonal, of an unsaturated log-linear model to the saturated model are the foundation of SAVE. The performance of SAVE in terms of the correct model parameter(s) detection is comparable with or even better than the commonly used metrics: Deviance, AIC, and BIC, as demonstrated in simulation studies.

Key words: Categorical; Log-linear; Model selection; Orthogonal extension; Unsaturated.

AMS Subject Classifications: 62H17, 62R07, 62B10.

1. Introduction

Let $X_1, \ldots, X_m$ denote m categorical variables, X_i with I_i categories, $i = 1, \ldots, m$. The n subjects selected in a study using a multinomial sample are cross-classified into $N = I_1 \times \cdots \times I_m$ possible combinations on m categorical variables $X_1, \ldots, X_m$ in a contingency table. The wth combination is represented by $(w_1, \ldots, w_m)$, where w_u is the level of X_u; $w_u = 0, \ldots, I_u - 1; u = 1, \ldots, m$. The number of subjects for the wth combination is a random variable Y_w having the observed value y_w and the expected value $E(Y_w) = \mu_w = np_w$, where p_w and μ_w are unknown parameters. The μ_w is the cell mean and p_w is the cell probability for the cell represented by the wth combination. We have $Y_w \geq 0$ and $Y_1 + \cdots + Y_N = n$, $p_1 + \cdots + p_N = 1$ and $\mu_1 + \cdots + \mu_N = n$. Also, $y_w \geq 0$, $w = 1, \ldots, N$, and $y_1 + \cdots + y_N = n$. The saturated log-linear model is

$$log(p_w) = \lambda + \delta_1 \lambda_{w_1}^{X_1} + \cdots + \delta_m \lambda_{w_m}^{X_m} + \delta_1\delta_2 \lambda_{w_1w_2}^{X_1X_2} + \cdots + \delta_1\delta_2\delta_3 \lambda_{w_1w_2w_3}^{X_1X_2X_3} + \cdots + \delta_1 \ldots \delta_m \lambda_{w_1 \ldots w_m}^{X_1 \ldots X_m}, \quad (1)$$

where $\{\lambda_{w_{i_1}w_{i_2}}^{X_{i_1}X_{i_2}}\}, \{\lambda_{w_{i_1}w_{i_2}w_{i_3}}^{X_{i_1}X_{i_2}X_{i_3}}\}, \ldots$, and $\lambda_{w_1 \ldots w_m}^{X_1 \ldots X_m}$, are the unknown association parameters. The $\{\lambda_{w_i}^{X_i}\}$ are the unknown effect parameters. The λ is the unknown overall effect parameter.

Correponding Author: Subir Ghosh
Email: subir.ghosh@ucr.edu

The $\delta_u, u = 1, \ldots, m$, are

$$\delta_u = \begin{cases} 0 & \text{if} \quad w_u = 0, \\ 1 & \text{if} \quad w_u \neq 0. \end{cases}$$

When at least one association parameter is zero or absent in the saturated model, the model becomes an unsaturated model in presence of the overall effect and the effect parameters. The unsaturated models considered in this paper consist of the overall effect, the effect parameters, and one or more association parameters. When all association parameters are absent in the unsaturated model, the categorical variables become mutually independent. "In practice, unsaturated models are preferable since their fit smoothes the sample data and has simpler interpretations" (page 341, Agresti (2013)). On the one hand, the over-fitted saturated model is unnecessary, but on the other hand, an under-fitted unsaturated model is deficient for describing the data. We propose a new method of finding the best fitted unsaturated log-linear model using the association parameters absent in the model considered but present in the saturated model. We compare the proposed method with the standard measures such as AIC, BIC, and Deviance using the 100,000 realizations of simulated data.

In Section 2, we present two saturated representations of standard and orthogonal extensions of unsaturated log-linear models. In Section 3, we explain the saturated representations with two illustrative examples in Sections 3.1 and 3.2. The data on the use of automobile seat-belt for lowering fatal injury is in Section 4. We propose the new metric, SAVE, in Section 5. We compare the new metric with the other available metrics AIC, BIC, and MDI in Section 5.1. Section 6 presents their performance comparison for the 100,000 simulated data from each of the six data-generating models. We conclude in Section 7 with some remarks.

2. Two Saturated Representations : $S1$ and $S2$

Let $\boldsymbol{p} = (p_1, \ldots, p_N)^\top$ be the column vector of expected counts for the N cells of the contingency table, $\boldsymbol{\lambda}^{(1)}$ $(k_1 \times 1)$ be the vector of the overall effect, the effect parameters, and the one or more association parameters in an unsaturated model considered for fitting to the collected data, and $\boldsymbol{X}_1$ $(N \times k_1)$ be the model matrix generated from the indicator variables for the parameters in $\boldsymbol{\lambda}_1^{(1)}$. Let $\boldsymbol{\lambda}_2$ $(k_2 \times 1)$ be the vector of association parameters that are absent in $\boldsymbol{\lambda}^{(1)}$ and $\boldsymbol{X}_2$ $(N \times k_2)$ be the model matrix generated from the indicator variables for the parameters in $\boldsymbol{\lambda}_2$. In the saturated model (1), the parameters in both $\boldsymbol{\lambda}^{(1)}$ and $\boldsymbol{\lambda}_2$ are present. The unsaturated model consists of the parameters in $\boldsymbol{\lambda}^{(1)}$ but not the parameters in $\boldsymbol{\lambda}_2$. The matrix representation of the unsaturated model considered is

$$log\boldsymbol{p} = \boldsymbol{X}_1 \boldsymbol{\lambda}_1^{(1)}, \tag{2}$$

where $\text{rank}(\boldsymbol{X}_1) = k_1$. We consider two representations of the saturated model. The first representation is the standard saturated model and we denote it by $S1$. The second representation is the orthogonal extension of the assumed unsaturated model in (2) and it is denoted by $S2$ (Klimova, Rudas and Dobra (2012), Klimova and Rudas (2016), Rudas (2018)). The

standard representation $S1$ of the saturated log-linear model is

$$log\boldsymbol{p} = \boldsymbol{X}_1\boldsymbol{\lambda}_1^{(1)} + \boldsymbol{X}_2\boldsymbol{\lambda}_2, \tag{3}$$

where $\text{rank}(\boldsymbol{X}_1, \boldsymbol{X}_2) = k_1 + k_2 = N$.

Let $\boldsymbol{D}$ $(N \times k_2)$ be a matrix which satisfies

$$rank(\boldsymbol{D}) = k_2, \boldsymbol{X}_1^\top \boldsymbol{D} = \boldsymbol{0}. \tag{4}$$

The matrix $\boldsymbol{D}$ is not unique. A simple form of the matrix $\boldsymbol{D}$ satisfying (4) is

$$\boldsymbol{D} = [\boldsymbol{I}_N - \boldsymbol{X}_1(\boldsymbol{X}_1^\top \boldsymbol{X}_1)^{-1}\boldsymbol{X}_1^\top]\boldsymbol{X}_2, \tag{5}$$

where $\boldsymbol{I}_N$ is the $(N \times N)$ identity matrix. Note that $\text{rank}([\boldsymbol{I}_N - \boldsymbol{X}_1(\boldsymbol{X}_1^\top \boldsymbol{X}_1)^{-1}\boldsymbol{X}_1^\top]) = N - k_1 = k_2 = \text{rank}(\boldsymbol{X}_2) = \text{rank}(\boldsymbol{D})$. From (5), it can be seen

$$\begin{aligned} \boldsymbol{D}\boldsymbol{\lambda}_2 &= [\boldsymbol{I}_N - \boldsymbol{X}_1(\boldsymbol{X}_1^\top \boldsymbol{X}_1)^{-1}\boldsymbol{X}_1^\top]\boldsymbol{X}_2\boldsymbol{\lambda}_2, \\ \boldsymbol{X}_2\boldsymbol{\lambda}_2 &= \boldsymbol{D}\boldsymbol{\lambda}_2 + \boldsymbol{X}_1(\boldsymbol{X}_1^\top \boldsymbol{X}_1)^{-1}\boldsymbol{X}_1^\top \boldsymbol{X}_2\boldsymbol{\lambda}_2. \end{aligned} \tag{6}$$

Let

$$\boldsymbol{\lambda}_1^{(2)} = \boldsymbol{\lambda}_1^{(1)} + (\boldsymbol{X}_1^\top \boldsymbol{X}_1)^{-1}\boldsymbol{X}_1^\top \boldsymbol{X}_2\boldsymbol{\lambda}_2. \tag{7}$$

The orthogonal saturated extension of the unsaturated model in (2), $S2$, is obtained from (3) and (7) as

$$\begin{aligned} log\boldsymbol{p} &= \boldsymbol{X}_1\boldsymbol{\lambda}_1^{(1)} + \boldsymbol{X}_2\boldsymbol{\lambda}_2 \\ &= \boldsymbol{X}_1\boldsymbol{\lambda}_1^{(1)} + \boldsymbol{D}\boldsymbol{\lambda}_2 + \boldsymbol{X}_1(\boldsymbol{X}_1^\top \boldsymbol{X}_1)^{-1}\boldsymbol{X}_1^\top \boldsymbol{X}_2\boldsymbol{\lambda}_2 \\ &= \boldsymbol{X}_1\left(\boldsymbol{\lambda}_1^{(1)} + (\boldsymbol{X}_1^\top \boldsymbol{X}_1)^{-1}\boldsymbol{X}_1^\top \boldsymbol{X}_2\boldsymbol{\lambda}_2\right) + \boldsymbol{D}\boldsymbol{\lambda}_2 \\ &= \boldsymbol{X}_1\boldsymbol{\lambda}_1^{(2)} + \boldsymbol{D}\boldsymbol{\lambda}_2. \end{aligned} \tag{8}$$

From (4) and (8), it follows that

$$\begin{aligned} \boldsymbol{\lambda}_1^{(2)} &= (\boldsymbol{X}_1^\top \boldsymbol{X}_1)^{-1}\boldsymbol{X}_1^\top log\boldsymbol{p} \\ \boldsymbol{\lambda}_2 &= (\boldsymbol{D}^\top \boldsymbol{D})^{-1}\boldsymbol{D}^\top log\boldsymbol{p}. \end{aligned} \tag{9}$$

Klimova, Rudas and Dobra (2012), Klimova and Rudas (2016), Rudas (2018) defined two kinds of relational models, dual and non-dual. For a dual representation of a relational model, we have $\boldsymbol{D}^\top log\boldsymbol{p} = 0$. In other words, from (9), $\boldsymbol{\lambda}_2 = 0$. Hence, the unsaturated model in (2) has a dual representation. On the other hand, for a non-dual representation of a relational model, we have $\boldsymbol{D}^\top log\boldsymbol{p} \neq 0$. Therefore, the saturated model in (3) has a non-dual representation.

3. Examples

3.1. Example 1

For a $2 \times 2 \times 2$ contingency table and the unsaturated model in (2) with three independent categorical variables X_1, X_2 and X_3, we have $m = 3$, $N = 8$, $k_1 = k_2 = 4$. Table 1 presents the cell representations.

Table 1: The cell representations for Example 1

Number w	Combination (w_1, w_2, w_3)	Probability p_w
1	(0, 0, 0)	p_1
2	(0, 0, 1)	p_2
3	(0, 1, 0)	p_3
4	(0, 1, 1)	p_4
5	(1, 0, 0)	p_5
6	(1, 0, 1)	p_6
7	(1, 1, 0)	p_7
8	(1, 1, 1)	p_8

The matrices $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are

$$\boldsymbol{X}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \boldsymbol{X}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \tag{10}$$

and the vectors $\boldsymbol{\lambda}_1^{(1)}$ and $\boldsymbol{\lambda}_2$ in (3) are

$$\boldsymbol{\lambda}_1^{(1)} = (\lambda, \lambda_1^{X_1}, \lambda_1^{X_2}, \lambda_1^{X_3})^\top, \boldsymbol{\lambda}_2 = (\lambda_{11}^{X_1X_2}, \lambda_{11}^{X_1X_3}, \lambda_{11}^{X_2X_3}, \lambda_{111}^{X_1X_2X_3})^\top. \tag{11}$$

Two $\boldsymbol{D}$ matrices, $\boldsymbol{D}_{(1)}$ and $\boldsymbol{D}_{(2)}$ in (12), are obtained by using (5) and (10). The last column of $\boldsymbol{D}_{(1)}$ is not orthogonal to its first three columns. The first three columns of $\boldsymbol{D}_{(1)}$ are mutually orthogonal. The first three columns of $\boldsymbol{D}_{(2)}$ are the same as the corresponding columns in $\boldsymbol{D}_{(1)}$. The four columns of $\boldsymbol{D}_{(2)}$ are mutually orthonormal. Thus, $\boldsymbol{D}_{(2)}^\top \boldsymbol{D}_{(2)} = \boldsymbol{I}_4$.

$$\boldsymbol{D}_{(1)} = (1/4) \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 0 \\ -1 & 1 & -1 & 0 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & 1 & 0 \\ -1 & 1 & -1 & -1 \\ 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 2 \end{bmatrix}, \boldsymbol{D}_{(2)} = (1/8) \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}. \tag{12}$$

For a dual relational model by using the expression of $\boldsymbol{D}_{(2)}$ in (12) for $\boldsymbol{D}$, we find

$$\begin{aligned}(i).\ log\left(\frac{p_1p_2p_7p_8}{p_3p_4p_5p_6}\right)=0,\ \ (ii).\ log\left(\frac{p_1p_3p_6p_8}{p_2p_4p_5p_7}\right)=0,\\ (iii).\ log\left(\frac{p_1p_4p_5p_8}{p_2p_3p_6p_7}\right)=0,\ \ (iv).\ log\left(\frac{p_1p_4p_6p_7}{p_2p_3p_5p_8}\right)=0.\end{aligned} \tag{13}$$

From the equations (iii) and (iv) in (13), it can be seen

$$(i).\ log\left(\frac{p_1p_4}{p_2p_3}\right)=0,\ \ (ii).\ log\left(\frac{p_5p_8}{p_6p_7}\right)=0. \tag{14}$$

In Table 1, we observe that $X_1 = 0$ for $w = 1, 2, 3, 4$ and $X_1 = 1$ for $w = 5, 6, 7, 8$. The equation (i) in (14) implies the conditional independence between the categorical variables X_2 and X_3 given $X_1 = 0$. The equation (ii) in (14) implies the conditional independence between X_2 and X_3 given $X_1 = 1$.

From the equations (i) and (iv) in (13), we observe

$$(i).\ log\left(\frac{p_1p_6}{p_2p_5}\right)=0,\ \ (ii).\ log\left(\frac{p_4p_7}{p_3p_8}\right)=0. \tag{15}$$

In Table 1, we observe that $X_2 = 0$ for $w = 1, 2, 5, 6$ and $X_2 = 1$ for $w = 3, 4, 7, 8$. The equation (i) in (15) implies the conditional independence between the categorical variables X_1 and X_3 given $X_2 = 0$. The equation (ii) in (15) implies the conditional independence between X_1 and X_3 given $X_2 = 1$.

From the equations (i) and (iv) in (13), we find

$$(i).\ log\left(\frac{p_1p_7}{p_3p_5}\right)=0,\ \ (ii).\ log\left(\frac{p_2p_8}{p_4p_6}\right)=0. \tag{16}$$

In Table 1, we observe that $X_3 = 0$ for $w = 1, 3, 5, 7$ and $X_3 = 1$ for $w = 2, 4, 6, 8$. The equation (i) in (16) implies the conditional independence between the categorical variables X_1 and X_2 given $X_3 = 0$. The equation (ii) in (16) implies the conditional independence between X_1 and X_2 given $X_3 = 1$.

3.2. Example 2

For a 3×2 contingency table and the unsaturated model in (2) with two independent categorical variables X_1 and X_2, we have $m = 2$, $N = 6$, $k_1 = 4, k_2 = 2$. Table 2 presents the cell representations.

The matrices $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are

$$\boldsymbol{X}_1=\begin{bmatrix}1&0&0&0\\1&0&0&1\\1&1&0&0\\1&1&0&1\\1&0&1&0\\1&0&1&1\end{bmatrix},\boldsymbol{X}_2=\begin{bmatrix}0&0\\0&0\\0&0\\1&0\\0&0\\0&1\end{bmatrix}, \tag{17}$$

Table 2: The cell representations for Example 2

Number w	Combination (w_1, w_2)	Probability p_w
1	(0, 0)	p_1
2	(0, 1)	p_2
3	(1, 0)	p_3
4	(1, 1)	p_4
5	(2, 0)	p_5
6	(2, 1)	p_6

and the vectors $\boldsymbol{\lambda}_1^{(1)}$ and $\boldsymbol{\lambda}_2$ in (3) are

$$\boldsymbol{\lambda}_1^{(1)} = (\lambda, \lambda_1^{X_1}, \lambda_2^{X_1}, \lambda_1^{X_2})^\top, \boldsymbol{\lambda}_2 = (\lambda_{11}^{X_1X_2}, \lambda_{21}^{X_1X_2})^\top. \tag{18}$$

The matrices $\boldsymbol{D}_{(1)}$ and $\boldsymbol{D}_{(2)}$ in (19) are obtained by using (5) and (17). The two columns of $\boldsymbol{D}_{(1)}$ are not mutually orthogonal. The two columns of $\boldsymbol{D}_{(2)}$ are mutually orthonormal. Thus, $\boldsymbol{D}_{(2)}^\top \boldsymbol{D}_{(2)} = \boldsymbol{I}_2$.

$$\boldsymbol{D}_{(1)} = (1/6)\begin{bmatrix} 1 & 1 \\ -1 & -1 \\ -2 & 1 \\ 2 & -1 \\ 1 & -2 \\ -1 & 2 \end{bmatrix}, \boldsymbol{D}_{(2)} = \begin{bmatrix} 1 & 1 \\ -1 & -1 \\ -2 & 0 \\ 2 & 0 \\ 1 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} (1/2\sqrt{3}) & 0 \\ 0 & (1/2) \end{bmatrix}. \tag{19}$$

For a dual relational model by using the expression of $\boldsymbol{D}_{(2)}$ in (19) for $\boldsymbol{D}$, we find

$$(i).\, log\Big(\frac{p_1p_5}{p_2p_6}\Big) = 2 \times log\Big(\frac{p_3}{p_4}\Big),\ \ (ii).\, log\Big(\frac{p_1p_6}{p_2p_5}\Big) = 0. \tag{20}$$

4. A Real Data

A research investigation started with a question (Agresti (2013)): Does seat-belt use in automobiles reduce injury? The collected data in Table 4 were on the injury outcomes of 68,694 passengers in autos and light trucks involved in accidents one year in the state of Maine, USA. Three factors each at two levels displayed in Table 3 were three categorical variables ($m = 3$) for the Table 4 data.

For the vectors $\boldsymbol{\lambda}_1^{(1)}$ and $\boldsymbol{\lambda}_2$ in (3) as

$$\boldsymbol{\lambda}_1^{(1)} = (\lambda, \lambda_1^{X_1}, \lambda_1^{X_2}, \lambda_1^{X_3}, \lambda_{11}^{X_1X_3}, \lambda_{11}^{X_2X_3})^\top, \boldsymbol{\lambda}_2 = (\lambda_{11}^{X_1X_2}, \lambda_{111}^{X_1X_2X_3})^\top, \tag{21}$$

Table 3: Three factors and their levels

Factors/ Categories	X_i	Levels 0	1
Location	X_1	Urban	Rural
Seat-belt use	X_2	No	Yes
Injury	X_3	No	Yes

Table 4: The number of subjects y_w

w	X_1, X_2, X_3	y_w
1	000	17,668
2	001	1,808
3	010	22,556
4	011	1,139
5	100	9,369
6	101	2,057
7	110	12,827
8	111	1,270

the matrices $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ in (3), and $\boldsymbol{D}$ in (5) are

$$\boldsymbol{X}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}, \boldsymbol{X}_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}, \boldsymbol{D} = \begin{bmatrix} 0.25 & 0.00 \\ 0.25 & 0.25 \\ -0.25 & 0.00 \\ -0.25 & -0.25 \\ -0.25 & 0.00 \\ -0.25 & -0.25 \\ 0.25 & 0.00 \\ 0.25 & 0.25 \end{bmatrix}. \tag{22}$$

5. SAVE - A New Model Selection Criterion

For the saturated log-linear model $S1$ in (3), assume

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 & \boldsymbol{X}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}_{11} & \boldsymbol{X}_{12} \\ \boldsymbol{X}_{21} & \boldsymbol{X}_{22} \end{bmatrix}, \tag{23}$$

where the matrix $\boldsymbol{X}_{11}(k_1 \times k_1)$ has rank k_1 and $\boldsymbol{X}_1^\top \boldsymbol{X}_2 \neq \boldsymbol{0}$. Recall from (2) and (3) that $\text{rank}(\boldsymbol{X}_1) = k_1$ and $\text{rank}(\boldsymbol{X}) = k_1 + k_2 = N$.

For the saturated log-linear model $S2$ in (8), assume

$$\boldsymbol{X}^* = \begin{bmatrix} \boldsymbol{X}_1 & \boldsymbol{D} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}_{11} & \boldsymbol{D}_1 \\ \boldsymbol{X}_{21} & \boldsymbol{D}_2 \end{bmatrix}, \tag{24}$$

where $\text{rank}(\boldsymbol{X}^*) = k_1 + k_2 = N$. Recall from (4) that $\text{rank}(\boldsymbol{D}) = k_2$ and $\boldsymbol{X}_1^\top \boldsymbol{D} = \boldsymbol{0}$.
Let $\boldsymbol{P}$ be an $(N \times N)$ lower-diagonal matrix

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{I}_{k_1} & \boldsymbol{0} \\ -\boldsymbol{X}_{21}\boldsymbol{X}_{11}^{-1} & \boldsymbol{I}_{k_2} \end{bmatrix}. \tag{25}$$

Pre-multiplying the matrices $\boldsymbol{X}$ in (23) and $\boldsymbol{X}^*$ in (24) by $\boldsymbol{P}$ in (25)

$$\boldsymbol{P}\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_{11} & \boldsymbol{X}_{12} \\ \boldsymbol{0} & \boldsymbol{X}_{22} - \boldsymbol{X}_{21}\boldsymbol{X}_{11}^{-1}\boldsymbol{X}_{12} \end{bmatrix}, \boldsymbol{P}\boldsymbol{X}^* = \begin{bmatrix} \boldsymbol{X}_{11} & \boldsymbol{D}_1 \\ \boldsymbol{0} & \boldsymbol{D}_2 - \boldsymbol{X}_{21}\boldsymbol{X}_{11}^{-1}\boldsymbol{D}_1 \end{bmatrix}. \tag{26}$$

Let $\widehat{\boldsymbol{\lambda}}_1^{(1)}$ be the estimator of $\boldsymbol{\lambda}_1^{(1)}$ and $\widehat{\boldsymbol{\lambda}}_2^{(1)}$ of $\boldsymbol{\lambda}_2$, for $S1$ in (3). Let $\widehat{\boldsymbol{\lambda}}_1^{(2)}$ be the estimator of $\boldsymbol{\lambda}_1^{(2)}$ and $\widehat{\boldsymbol{\lambda}}_2^{(2)}$ of $\boldsymbol{\lambda}_2$, for $S2$ in (8). From (8) and (26), it can be seen that

$$\begin{aligned} \boldsymbol{X}_{11}\widehat{\boldsymbol{\lambda}}_1^{(1)} + \boldsymbol{X}_{12}\widehat{\boldsymbol{\lambda}}_2^{(1)} &= \boldsymbol{X}_{11}\widehat{\boldsymbol{\lambda}}_1^{(2)} + \boldsymbol{D}_1\widehat{\boldsymbol{\lambda}}_2^{(2)}, \\ (\boldsymbol{X}_{22} - \boldsymbol{X}_{21}\boldsymbol{X}_{11}^{-1}\boldsymbol{X}_{12})\widehat{\boldsymbol{\lambda}}_2^{(1)} &= (\boldsymbol{D}_2 - \boldsymbol{X}_{21}\boldsymbol{X}_{11}^{-1}\boldsymbol{D}_1)\widehat{\boldsymbol{\lambda}}_2^{(2)}. \end{aligned} \tag{27}$$

Clearly from (27),

$$\begin{aligned} \widehat{\boldsymbol{\lambda}}_2^{(1)} &= (\boldsymbol{X}_{22} - \boldsymbol{X}_{21}\boldsymbol{X}_{11}^{-1}\boldsymbol{X}_{12})^{-1}(\boldsymbol{D}_2 - \boldsymbol{X}_{21}\boldsymbol{X}_{11}^{-1}\boldsymbol{D}_1)\widehat{\boldsymbol{\lambda}}_2^{(2)}, \\ \widehat{\boldsymbol{\lambda}}_1^{(2)} - \widehat{\boldsymbol{\lambda}}_1^{(1)} &= \boldsymbol{X}_{11}^{-1}(\boldsymbol{X}_{12}\widehat{\boldsymbol{\lambda}}_2^{(1)} - \boldsymbol{D}_1\widehat{\boldsymbol{\lambda}}_2^{(2)}). \end{aligned} \tag{28}$$

Theorem 1: For two matrices, $\boldsymbol{X}$ in (23) in the standard representation $S1$ of the saturated log-linear model in (3) and $\boldsymbol{X}^*$ in (24) in the orthogonal extension representation $S2$ of the saturated log-linear model in (5), the estimators $\widehat{\boldsymbol{\lambda}}_1^{(1)}$ of $\boldsymbol{\lambda}_1^{(1)}$ and $\widehat{\boldsymbol{\lambda}}_2^{(1)}$ of $\boldsymbol{\lambda}_2$ for $S1$ in (3), $\widehat{\boldsymbol{\lambda}}_1^{(2)}$ of $\boldsymbol{\lambda}_1^{(2)}$ and $\widehat{\boldsymbol{\lambda}}_2^{(2)}$ of $\boldsymbol{\lambda}_2$ for $S2$ in (8), satisfy
(i) $\widehat{\boldsymbol{\lambda}}_2^{(2)} = \widehat{\boldsymbol{\lambda}}_2^{(1)}$ if $(\boldsymbol{X}_{22} - \boldsymbol{X}_{21}\boldsymbol{X}_{11}^{-1}\boldsymbol{X}_{12}) = (\boldsymbol{D}_2 - \boldsymbol{X}_{21}\boldsymbol{X}_{11}^{-1}\boldsymbol{D}_1)$,
(ii) $\widehat{\boldsymbol{\lambda}}_1^{(2)} = \widehat{\boldsymbol{\lambda}}_1^{(1)}$ if and only if $\boldsymbol{X}_{12}\widehat{\boldsymbol{\lambda}}_2^{(1)} = \boldsymbol{D}_1\widehat{\boldsymbol{\lambda}}_2^{(2)}$.

Proof: The proof follows from (28).

Theorem 2: For the orthogonal extension representation $S2$ of the saturated log-linear model in (8), the matrix $\boldsymbol{D}$ is not unique but $\boldsymbol{D}\widehat{\boldsymbol{\lambda}}_2^{(2)}$ is unique.

Proof: From (3), (8), (9), and the condition $\boldsymbol{X}_1^\top \boldsymbol{D} = \boldsymbol{0}$ in (4),

$$\begin{aligned} log\widehat{\boldsymbol{p}} &= \boldsymbol{X}_1\widehat{\boldsymbol{\lambda}}_1^{(1)} + \boldsymbol{X}_2\widehat{\boldsymbol{\lambda}}_2^{(1)} \\ &= \boldsymbol{X}_1\widehat{\boldsymbol{\lambda}}_1^{(2)} + \boldsymbol{D}\widehat{\boldsymbol{\lambda}}_2^{(2)}, \\ \widehat{\boldsymbol{\lambda}}_2^{(2)} &= (\boldsymbol{D}^\top\boldsymbol{D})^{-1}\boldsymbol{D}^\top\boldsymbol{X}_2\widehat{\boldsymbol{\lambda}}_2^{(2)} \\ &= (\boldsymbol{D}^\top\boldsymbol{D})^{-1}\boldsymbol{D}^\top log\widehat{\boldsymbol{p}}, \\ \widehat{\boldsymbol{\lambda}}_1^{(2)} &= (\boldsymbol{X}_1^\top\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1^\top log\widehat{\boldsymbol{p}}, \\ \boldsymbol{D}\widehat{\boldsymbol{\lambda}}_2^{(2)} &= log\widehat{\boldsymbol{\mu}} - \boldsymbol{X}_1\widehat{\boldsymbol{\lambda}}_1^{(2)} \\ &= [\boldsymbol{I}_N - \boldsymbol{X}_1(\boldsymbol{X}_1^\top\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1^\top]log\widehat{\boldsymbol{p}}. \end{aligned} \tag{29}$$

The right hand side of $\boldsymbol{D}\widehat{\boldsymbol{\lambda}}_2^{(2)}$ in (29) depends only on $\boldsymbol{X}_1$ and $\widehat{\boldsymbol{p}}$ but not $\boldsymbol{D}$ since the elements of $\widehat{\boldsymbol{p}}$ are $y_w/n, w = 1, \ldots, N$. Hence, $\boldsymbol{D}\widehat{\boldsymbol{\lambda}}_2^{(2)}$ is unique.

Theorem 3: The sum of the elements in $\boldsymbol{D}\widehat{\boldsymbol{\lambda}}_2^{(2)}$ is zero.

Proof: Since the first column of $\boldsymbol{X}_1$ is an $(N \times 1)$ column vector $\boldsymbol{j}_N = (1, 1, \ldots, 1)^\top$ with the elements equal to one, it follows from (4) that $\boldsymbol{j}_N^\top\boldsymbol{D} = \boldsymbol{0}$ and therefore, $\boldsymbol{j}_N^\top\boldsymbol{D}\boldsymbol{\lambda}_2^{(2)} = 0$. In other words, the sum of elements of $\boldsymbol{D}\widehat{\boldsymbol{\lambda}}_2^{(2)}$ is zero.

It follows from Theorem 3 that the non-zero elements of $\boldsymbol{D}\widehat{\boldsymbol{\lambda}}_2^{(2)}$ are either positive or negative. Moreover, the sum of the positive elements is negative of the sum of the negative values. A new model comparison criterion is proposed as

$$\begin{aligned} \text{SAVE} &= \text{The sum of the absolute values of the elements in } \boldsymbol{D}\widehat{\boldsymbol{\lambda}}_2^{(2)} \\ &= 2 \times \text{The sum of the positive elements in } \boldsymbol{D}\widehat{\boldsymbol{\lambda}}_2^{(2)}. \end{aligned} \tag{30}$$

for comparing a class of unsaturated log-linear models. Smaller the value of SAVE for a model means the better fit to describe the data. The unsaturated model having the smallest value of SAVE means the elements of $\boldsymbol{D}\widehat{\boldsymbol{\lambda}}_2^{(2)}$ are overall individually small. In other words, the unsaturated model provides the closest fitted values of p_w to their corresponding observed values y_w/n, for $w = 1, \ldots, N$.

5.1. Comparison of unsaturated models fitted to the seat-belt use data

Table 5 compares the seven unsaturated models in fitting to the Section 4 data using the four criterion functions: AIC, BIC, MDI, and SAVE. From now on, $\lambda_1^{X_1}, \lambda_1^{X_2}, \lambda_1^{X_3}, \lambda_{11}^{X_1X_2}, \lambda_{11}^{X_1X_3}$, $\lambda_{11}^{X_2X_3}$, and $\lambda_{111}^{X_1X_2X_3}$ are denoted by $\lambda_1, \lambda_2, \lambda_3, \lambda_{12}, \lambda_{13}, \lambda_{23}$, and λ_{123}, respectively.

Table 5: The comparison of seven unsaturated log-linear models

	Model	AIC	BIC	MDI	SAVE
1	$\lambda_{123} = 0$	99.14	99.69	85.14	0.09
2	$\lambda_{23} = \lambda_{123} = 0$	878.96	879.43	866.96	1.50
3	$\lambda_{13} = \lambda_{123} = 0$	830.53	831.01	818.53	1.44
4	$\lambda_{12} = \lambda_{123} = 0$	111.70	112.18	99.70	0.09
5	$\lambda_{13} = \lambda_{23} = \lambda_{123} = 0$	1596.57	1596.96	1586.57	1.51
6	$\lambda_{12} = \lambda_{23} = \lambda_{123} = 0$	877.73	878.13	867.73	1.50
7	$\lambda_{12} = \lambda_{13} = \lambda_{123} = 0$	829.31	829.71	819.31	1.47

The criterion functions AIC and BIC (Akaike (1973), Schwarz (1978)), Konishi and Kitagawa (2008)) penalize the bigger model, while the Minimum Discrimination Information (MDI) (Kullback and Leibler (1951), Kullback (1959), Csiszár (1975), Gokhale and Kullback (1978), Haberman (1984), Kullback, Keegel, and Kullback (2013)) and SAVE do not. The best-fitted model having the smallest values of all four criterion functions is the model with $\lambda_{123} = 0$. The second-best model under all four criterion functions, is the model having $\lambda_{12} = \lambda_{123} = 0$, which means the conditional independence between X_1 and X_2 given X_3. The proposed criterion function SAVE does not discriminate visibly between the top two models by the other three criterion functions numerically for the data considered.

6. A Performance Evaluation Simulation Study for a $2\times2\times2$ Contingency Table

The 100,000 multinomial random samples are generated from the six log-linear models satisfying (1) for a $2 \times 2 \times 2$ contingency table. The eight λ values for the data generating six models are given in Table 6 so that the sum of $p_w, w = 1, \ldots, 8$, is 1. The p_w values are displayed in Table 7.

Table 6: The λ parameters of the six data generating models

Parameters	$M1$	$M2$	$M3$	$M4$	$M5$	$M6$
λ	-2.4654	-4.3262	-1.3008	-2.0844	-0.7839	-3.9759
λ_1	-1.6094	0.5000	-1.6094	0.5000	-1.6094	-1.6094
λ_2	-0.9163	-0.9163	-0.9163	-0.9163	-0.9163	-0.9163
λ_3	-1.2040	-1.2040	-1.2040	-1.2040	-1.2040	-1.2040
λ_{12}	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
λ_{13}	3.2834	3.2834	0.0500	0.0500	0.0150	3.2834
λ_{23}	2.3434	2.3434	2.3434	2.3434	0.0200	2.3434
λ_{123}	0.0300	0.0300	0.0300	0.0300	0.0300	1.9738

Table 8 displays the unsaturated models fitted to the 100,000 datasets generated using each model in Table 6. The best-fitted models satisfy the criterion functions Deviance, AIC, BIC, and SAVE.

The number or proportion of times a parameter appearing or not-appearing in the best-fitted models is a measure of correct detection. For the data generating six models in Table 6, the values of λ_{12} are identical, smallest, and close to zero. Hence, smaller the number or proportion of times λ_{12} appearing in the best-fitted models is better and larger the number or proportion of times λ_{12} not-appearing in the best fitted models is better, are two equivalent measures of correct detection. Table 9 provides the comparison between Deviance Statistic/AIC/BIC and SAVE in terms of the number of times λ_{12} does not appear in the best fitted models of three groups ($g = 1, 2, 3$ in Table 8) to 100,000 datasets generated by the six models ($M_i, i = 1, \ldots, 6$, in Table 7). Table 9 demonstrates that the number of times λ_{12} does not appear in the best fitted models using the criterion function SAVE, is greater than or equal to the corresponding number which is the common value of the criterion functions Deviance, AIC, and BIC. In other words, the new criterion function SAVE makes the correct detection more frequently than the three popular criterion functions: Deviance, AIC, and BIC.

Table 7: The cell probabilities p_w of the six data generating models

w	$M1$	$M2$	$M3$	$M4$	$M5$	$M6$
(0,0,0)	0.0850	0.0132	0.2803	0.1244	0.4566	0.0188
(0,0,1)	0.0255	0.0040	0.0841	0.0373	0.1370	0.0056
(0,1,0)	0.0340	0.0053	0.1121	0.0498	0.1826	0.0075
(0,1,1)	0.1062	0.0165	0.3504	0.1555	0.0559	0.0235
(1,0,0)	0.0170	0.0218	0.0561	0.2051	0.0913	0.0038
(1,0,1)	0.1360	0.1743	0.0177	0.0647	0.0278	0.0300
(1,1,0)	0.0069	0.0088	0.0227	0.0829	0.0369	0.0015
(1,1,1)	0.5895	0.7561	0.0767	0.2825	0.0118	0.9094

For the data generating six models $M1, \ldots, M6$, the values of λ_{13} are equal and largest among the association parameters for $M1$, $M2$, and $M6$. Therefore, larger the number of times λ_{13} appearing in the best fitted models is better. Table 10 presents the comparison between Deviance Statistic/AIC/BIC and SAVE with respect to the number of times λ_{13} appears in the best fitted models of three groups($g = 1, 2, 3$ in Table 8) to 100,000 datasets generated by $M1$, $M2$, and $M6$. The SAVE makes the correct detection more frequently than Deviance/AIC/BIC for the datasets generated by $M1$ in the group $g = 1$ and for the datasets generated by $M6$ in the group $g = 2$. The performances are equal for the other cases in Table 10. The Deviance/AIC/BIC makes the correct detection more frequently than SAVE for the datasets generated by $M2$ in the group $g = 1$. Overall, SAVE performs better than Deviance/AIC/BIC.

7. Concluding Remarks

We constructed the new metric SAVE from the standard and orthogonal extensions of the unsaturated models. The construction process is simple and meaningful. We made the comparison of the metric SAVE with its competitors Deviance, AIC, and BIC. The SAVE

Table 8: The fitted models for $k = 1,\ 2,$ and 3

g	h	The fitted Model $g.h$	The common λ parameters present	The other λ parameters present	The λ parameters absent
1	1	1.1	λ, λ_1, λ_2, λ_3	λ_{123}	λ_{12}, λ_{13}, λ_{23}
	2	1.2	λ, λ_1, λ_2, λ_3	λ_{12}	λ_{123}, λ_{13}, λ_{23}
	3	1.3	λ, λ_1, λ_2, λ_3	λ_{13}	λ_{123}, λ_{12}, λ_{23}
	4	1.4	λ, λ_1, λ_2, λ_3	λ_{23}	λ_{123}, λ_{13}, λ_{12}
2	1	2.1	λ, λ_1, λ_2, λ_3	$\lambda_{123}, \lambda_{23}$	λ_{12}, λ_{13}
	2	2.2	λ, λ_1, λ_2, λ_3	$\lambda_{123}, \lambda_{13}$	λ_{12}, λ_{23}
	3	2.3	λ, λ_1, λ_2, λ_3	$\lambda_{123}, \lambda_{12}$	λ_{13}, λ_{23}
	4	2.4	λ, λ_1, λ_2, λ_3	$\lambda_{12}, \lambda_{13}$	λ_{123}, λ_{23}
	5	2.5	λ, λ_1, λ_2, λ_3	$\lambda_{12}, \lambda_{23}$	λ_{123}, λ_{13}
	6	2.6	λ, λ_1, λ_2, λ_3	$\lambda_{13}, \lambda_{23}$	λ_{123}, λ_{12}
3	1	3.1	λ, λ_1, λ_2, λ_3	λ_{123}, λ_{13}, λ_{23}	λ_{12}
	2	3.2	λ, λ_1, λ_2, λ_3	λ_{123}, λ_{12}, λ_{23}	λ_{13}
	3	3.3	λ, λ_1, λ_2, λ_3	λ_{123}, λ_{13}, λ_{12}	λ_{23}
	4	3.4	λ, λ_1, λ_2, λ_3	λ_{12}, λ_{13}, λ_{23}	λ_{123}

Table 9: The number of best fitted unsaturated models without λ_{12}

g	Data generated by	Deviance/ AIC/BIC	SAVE	g	Data generated by	Deviance/ AIC/BIC	SAVE
1	$M1$	100,000	100,000	2	$M1$	100,000	100,000
	$M2$	100,000	100,000		$M2$	100,000	100,000
	$M3$	100,000	100,000		$M3$	97,211	98,658
	$M4$	100,000	100,000		$M4$	94,223	96,230
	$M5$	100,000	100,000		$M5$	56,993	62,540
	$M6$	100,000	100,000		$M6$	99,987	100,000
3	$M1$	51,143	62,773				
	$M2$	55,043	58,533				
	$M3$	46,476	55,344				
	$M4$	40,095	49,677				
	$M5$	27,104	32,107				
	$M6$	100,000	100,000				

Table 10: The number of best fitted unsaturated models including λ_{13}

g	Data generated by	Deviance/ AIC/BIC	SAVE	g	Data generated by	Deviance/ AIC/BIC	SAVE
1	$M1$	0	36,684	2	$M1$	100,000	100,000
	$M2$	100,000	29, 153		$M2$	100,000	100,000
	$M6$	100,000	100,000		$M6$	99,987	100,000
3	$M1$	100,000	100,000				
	$M2$	100,000	100,000				
	$M6$	100,000	100,000				

performed as well as or even better than Deviance, AIC, and BIC. We compared them in terms of the correct identification of parameters of unsaturated log-linear models.

Acknowledgements

We thank a reviewer for the careful reading of an earlier version of our paper, the thoughtful comments, and the editors for their valuable suggestions.

References

Agresti, A. (2013). *Categorical Data Analysis.* Wiley, New York, Third Edition.

Csiszár, I. (1975). i-divergence geometry of probability distributions and minimization problems. *Annals of Probability*, **3(1),** 146-158.

Gokhale, D. V. and Kullback, S. (1978). *The Information in Contingency Tables.* Marcel Dekker, New York.

Haberman, S. J. (1984). Adjustment by minimum discriminant information. *Annals of Statistics,* **12(3),** 971-988.

Klimova, A., Rudas, T., and Dobra, A. (2012). Relational models for contingency tables. *Journal of Multivariate Analysis,* **104,** 159-173.

Klimova, A. and Rudas, T. (2016). On the closure of relational models. *Journal of Multivariate Analysis,* **143** 440-452.

Konishi, S. and Kitagawa, G. (2008). *Information Criteria and Statistical Modeling.* Springer, New York.

Kullback, S., Keegel, J., and Kullback, J. (2013). *Topics in Statistical Information Theory.* Springer, New York.

Kullback, S. (1959). *Information Theory and Statistics.* Wiley, New York. (Dover edition, 1997.)

Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *Annals of Mathematical Statistics,* **22**, 79-86.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, **6**, 461-464.

Rudas, T. (2018). *Lectures on Categorical Data Analysis.* Springer, New York.

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), 431-442

Use of Linear Combination Test to Identify Gene Signatures of Human Embryonic Development in Single Cell RNA-Seq Experiments

I. Dinu[1], E. Khodayari Moez[1], M. Hajihosseini[1], A.P. Leite[2] and S. Pyne[3,4]*
[1]*School of Public Health, University of Alberta, Edmonton, AB, Canada.*
[2]*University College London Cancer Institute, London, UK.*
[3]*Public Health Dynamics Laboratory; and Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA.*
[4]*Health Analytics Network, Pittsburgh, PA, USA.*

Received: 05 March 2021; Revised: 06 April 2021; Accepted: 11 April 2021

Abstract

Data on human systems biology are being generated at a rapid pace due to technological advances in not only high-throughput, but also high-resolution, platforms. Increasing availability of single cell omic data have motivated complex experiments with the intention to gain deeper insights into complex biological systems such as those involved in the development of organisms. Individual, technological and biological sources of heterogeneity of outcomes that are observed among the different populations of cells that are sampled in such experiments require robust analysis. We describe our Linear Combination Test (LCT) methodology, and briefly review its applications to binary, multivariate continuous and longitudinal outcomes in a wide range of omic studies. It allows us to test hypothesis not just about the role of single genes in discrete outcomes, but of large sets of genes in multivariate continuous outcomes, which are representative of dynamic biological phenomena such as embryogenesis, degenerative diseases, *etc*. LCT, which uses a shrinkage covariance matrix estimator, has been shown to be effective at a small computational cost in both simulated omic studies and real-life biomedical applications. In this study, we applied LCT to analyze a new collection of stem cell gene signatures associated with single cell RNA-Seq data measured during human preimplantation embryonic development.

Key words: Linear combination test; Single cell analysis, RNA-Seq; Stem cell gene signatures.

1. Introduction

In early microarray data analysis, individual genes that were differentially expressed across 2 or more classes or conditions were identified using traditional statistical methods such as the t-test, ANOVA, *etc*., Drăghici (2012). Then, the most significant genes were selected based on a predefined threshold and validated for biological patterns. However, given the heterogeneity of gene expression levels, biological interpretation of the results was sensitive to the choice of the threshold, and this subjectivity remains an important concern in such analysis of individual genes. In order to overcome this problem, Gene Set Analysis (GSA) uses existing experimentally obtained knowledge of genes and their pathways to test for significant regulation of sets of multiple genes (called *geneset*s) instead of individual genes. Since the genes within such genesets share a common biological function, considering the correlations

Corresponding Author: Saumyadipta Pyne
E-mail: spyne@pitt.edu, spyne@healthanalytics.net

within each set is a key aspect of a sound GSA method. However, it was shown by Tsai and Chen (2009) that many of the known GSA methods, *e.g*., Wang *et al*. (2011), Dinu *et al*. (2007), Goeman *et al*. (2004), Mansmann and Meister (2005), Kong *et al*. (2006), Subramanian (2005), Efron and Tibshirani (2007), were affected by large type II errors. Another important limitation of many GSA methods is that they can only accommodate binary outcomes, such as disease versus control. Our method, Linear Combination Test (LCT) is a GSA method that was designed to address these limitations by taking into account correlations across genes and outcomes, and dealing with binary, univariate or multivariate continuous outcomes, measured either at a single point in time or at multiple time points, and therefore, allow us to analyze a wider range of studies involving complex study designs (Wang *et al*. 2014).

Single cell omic studies have become increasingly popular over the past decade, due to their powerful ability to profile from a panel of selected few dozen genes up to the entire transcriptome of a large number of individual cells in parallel. A typical example of a basic experiment on single-cell analysis (SCA) of gene expression is described in Figure 1 due to Kalisky *et al*. (2018). SCA involves experiments on individual cells that are typically isolated from a biological sample and then physically separated from each other and assayed upon DNA amplification. For each cell, the expressions of genes (or their products such as proteins) are measured using different well-established (or newly emerging) omic platforms such as RNA sequencing as reviewed by Lafzi (2018), Dal Molin (2019), Kalisky (2018). In the resulting data, the single cells could be considered as samples that are assumed independent and hence possibly affected by different sources of variation among the expression-levels of even the same genes.

The large volume of data measured by single cell omic studies calls for sound statistical and computational analysis methods. Various methods at the individual gene level, have been reviewed by Andrews *et al*. (2021). While most of such methods are focused on differential expression of individual genes between cells representing (generally two) different states, here we reason that an analysis using sets of genes, *i.e*., GSA, has important advantages over the individual gene level analysis when applied to single cell omic studies. The stochasticity of expression levels of the same genes in individual cells could be due to different factors ranging from biological (*e.g*., the cell cycle phase of a particular cell) to technical (*e.g*., missing data). While specific genes may not show consistent expression across single cells, if we consider testing the differential expression of not one gene at a time but multiple genes together that are known to belong to a carefully selected geneset, then our LCT method is more likely than traditional approaches to detect the regulation of a functional process or biological pathway that is significantly associated with the outcomes of a given SCA experiment.

Interestingly, LCT allows multivariate and continuous outcomes that could be more realistic representations of single cell level stochasticity of behaviors than univariate and discrete class labels as used in traditional studies of bulk samples. For such reasons, LCT can provide an overall more robust analytical approach for SCA experiments. In addition, LCT type I error, power and computational efficiency were compared to top GSA methods in simulations and real data analysis studies (Wang *et al.*, 2014). LCT type I error and power were comparable to MANOVA-GSA (Tsai and Chen 2009), and superior to SAM-GS (Dinu *et al*., 2007), especially at higher magnitudes of the correlations values across sets of genes, which is a common scenario in GSA. LCT was superior to both methods in terms of computational efficiency. LCT performed better than GSEA in a simulation study presented by Khodayari *et al*. (2018). However, we would like to point out that GSEA uses information from genes exterior to a pre-defined set or pathway. Based on methodological considerations, Goeman and

Buhlmann (2007) discourage comparing methods involving only genes belonging to the pre-defined set to methods involving genes outside the pre-defined set, as these two categories are conceptually different, and they are testing different hypotheses.

GSA focuses on analysis of biological pathways, or genesets sharing a common biological function. Well-known examples of such collections of pre-defined, often expert-curated, genesets include The Cancer Genome Atlas (TCGA), Tomczak *et al*. (2015); Gene Expression Omnibus (GEO), Edgar *et al*. (2002); Kyoto Encyclopedia of Genes and Genomes (KEGG), Kanehisa *et al*. (2000); BioCarta, Nishimura *et al*. (2001); Molecular Signature Database (MSigDB), Liberzon *et al*. (2015). The use of a carefully selected collection of genesets relevant to the outcomes of interest is a key aspect in GSA. In this study, we compiled a new, large collection of genesets that were reported by several past embryonic stem cell gene expression experiments, and used them to test their association with different stages of early human embryonic development. The remaining of this paper consists of a presentation of the LCT methodology, followed by its application to single cell embryonic genome-wide expression (RNA-Seq) data. We will also discuss various extensions of LCT, including its applications beyond gene expression studies.

2. Data and Methods

Data: In this study, we used 2 types of data. First, for our genesets, we introduce a large collection of 457 curated genesets that were derived from experimentally identified signatures of gene expression in human embryonic stem cells. Hence, we call the collection "stem cell signatures". These genesets were compiled from the Molecular Signature DataBase (MigDB), Liberzon *et al*. (2015); the Differentiation Map portal, Novershtern *et al*. (2011); Ingenuity Pathway Analysis tool (2020); and ChIP-X database, Lachmann *et al*. (2010). We restricted the size of genesets to be between 5 and 500. There are 281 genesets in this range (The full collection of these stem cell signatures is available from the authors upon request). Second, we downloaded the single cell RNA-Seq data from ArrayExpress database, Athar *et al*. (2019); access number E-MTAB-3929 ArrayExpress (2020). The dataset consists of 17855 genes measured in each of 1529 individual cells from 88 human preimplantation embryos. The total of 1529 individual cells is broken down during days 3 to 7 of the embryonic development as follows: 81 cells measured on day 3; 190 on day 4; 377 on day 5; 415 on day 6; and 466 on day 7. During the first 7 days of human development, the zygote undergoes cellular division and establishes the first three distinct cell types of the mature blastocyst: trophectoderm (TE), primitive endoderm (PE), and epiblast (EPI). Petropoulos *et al*. (2016) While the analysis of these data at individual gene level was conducted previously by Petropoulos (2016), in this study, we performed LCT analysis at geneset level of the same data using the above-mentioned stem cell signatures.

Methods: LCT tests if there is a linear relationship between the geneset $X = \{x_1, \dots, x_p\}$ consisting of p genes and a set of q multivariate outcomes $Y = \{Y_1, \dots Y_q\}$. The multivariate null hypothesis can be expressed linearly and univariately as

H_0: There is no association between any linear combination of gene expressions of the members of a geneset X and any linear combination of multivariate outcomes Y.

If $Z(X, A)$ is a linear combination of gene expression measurements within a set of x_is with coefficient vector A and $Z(Y, B)$ is a linear combination of outcomes y_is with coefficient vector B, then we calculate the following statistic to test the null hypothesis

$$T^2 = max|\rho((Z(X,A),Z(Y,B))^2|. \quad (1)$$

The coefficient vectors A and B are estimated in a way that maximizes the Pearson correlation between $Z(X,A)$ and $Z(Y,B)$. T^2 can be rewritten as

$$T^2 = max\frac{\left(A^T Cov(X,Y)B\right)^2}{(A^T Cov(X,X)A).(B^T Cov(Y,Y)B)} = \frac{\left(A^T \Sigma_{XY} B\right)^2}{(A^T \Sigma_{XX} A).(B^T \Sigma_{YY} B)}. \quad (2)$$

In the procedure for estimation of the coefficient vectors, two problems arise: singularity caused by the high dimensionality of data (solved by shrinkage methods) and computational efficiency (solved by eigenvalue decomposition). Then, the *p*-value is calculated using sample permutations. Sample permutation method preserves the correlation structure within geneset and the correlation structure within multivariate outcomes, see Schäfer and Strimmer (2005).

Specifically, the (i,j)th entry of the shrinkage covariance matrix Σ^*_{XX} is given by

$$\sigma^*_{ij} = \gamma_{ij}\sqrt{\sigma_{ii}\sigma_{jj}}$$

with shrinkage coefficients 1 for the diagonal terms, and the off-diagonal terms

$$\gamma_{ij} = \rho_{ij}\, min(\,1, max(\,0, 1-\lambda^*))$$

where ρ_{ij} is the sample correlation between x_i and x_j. The optimal shrinkage intensity can be estimated by

$$\lambda^* = \frac{\sum_{i\neq j} var(\rho_{ij})}{\sum_{i\neq j}\rho^2_{ij}}.$$

Based on this shrinkage strategy, we get the shrinkage version of the test statistic

$$T^{2*} = \max_{A,B}\frac{(A^T \Sigma_{XY} B)^2}{(A^T \Sigma^*_{XX} A)\,.\,(B^T \Sigma^*_{YY} B)}. \quad (3)$$

The computational cost of calculating (3) has to be taken into consideration, since the right-hand side is a nonlinear programming problem involving $p+q$ parameters. The computational cost can be very high for maximizing directly the right-hand side of (3), especially when permutations are used for calculating the p-values of the test. To address the computational efficiency problem, we adopt a strategy of using two groups of normalized orthogonal bases, instead of using the original observation vectors of *X* and *Y*. We perform eigenvalue decompositions for the two shrinkage covariance matrices, $\Sigma^*_{XX} = UD_X U^T$ and $\Sigma^*_{YY} = VD_Y V^T$, and obtain two groups of orthogonal basis vectors $\tilde{X} = (\tilde{x}_1 \ldots, \tilde{x}_p) = (x_1 - \bar{x}_1, \ldots, x_p - \bar{x}_p)UD_X^{-1/2}$ and vectors $\tilde{Y} = (\tilde{y}_1 \ldots, \tilde{y}_q) = (y_1 - \bar{y}_1, \ldots, y_q - \bar{y}_q)VD_Y^{-1/2}$.

The test statistic in (3) can further be rewritten as

$$T^{2*} = \max_{\alpha,\beta}\frac{(\alpha^T \Sigma_{\tilde{X}\tilde{Y}}\beta)^2}{||\alpha||_2^2 \cdot ||\beta||_2^2}, \quad (4)$$

where $\alpha = D_X^{1/2} U^T A$ and $\beta = D_Y^{1/2} V^T B$, and $\Sigma_{\tilde{X}\tilde{Y}}$ is the covariance matrix between $\tilde{X}$ and $\tilde{Y}$, with its (i,j)th entry being $cov(\tilde{x}_i, \tilde{y}_j)$.

The optimization problem in (4) can be solved in two steps. First, for a given β, we find the optimal, α which is proportional to $\Sigma_{\tilde{X}\tilde{Y}}\beta$; second, substitute the optimal α into (4), and find the global optimal, which is proportional to the first eigenvector of the matrix $\Sigma_{\tilde{X}\tilde{Y}}^T \Sigma_{\tilde{X}\tilde{Y}}$ corresponding to the largest eigenvalue. We note that the value of T^{2*} equals to the largest eigenvalue of either the $q \times q$ matrix $\Sigma_{\tilde{X}\tilde{Y}}^T \Sigma_{\tilde{X}\tilde{Y}}$ or the $p \times p$ matrix $\Sigma_{\tilde{Y}\tilde{X}}^T \Sigma_{\tilde{Y}\tilde{X}}$. The cost of obtaining the largest eigenvalue is low, providing $\min(p,q)$ is not large.

The computational advantage is obvious when sample permutations are used to calculate p-values of the test. Since sample permutation changes neither the correlation structure within genesets nor that within the outcomes, we do not need to repeat the same eigenvalue decompositions of the two shrinkage covariance matrices in (3) for the permuted data, but only for the original outcome. That is, after performing the eigenvalue decompositions for the two shrinkage covariance matrices Σ_{XX}^* and Σ_{YY}^* and creating two groups of orthogonal basis vectors $\tilde{X}$ and $\tilde{Y}$, permutations can be done only on $\tilde{Y}$ directly, instead of on the original outcome Y.

For multiple comparisons over large collections of genesets, False Discovery Rate (FDR) is a commonly used method that can provide a better alternative to the more conservative Bonferroni approach. In this study, we used q-value, which is the expected positive FDR, to identify the significantly regulated genesets at different q-value levels (Storey and Tibshirani 2003).

3. Results and Discussion

In this study, we used LCT for testing associations of a new, large collection of curated stem cell signatures with a single-cell RNA-Seq based genome-wide expression dataset on human embryo development. We conducted a quick confirmation of the relevance of these signatures in stem cell gene regulation during human embryo development by applying LCT to single cell data across each pair of consecutive days, from day 3 to day 7. Petropoulos et al. (2016) reported results of an analysis at the individual gene level, across the three distinct cell types of the mature blastocyst. We note that segregation of EPI, PE and TE cell types appears at day 5. The breakdown of sample sizes by day and cell type is as follows: 41 EPI, 32 PE and 142 TE for day 5; 45 EPI, 39 PE and 331 TE for day 6; 41 EPI, 37 PE and 388 TE for day 7.

Differential expression analysis between the EPI cell types and PE cell types performed by Petropoulos *et al*. (2016) identified 43, 1,412, and 542 differentially expressed genes at days 5, 6 and 7 respectively (at FDR ≤ 0.05), with earlier days' (5 and 6) significance being maintained through later days (6 and 7). Our analysis at the geneset level identified 126 differentially regulated stem cell signatures between EPI and PE at day 5 (q-value ≤ 0.001), and a selected subset of 105 genesets that were the most significant (q-value ≤ 0.0001) on days 6 and 7 is shown in Table 1. Regarding the other two cell type pairs, TE versus PE, and TE versus EPI, our analysis at the geneset level indicated more obvious differences compared to EPI versus PE. More importantly, all 281 genesets, which are known stem cell signatures, were found to be significant on each pair of consecutive days from 5 to 7, and for each pair of cell types of TE versus PE, and TE versus EPI, which is in agreement with the individual gene analysis results of Petropoulos et al. (2016).

Finally, we focused on the endothelial transcription factor genes *GATA2* and *GATA3*, which have been previously reported as known markers of TE segregation, and thus, for playing an important role in embryonic development. Ortega *et al*. (2018) We performed LCT using the expressions of these two genes as a bivariate continuous outcome. One hundred and thirty-eight stem cell signatures were significantly associated (q-value ≤ 0.001) with the bivariate continuous phenotype at day 3. For the subsequent days, this count increased to 234, 278 and 281 respectively.

We presented here a new application of LCT methodology to SCA experiments with an illustration on human embryonic development gene expression data. Our approach extends the individual gene analysis to identification of sets of genes that share a common biological function. Such collections of sets exhibit higher reproducibility across studies, and are more robust for addressing complex questions in systems biology. Notably, our new collection of stem cell signatures presented the opportunity to confirm their relevance to the dynamic gene regulation during human embryonic development. As stem cells are an active area of research in biology and medicine, multivariate dynamic outcomes and associated markers (and combinations thereof) can be analyzed by LCT, which can also be extended to testing of newer gene signatures such as those reported by novel experiments to chart a single cell level transcriptional roadmap of human development, *e.g*., Blakeley *et al*. (2015), Durruthy-Durruthy *et al*. (2016).

Unlike the traditional analysis of bulk samples composed of thousands of different cells, experiments that can measure the expressions of selected markers in individual cells are capable of revealing not only the occurrence but also the dynamic states of diverse cell populations, including rare ones, as shown by Pyne *et al*. (2009), Pyne *et al*. (2014), Qi *et al.* (2020), *etc*. In order to characterize the cellular heterogeneity of a given sample (say, a tumor) with precision, the experimenter will need to select the corresponding panels of marker genes. Different choices of markers are given by genesets which must be compiled carefully as we have done for characterizing embryonic stem cell signatures in this study.

LCT is a powerful correlation-based test that can be used to explore thousands of genesets in an automated yet computationally efficient manner. However, we note that the linear combinations identified by LCT are not unique, and there is no direct interpretation of the linear combination coefficients that one can achieve through classical linear regression techniques. To address this, statistical models for high-dimensional analysis can be applied *post hoc* to the identified sets of genes, based on their LCT significance, and selected for their biological relevance. Such methods may provide further interpretation and insight into the selected sets. LCT was extended to longitudinal multivariate outcomes by Khodayari *et al*. (2019). LCT has also proven to be effective beyond gene expression data. Analogous to genesets, we have used LCT on collections of metabolite-sets to test for associations between oncogenic outcomes and high-throughput metabolomic data from prostate cancer patients in Khodayari *et al.* (2018).

References

Andrews, T. S., Kiselev, V. Y., McCarthy, D. and Hemberg, M. (2021). Tutorial: guidelines for the computational analysis of single-cell RNA sequencing data. *Nature Protocols*, **16**, 1-9.

ArrayExpress. (2020). https://www.ebi.ac.uk/arrayexpress/.

Athar, A., Füllgrabe, A., George, N., Iqbal, H., Huerta, L., Ali, A., Snow, C., Fonseca, N. A., Petryszak, R., Paptheodorou, I., Sarkans, U. and Brazma, A. (2019). ArrayExpress update - from bulk to single-cell expression data. *Nucleic Acids Research*, **47**(**D1**), D711-D715.

Blakeley, P., Fogarty, N. M., del Valle, I., Wamaitha, S. E., Hu, T. X., Elder, K., Snell, P., Christie, L., Robson, P. and Niakan, K. K. (2015). Defining the three cell lineages of the human blastocyst by single-cell RNA-Seq. *Development*, **142**(**18**), 3151-3165.

Dal Molin, A. and Di Camillo, B. How to design a single-cell RNA-sequencing experiment: pitfalls, challenges and perspectives. (2019). *Brief Bioinformatics*, **20**(**4**),1384-1394.

Dinu, I., Potter, J. D., Mueller, T., Liu, Q., Adewale, A. J., Jhangri G. S., Einecki, G., Famulski, K. S., Halloran, P., Yasui, Y. (2007). Improving gene set analysis of microarray data by SAM-GS. *BMC Bioinformatics*, **8**, 242

Drăghici, S. (2012). *Statistics and Data Analysis for Microarrays Using R and Bioconductor. 2nd Edition*. Chapman & Hall/CRC Mathematical and Computational Biology Series (ISSN: 978-1439809754).

Durruthy-Durruthy, J., Wossidlo, M., Pai, S., Takahashi, Y., Kang, G., Omberg, L., Chen, B., Nakauchi, H., Reijo Pera, R. and Sebastiano, V. (2016). Spatiotemporal reconstruction of the human blastocyst by single-cell gene-expression analysis informs induction of Naive Pluripotency. *Development Cell*, **38**(**1**), 100-115.

Edgar, R., Domrachev, M. and Lash, A. E. (2002). Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Research*, **30**, 207–210.

Efron, B. and Tibshirani, R. (2007). On testing the significance of sets of genes. *Annals of Applied* Statistics, **1**, 107–129.

Goeman, J. J. and Buhlmann, P. (2007). Analyzing gene expression data in terms of gene sets: methodological issues. *Bioinformatics*, **23**, 980-987.

Goeman J. J., van de Geer S. A., de Kort, F. and van Houwelingen, H. C. (2004). A global test for groups of genes: testing association with a clinical outcome. *Bioinformatics*, **20,** 93–99.

Ingenuity Pathway Analysis Tool. (2020). http://www.ingenuity.com.

Kalisky, T., Oriel, S., Bar-Lev, T. H., Ben-Haim, N., Trink, A., Wineberg, Y., Kanter, I., Gilad, S. and Pyne, S. (2018). A brief review of single-cell transcriptomic technologies. *Brief Functional Genomics*, **17**(**1**), 64-76.

Kanehisa, M. and Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, **28**, 27–30.

Khodayari, M. E., Hajihosseini, M., Andrews, J. L. and Dinu, I. (2019). Longitudinal linear combination test for gene set analysis. *BMC Bioinformatics*, **20**(**1**), 650.

Khodayari, M. E., Pyne, S. and Dinu, I. (2018). Association between bivariate expression of key oncogenes and metabolic phenotypes of patients with prostate cancer. *Computers in Biology and Medicine*, **103**, 55-63.

Kong, S. W., Pu, W. T. and Park, P. J. (2006). A multivariate approach for integrating genome wide expression data and biological knowledge. *Bioinformatics*, **22,** 2373-2380.

Lachmann, A., Xu, H., Krishnan, J., Berger, S. I., Mazloom, A. R. and Ma'ayan, A. (2010). ChEA: transcription factor regulation inferred from integrating genome-wide ChIP-X experiments. *Bioinformatics*, **26**(**19**), 2438-2444.

Lafzi, A., Moutinho, C., Picelli, S. and Heyn, H. (2018). Tutorial: guidelines for the experimental design of single-cell RNA sequencing studies. *Natural Protocols*, **13**(**12**), 2742-2757.

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J. P. and Tamayo, P. (2015). The molecular signatures database (MSigDB) hallmark gene set collection. *Cell Systems*, **1**(**6**), 417-425.

Mansmann, U. and Meister, R. (2005). Testing differential gene expression in functional groups. *Methods of Information in Medicine*, **44**, 449-53.

Nishimura, D. (2001). BioCarta. *Biotech Software and Internet Report*, **2**, 117–120.

Novershtern, N., Subramanian, A., Lawton, L. N., Mak, R. H., Haining, W. N., McConkey, M. E., Habib, N., Yosef, N., Chang, C. Y., Shay, T., Frampton, G. M., Drake, A. C., Leskov, I., Nilsson, B., Preffer, F., Dombkowski, D., Evans, J. W., Liefeld, T., Smutko, J. S., Chen, J., Friedman, N., Young, R. A., Golub, T. R., Regev, A. and Ebert, B. L. (2011). Densely interconnected transcriptional circuits control cell states in human hematopoiesis. *Cell*, **144**(**2**), 296-309.

Ortega, M. N., Winblad, N., Reyes, A. P., Lanner, F. (2018). Functional genetics of early human development. *Current Opinion in Genetics and Development*, **52**, 1-6.

Petropoulos, S., Edsgard, D.,, Reinius, B., Linnarsson, S., Sandberg, R. and Lanner, F. (2016). Single-cell RNA-Seq reveals lineage and X chromosome dynamics in human preimplantation embryos. *Cell*, **165**, 1012-1026.

Pyne, S., Hu, X., Wang, K., Rossin, E., Lin, T-I., Maier, L. M., Beacher-A. C., *et al*. (2009). Automated high-dimensional flow cytometric data analysis. *Proceedings of the National Academy of Sciences USA*, **106**(**21**), 8519-8524.

Pyne, S., Lee, S. X., Wang, K., Irish, J., Tamayo, P., Nazaire, M-D., Duong, T., Ng, S-K., Hafler, D., *et al*. (2014). Joint modeling and registration of cell populations in cohorts of high-dimensional flow cytometric data. *PLoS ONE*, **9**(**7**), e100334.

Qi, Y., Fang, Y., Sinclair, D., Guo, S., Alberich-Jorda, M., Lu, J., Tenen, D., Kharas, M. and Pyne, S. (2020). High-speed automatic characterization of rare events in flow cytometric data. *PLoS ONE*, **15**(**2**): e0228651.

Schäfer, J. and Strimmer, K. (2005). A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Statistical Applications in Genetics and Molecular Biology*, **4**, 32.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy S. L., Golub T. R., Lander, E. S. and Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, **102**, 15545–15550.

Storey, J. D. and Tibshirani, R. (2003). Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences*, **100**, 9440–9445.

Tomczak, K., Czerwińska, P. and Wiznerowicz, M. (2015). The cancer genome Atlas (TCGA): an immeasurable source of knowledge. *Contemporary Oncology*, **19**, A68–A77.

Tsai, C. and Chen, J. J. (2009). Multivariate analysis of variance test for geneset analysis. *Bioinformatics.* **25**(**7**), 897-903.

Wang, X., Dinu, I., Liu, W. and Yasui, Y. (2011). Linear combination test for Hierarchical Gene Set Analysis. *Statistical Applications in Genetics and Molecular Biology*, **10**(**1**), Article 13.

Wang, X., Pyne, S. and Dinu, I. (2014). Gene set enrichment analysis for multiple continuous phenotypes. *BMC Bioinformatics*, **15**, 260.

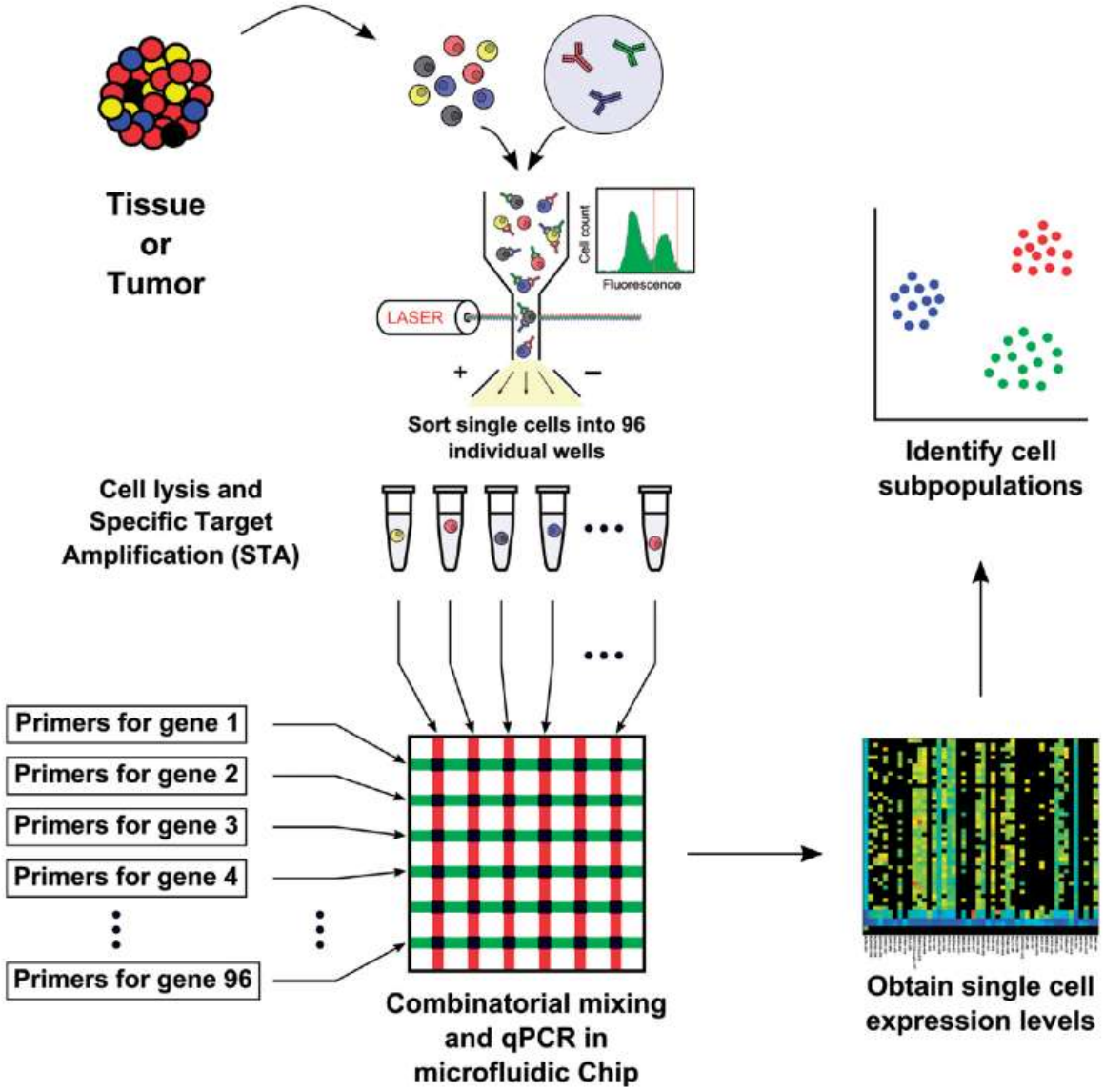


Figure 1: A schematic diagram showing a basic SCA experiment. Single-cell gene expression measurement using qPCR workflow is performed with the Fluidigm Dynamic Array microfluidic chip. Reproduced from Kalisky *et al*. (2018) with permission.

Table 1: Stem cell gene signatures differentially regulated between 41 epiblast and 32 primitive endoderm lineages at day 5 in human embryo development (shown in the increasing order of size)

Geneset	Size	*q*-value
StemCell_Lian07_20genes_17053208.table1a	20	<0.0001
StemCell_Shim04_23genes_15246160.table4	22	<0.0001
IPA_affects.epithelial.mesenchymal.transition.of.cells	22	<0.0001
IPA_increases.epithelial.mesenchymal.transition.of.cells	24	<0.0001
StemCell_Kocer08_44genes_18667080.TableS3	26	<0.0001
StemCell_Matushansky08_35genes_18310505.TableS6	28	<0.0001
StemCell_Lottaz10_30genes_20145155.Table1	29	<0.0001
IPA_increases.differentiation.of.embryonic.stem.cells	29	<0.0001
StemCell_Lim08_35genes_18510698.Table3	34	<0.0001
Ben.Porath_ES_2	39	<0.0001
StemCell_Seo07_61genes_18034892.Table1	40	<0.0001
DMAP_TCELLA3_DN	40	<0.0001
Marson_H3K4me3	41	<0.0001
DMAP_PRE_BCELL2_UP	42	<0.0001
DMAP_PRE_BCELL3_DN	42	<0.0001
IPA_affects.differentiation.of.embryonic.stem.cells	43	<0.0001
DMAP_EOS_DN	44	<0.0001
DMAP_BCELLA3_DN	44	<0.0001
DMAP_CMP_DN	44	<0.0001
DMAP_NKA2_DN	44	<0.0001
DMAP_TCELL_DN	45	<0.0001
DMAP_GRAN3_UP	45	<0.0001
DMAP_MEGA2_DN	45	<0.0001
DMAP_NKA3_UP	45	<0.0001
DMAP_ERY1_DN	46	<0.0001
DMAP_ERY2_UP	46	<0.0001
DMAP_HSC3_DN	46	<0.0001
DMAP_BCELLA1_DN	47	<0.0001
DMAP_TCELLA2_UP	47	<0.0001
DMAP_TCELLA6_DN	47	<0.0001
StemCell_Lim08_50genes_18510698.Table1	48	<0.0001
DMAP_MYP_UP	48	<0.0001
DMAP_BCELLA2_UP	49	<0.0001
StemCell_Duhagon10_60genes_20500816.Table1	56	<0.0001
IPA_affects.differentiation.of.hematopoietic.progenitor.cells	56	<0.0001
IPA_affects.differentiation.of.hematopoietic.cells	62	<0.0001
IPA_increases.differentiation.of.stem.cells	68	<0.0001
StemCell_Kocer08_87genes_18667080.TableS6	71	<0.0001
IPA_affects.differentiation.of.stem.cells	73	<0.0001

StemCell_Hao09_97genes_20077526.TableS7	74	<0.0001
Kim_CORE	74	<0.0001
StemCell_Almstrup04_138genes_15256440.tableS1	80	<0.0001
DB_KLF2.18264089	83	<0.0001
DB_KLF4.18264089	83	<0.0001
DB_KLF5.18264089	83	<0.0001
Taube.et.al._EMT_upregulated_genes	86	<0.0001
StemCell_Colombo09_111genes_19123479.TableS1	90	<0.0001
IPA_affects.differentiation.of.bone.marrow.cells	90	<0.0001
DB_TRP63.18441228	94	<0.0001
Ben.Porath_ES_CORE_NINE_CORRELATED	99	<0.0001
StemCell_Korkola05_146genes_15870693.SuppTable1	100	<0.0001
StemCell_Bohgaki05_118genes_16014681.table2	113	<0.0001
DB_NOTCH1.17114293	121	<0.0001
StemCell_Hao09_173genes_20077526.TableS5	126	<0.0001
DB_HOXD13.18407260	130	<0.0001
StemCell_Kocer08_185genes_18667080.TableS4	154	<0.0001
Ben.Porath_NOS_TARGETS	168	<0.0001
DB_IRF1.19129219	173	<0.0001
DB_TP63.19390658	176	<0.0001
DB_ESR1.20079471	187	<0.0001
DB_PPARG.19300518	187	<0.0001
StemCell_Kocer08_236genes_18667080.TableS8	194	<0.0001
DB_VDR.20736230	196	<0.0001
DB_WT1.19549856	197	<0.0001
DB_SCL.19346495	206	<0.0001
Kim_GCN5L2	211	<0.0001
Ben.Porath_MYC_TARGETS_WITH_EBOX	222	<0.0001
StemCell_Hao09_359genes_20077526.TableS4	239	<0.0001
DB_EGR1.19032775	242	<0.0001
DB_CDX2.19796622	253	<0.0001
Kim_CTR9	258	<0.0001
StemCell_Matushansky08_297genes_18310505.TableS8	261	<0.0001
DB_ZIC3.20872845	266	<0.0001
Ben.Porath_OCT4_TARGETS	272	<0.0001
StemCell_Hao09_612genes_20077526.TableS3	302	<0.0001
StemCell_Matushansky08_886genes_18310505.TableS1	306	<0.0001
StemCell_Bhattacharya05_2471genes_16207381.Table1Sb	308	<0.0001
StemCell_Kocer08_575genes_18667080.TableS9	309	<0.0001
StemCell_Kocer08_864genes_18667080.TableS2	310	<0.0001
StemCell_Bhattacharya05_2843genes_16207381.Table1Sa	310	<0.0001
StemCell_Majeti09_3024genes_19218430.TableS3	313	<0.0001
StemCell_Matushansky08_1453genes_18310505.TableS7	315	<0.0001
DB_POU5F1.18700969	322	<0.0001

DB_RARG.19884340	326	<0.0001
Zola07_426genes_CellDifferentiationMarkers_17174972.TableS1	332	<0.0001
StemCell_Qi03_534genes_12631704.table1	334	<0.0001
StemCell_Hassan09_1544genes_19808871.TableS3	340	<0.0001
DB_ESR1.17901129	347	<0.0001
DB_TP53.16413492	349	<0.0001
DB_HTT.18923047	355	<0.0001
Ben.Porath_ES_1	358	<0.0001
DB_SMAD4.19686287	367	<0.0001
DB_STAT6.20620947	369	<0.0001
DB_SOX2.18555785	388	<0.0001
DB_CLOCK.20551151	399	<0.0001
DB_NANOG.18555785	412	<0.0001
DB_CTNNB1.20615089	416	<0.0001
DB_TCF4.18268006	420	<0.0001
DB_POU5F1.18555785	438	<0.0001
DB_ZFP281.18358816	441	<0.0001
Kim_PRC	444	<0.0001
DB_CDX2.20551321	446	<0.0001
Kim_ZFP281	461	<0.0001
DB_SMAD1.18555785	465	<0.0001
DB_PDX1.19855005	493	<0.0001

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 443–451

On a Process of Rumour Propagation

Farkhondeh Alsadat Sajadi[1] **and Rahul Roy**[2]
[1] *Department of Statistics, Faculty of Mathematics and Statistics*
University of Isfahan, Iran
[2] *Theoretical Statistics and Mathematics Unit*
Indian Statistical Institute, Delhi

Received: 20 March 2021; Revised: 16 April 2021; Accepted: 19 April 2021

Abstract

In recent years there has been a vast amount of work to model the spread of rumour. Here we review some of these mathematical models and present some of the main results.

Key words: Rumour process; Firework processes; Renewal processes; Double coverage.

AMS Subject Classifications: 60K35

1. Introduction

The Oxford Dictionary defines rumour as 'a statement or report circulating in a community, of the truth of which there is no evidence'. Mathematically, Gilbert (1961) used the Poisson Boolean model and Maki and Thompson (1973) used a slight variant of this model to study the transmission of information/rumour. This model consisted of a signal being transmitted through a relay of transmitters to its recipient. Two such versions are the Poisson Boolean model and the rumour processes. We present a brief description of these processes here.

Poisson Boolean model: Let $\Xi := (\xi_1, \xi_2, \ldots)$ on $\mathbb{R}^d$ be a homogeneous Poisson point process of intensity λ and $\{\rho_1, \rho_2, \ldots\}$ an independent collection of i.i.d. positive real valued random variables. This is the Poisson Boolean model and its covered region is defined to be the random set $C := \cup_{i=1}^{\infty} B(\xi_i, \rho_i)$, where $B(\xi, \rho)$ is the closed ball centred at ξ and of radius ρ in the Euclidean norm. Geometric properties of this Boolean model has been studied by Matheron (1968), Hall (1988) and Chiu, *et al.* (2013). Kertesz and Vicsek (1982) used this model to study a continuum version of percolation whose parameter is the intensity λ with the radius random variables $\rho_1, \rho_2, \ldots$ being either constants or of a fixed distribution (see Meester and Roy (1996) and Penrose (2003) for a review of the percolation properties of this model). Gupta and Kumar (1998) used this model to study questions of signal-to-interference-ratio (SINR) and other such problems in wireless transmission, see Francisccetti and Meester (2007) for a review.

Rumour process: Sudbury (1985) studied the variant of the information-transmission model introduced in Maki and Thompson (1973). Subsequently, Junior, *et al.*(2011) renamed

Corresponding Author: Rahul Roy
Email: rahul@isid.ac.in

the rumour process as the 'firework process' and introduced a different variant the 'reverse firework process'.

Firework process: Let $\{R_i : i \geq 0\}$ be a sequence of non-negative integer valued i.i.d. random variables . At time 0 the origin starts a rumour and passes it onto all individuals in the interval $[0, R_0]$. At time t, all individuals who heard the rumour for the first time at time $t-1$ spread the rumour, with the individual at site j spreading it among all individuals in the region $[j, j+R_j]$. Note that allowing $\mathbb{P}\{R_j = 0\} > 0$ ensures that there are individuals who are inactive.

Reverse firework process: The reverse firework process consists of the origin who knows the rumour at time 0, and at time t an individual located at site j listens to individuals in the interval $[j - R_j, j]$. If there is an individual at a site in this interval who has heard the rumour by time $t-1$, then the individual at site j gets to know the rumour. Here the random variables $\{R_i : i \geq 0\}$ are as in the firework process.

1.1. Definitions

For each individual at site $i \in \mathbb{N}$ associate the pair (X_i, ρ_i) where $(X_i)_{i\geq 1}$, is a sequence of Bernoulli (p) random variables, i.e.,

$$X_{\mathbf{i}} = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1-p. \end{cases} \tag{1}$$

and $(\rho_i)_{i\geq 1}$ a sequence of i.i.d. copies of some $\mathbb{N}$-valued random, independent of the random variables $(X_i)_{i\geq 1}$. Let ρ denote a generic random variable with the same distribution as ρ_i. In addition, let ρ_0 an independent $\mathbb{N}$ valued random variables, independent of the collections $(X_i)_{i\geq 1}$ and $(\rho_i)_{i\geq 1}$, with ρ_0 having the same distribution as ρ. Whenever $X_i = 1$, the individual at site i starts to spread rumour within a random distance to its right (an interval of length ρ_i). Coverage occurs if every site of $\mathbb{N}$ is covered by some interval. Set $X_0 \equiv 1$ and let

$$C := \bigcup_{\{i\geq 0 : X_i = 1\}} [i, i+\rho_i],$$

and

$$\begin{aligned} D := \{x \in \mathbb{R} : &\text{there exist } j,\ k \geq -1 \text{ with } j \neq k,\ X_j = X_k = 1 \\ &\text{and } x \in ([j, j+\rho_j] \cap [k, k+\rho_k])\}. \end{aligned}$$

We say that $\mathbb{N}$ is *eventually covered* by C if there exists a $t \geq 1$ such that $[t, \infty) \subseteq C$. We say that $\mathbb{N}$ is *eventually doubly covered* by D if it contains a region $[t, \infty)$, for some $t \geq 1$.

Let $\{X_{\mathbf{i}} : \mathbf{i} \in \mathbb{N}^d\}$ be a collection of Bernoulli (p) random variables and $\{\rho_{\mathbf{i}} : \mathbf{i} \in \mathbb{N}^d\}$ a collection of i.i.d. $\mathbb{N}$ valued random variables, independent of the collection $\{X_{\mathbf{i}} : \mathbf{i} \in \mathbb{N}^d\}$. Let ρ denote a generic random variable with the same distribution as $\rho_{\mathbf{i}}$ and

$$\mathbf{C} := \cup_{\{\mathbf{i} : X_{\mathbf{i}} = 1\}} (\mathbf{i} + [0, \rho_{\mathbf{i}}]^d)$$

denote the *covered region* of $\mathbb{N}^d$; here and subsequently $\mathbf{i}+[0,\rho_{\mathbf{i}}]^d = [i_1, i_1+\rho_{\mathbf{i}}]\times\cdots\times[i_d, i_d+\rho_{\mathbf{i}}]$, where $\mathbf{i} = (i_1, \ldots, i_d)$. We say that $\mathbb{N}^d$ is *eventually covered* if there exists $\mathbf{t} \in \mathbb{N}^d$ such that $\mathbf{t} + \mathbb{N}^d \subseteq \mathbf{C}$. Note that this definition may be seen to be equivalent to percolation of the homogenous firework process for $d = 1$, and in that sense, it extends the definition of percolation for a homogenous firework process in $\mathbb{N}^d$. We say that $\mathbb{N}^d$ is *eventually doubly covered* if there exists $\mathbf{t} \in \mathbb{N}^d$ such that $\mathbf{t} + \mathbb{N}^d \subseteq \mathbf{D}$, where

$$\mathbf{D} := \{\mathbf{x} \in \mathbb{R}^d : \text{there exist } \mathbf{i},\ \mathbf{j} \in \mathbb{N}^d \text{ with } \mathbf{i} \neq \mathbf{j} \text{ and } X_{\mathbf{i}} = X_{\mathbf{j}} = 1 \text{ such that } x \in (\mathbf{i} + [0,\rho_{\mathbf{i}}]^d) \cap (\mathbf{j} + [0,\rho_{\mathbf{j}}]^d)\}.$$

The probability of coverage in terms of stochastic geometry or probability of survival for the original rumour process depends on both, the marginal distribution of the radius of influence ρ, and the joint distribution of the X_i's. There are three types of scenarios for random variables X_i's for which we have a necessary and sufficient condition to guarantee a positive probability of survival of the rumour:

(1): X_i's are i.i.d. random variables.

(2): X_i's are a $\{0,1\}$-valued Markov chain.

(3): X_i's are a one-dimensional undelayed discrete renewal point process.

2. The i.i.d Case

Suppose $\{X_{\mathbf{i}} : \mathbf{i} \in \mathbb{N}^d\}$ is a collection of $\{0,1\}$-valued i.i.d. random variable with $p = \mathbb{P}(X_i = 1)$. We assume that this collection is independent of the the collection of i.i.d. positive integer-valued random variables $\{\rho_{\mathbf{i}} : \mathbf{i} \in \mathbb{N}^d\}$. Let $\mathbb{P}_p$ denote the product probability law of X and ρ. When the individuals are not sceptical we have:

Proposition 1: (Athreya, *et al.* (2004))

(i): For $d = 1$

$$\mathbb{P}_p(C \text{ eventually covers } \mathbb{N}) = \begin{cases} 1 & \text{if } p > 1/l \\ 0 & \text{if } p < 1/L. \end{cases}$$

where

$$l := \liminf_{j\to\infty} j\mathbb{P}(\rho > j) > 1 \text{ and } L := \limsup_{j\to\infty} j\mathbb{P}(\rho > j) < \infty.$$

(ii): For $d > 1$ and $0 < p < 1$, we have

$$\mathbb{P}_p(\mathbf{C} \text{ eventually covers } \mathbb{N}^d) = \begin{cases} 1 & \text{if } \liminf_{j\to\infty} j\mathbb{P}(\rho > j) > 0 \\ 0 & \text{if } \lim_{j\to\infty} j\mathbb{P}(\rho > j) = 0. \end{cases}$$

A priori it may be the case that 'single coverage' occurs, i.e. $\mathbf{C} \supseteq \mathbf{t} + \mathbb{N}^d$, but double coverage does not occur. Equivalently, in terms of the rumour process, a rumour may have a positive probability of spreading in a population consisting of only disbelievers or

gullible persons. While if among the gullible persons there is also a further group who are sceptics, then the rumour may not spread with positive probability. However, the following proposition shows that this is not the case:

Proposition 2: (Sajadi and Roy (2019))

(i): For $d = 1$,

$$\mathbb{P}_p(D \text{ eventually covers } \mathbb{N}) = \begin{cases} 1 & \text{if } p > 1/l \\ 0 & \text{if } p < 1/L. \end{cases}$$

where

$$l := \liminf_{j\to\infty} j\mathbb{P}(\rho \geq j) > 1 \text{ and } L := \limsup_{j\to\infty} j\mathbb{P}(\rho \geq j) < \infty.$$

(ii): For $d > 1$ and $p > 0$, we have

$$\mathbb{P}_p(\mathbf{D} \text{ eventually covers } \mathbb{N}^d) = \begin{cases} 1 & \text{if } \liminf_{j\to\infty} j\mathbb{P}(\rho \geq j) > 0 \\ 0 & \text{if } \lim_{j\to\infty} j\mathbb{P}(\rho \geq j) = 0. \end{cases}$$

The key to the proof of Proposition 1 (i) is to note that, for $d = 1$, the coverage process forms a renewal process, with renewal happening at every site $i \in \mathbb{N}$ such that $i \notin C$. Part (ii) of the above two propositions exhibits a dichotomy in the behaviour of the process in dimension 1 and in dimensions 2 or more. If $\mathbb{P}(\rho \geq j) = O(j)$ as $j \to \infty$, then in dimension 1, depending on the value of p, there may not be coverage or double coverage, with probability 1. However, for dimensions 2 or more, the only case when there is no coverage (and hence no double coverage) with probability 1 when $p = 0$, i.e. there are no gullible people in the population.

In particular, for $p > 0$ and $i \geq 1$, let

$$A_i := \{i \notin C\} \text{ and } B_i := \{i \notin D\}.$$

Taking $G(i) = \mathbb{P}(\rho \geq i)$ and $g_p(i) = 1 - pG(i)$, we observe that

$$\begin{aligned}
&\mathbb{P}_p(B_i) \\
&\quad = \mathbb{P}\left(A_i \cup \{(\text{there exists exactly one } j \text{ with } X_j = 1 \text{ such that } i \in [j, j + \rho_j]\}\right) \\
&\quad = \mathbb{P}_p(A_i) + \sum_{l=0}^{i-1} \mathbb{P}_p(X_{i-l} = 1,\ i \leq i - l + \rho_{i-l},\ i \notin \cup_{\{j\neq i-l, X_j=1\}}[j, j + \rho_j]) \\
&\quad = \mathbb{P}_p(A_i) + p\prod_{l=1}^{i-1} g_p(l) + p(1-p)\sum_{k=1}^{i-1} G(k) \prod_{l\neq k, l=1}^{i-1} g_p(l). \qquad (2)
\end{aligned}$$

We next show that for $p > 1/l$, where l is as in Proposition 1

$$\sum_i \mathbb{P}_p(B_i) < \infty, \qquad (3)$$

which, by Borel-Cantelli lemma yields $\mathbb{P}_p(B_i \text{ occurs finitely often}) = 1$, i.e. there is a random variable T, with $T < \infty$ almost surely, such that $\mathbb{P}_p\{D \supseteq [T, \infty)\} = 1$.

Also, for $p < 1/L$, where L is as in Proposition 1 we have

$$\sum_i \mathbb{P}_p(A_i) = \infty. \tag{4}$$

However, since A_i's are not independent events so we cannot apply the converse of the Borel–Cantelli lemma. Our observation A_i's are renewal events (in the sense that, for $k > i$, $P_p(A_k \cap A_i) = P_p(A_{k-i})P_p(A_i)$) allows us to use Theorem 3, on page 312 of Feller (1971) to conclude that (4) implies that A_i occurs for infinitely many i's. Thus single coverage (and hence double coverage) does not occur almost surely

3. The Markovian Case

Suppose $(X_i)_{i\geq 1}$ is a $\{0,1\}$-valued time-homogeneous Markov chain with $p_{ij} = \mathbb{P}(X_{n+1} = j|X_n = i)$, for $i, j \in \{0,1\}$ and $n \geq 0$. Also suppose $(\rho_i)_{i\geq 1}$ is an independent and identically distributed sequence of random variables assuming values on $\mathbb{N}$, independent of the Markov chain. Let $l := \liminf_{j\to\infty} j\mathbb{P}(\rho > j) > 1$ and $L := \limsup_{j\to\infty} j\mathbb{P}(\rho > j) < \infty$. We have

Theorem 1: (Athreya, *et al.* (2004)) For $0 < p_{00}, p_{10} < 1$, we have

$$\mathbb{P}(C \text{ eventually covers } \mathbb{N}) \begin{cases} > 0 & \text{if } \frac{p_{01}}{p_{10}+p_{01}} > 1/l \\ = 0 & \text{if } \frac{p_{01}}{p_{10}+p_{01}} < 1/L. \end{cases}$$

Theorem 2: (Esmaeeli and Sajadi (2020)) For $0 < p_{00}, p_{10} < 1$, we have

$$\mathbb{P}(D \text{ eventually covers } \mathbb{N}) \begin{cases} > 0 & \text{if } \frac{p_{01}}{p_{10}+p_{01}} > 1/l \\ = 0 & \text{if } \frac{p_{01}}{p_{10}+p_{01}} < 1/L. \end{cases}$$

The proofs of the above two theorems require intricate analysis using probability generating functions. In particular, for $k \geq 1$ and the event A_i as before, let $\mathbb{P}_0(A_k) = \mathbb{P}(A_k|X_1 = 0)$ and $\mathbb{P}_1(A_k) = \mathbb{P}(A_k|X_1 = 1)$. We observe that

$$\mathbb{P}_0(A_{k+1}) = p_{00}\mathbb{P}_0(A_k) + p_{01}\mathbb{P}_1(A_k),\ \mathbb{P}_1(A_{k+1}) = \mathbb{P}(\rho_0 \leq k-1)[p_{10}\mathbb{P}_0(A_k) + p_{11}\mathbb{P}_1(A_k)]. \tag{5}$$

Taking k_0 such that $k_0 + (1 - L) > 0$, and considering the functions $A(s) := \sum_{k\geq k_0} \mathbb{P}_0(A_k)s^k$ and $B(s) := \sum_{k\geq k_0} \mathbb{P}_1(A_k)s^k$ we show that

$$A(1) = \sum_{k\geq k_0} \mathbb{P}_0(A_k) = B(1) = \sum_{k\geq k_0} \mathbb{P}_1(A_k) = \infty.$$

This along with the observation that A_k's are delayed renewal events allow us to use the theorem in Feller (1971). The proof of the double coverage case involves the same ideas, except that the relation (5) is more complicated and as such the calculations are more intricate.

4. The Renewal Case

Let $(T_i)_{i\geq 1}$ be a sequence of independent copies of some $\mathbb{N}$-valued random variable T. Taking $X_0 = 1$, define the $\{0,1\}$-valued random variables $\mathbf{X} = (X_i)_{i\geq 1}$ as follows:

$$X_{\mathbf{i}} = \begin{cases} 1 & \text{if and only if there exists } n \geq 1 \text{ such that } \sum_{k=1}^{n} T_k = i \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

Observe that $\mathbf{X}$ is a binary undelayed renewal sequence with inter-arrival times T. Let $(\rho_i)_{i\in\mathbb{N}}$ be a sequence of independent copies of some $\mathbb{N}$-valued random variable ρ, independent of the sequence $\mathbf{X}$ and satisfying $\mathbb{P}(\rho = 0) > 0$. We say that there is coverage if the event

$$\mathcal{A} := \Big(\bigcup_{\{i\geq 0: X_i = 1 \& \rho_i \geq 1\}} [i+1, i+\rho_i] = \mathbb{N} \Big),$$

occurs. The main objective is to study $\mathbb{P}(\mathcal{A})$. We start with conditions under which this probability is null.

Proposition 3: (Gallo and Garcia (2018)) If $\mathbb{E}[T] = \infty$, then $\mathbb{P}(\mathcal{A}) = 0$.

Also

Proposition 4: (Gallo and Garcia (2018)) If $\limsup_{n\to\infty} \dfrac{n\mathbb{P}(\rho > n)}{\mathbb{E}[T]} < 1$, then $\mathbb{P}(\mathcal{A}) = 0$.

But the case where $\mathbb{E}[T] = \infty$, is not interesting and usually it assumes that the process $(X_i)_{i\geq 1})$ is positive recurrent. If we further assume that $\mathbf{X}$ is aperiodic, we have the following formula from Gallo and Garcia (2018), for the probability of coverage:

Theorem 3: (Gallo and Garcia (2018))

$$\mathbb{P}(\mathcal{A}) = \left(1 + \sum_{n\geq 1} \mathbb{E} \prod_{i=0}^{n-1} [\mathbb{P}(\rho \leq i)]^{X_{i+1}} \right)^{-1}.$$

This above quantity is difficult to handle in general and Gallo and Garcia (2018) presented explicit bounds for the probability of coverage.

To guarantee positive probability of coverage, Gallo and Garcia (2018) needed an extra assumption.

Proposition 5: (Gallo and Garcia (2018)) Let $q_i := \max_{n\leq i} \mathbb{P}(T \geq n+2 | T \geq n+1)$, $i \geq 0$. If

$$\sum_{j=1}^{k} \prod_{i=k}^{k+j-2} q_i = o(k) \tag{7}$$

and $\limsup_{n\to\infty} \dfrac{n\mathbb{P}(\rho > n)}{\mathbb{E}[T]} > 1$ then $\mathbb{P}(\mathcal{A}) > 0$.

If the rumour process satisfies condition (7), then by Propositions 3 and 4, we have a sharp phase transition between null and positive probability of coverage.

Gallo and Garcia (2018) conjectured that if the renewal process is positive recurrent, then $\liminf_{n\to\infty} \dfrac{n\mathbb{P}(\rho > n)}{\mathbb{E}[T]} > 1$ should guarantee that $\mathbb{P}(\mathcal{A}) > 0$.

5. Fireworks Version of Rumour Processes

To study rumour propagation in terms of the fireworks process, the main goal is to find out the probability of having an infinite set of individuals knowing the rumour is positive. Junior, *et al.* (2011) presented the survival event as a limit of an increasing sequence of events whose probability can be bounded by a use of the FKG inequality. To find conditions under which the process dies out, they used a non-standard version of the Borel-Cantelli lemma. Gallo, *et al.* (2014) used a technique based on the relationship between the rumour process and a certain discrete time renewal process to obtain more precise results for the homogeneous versions of the fireworks process.

Suppose at time 0, the origin spreads a rumour to all individuals in the interval $[0, \rho_0]$. At time t all individuals, who received the rumour at time $t-1$, spread the rumour, with an individual j spreading the rumour to all individuals in the interval $[j, j+\rho_j]$ who have not been activated before. Define the following monotone decreasing event and its limit:

$$V_n := \{\text{the vertex } n \text{ is hit by an explosion}\} \text{ and } V = \lim_{n\to\infty} V_n.$$

Theorem 4: (Junior, *et al.* (2011)) For the homogeneous firework process we have

$$\sum_{n=1}^{\infty}\prod_{i=0}^{n}\mathbb{P}(\rho \le i) = \infty \text{ if and only if } \mathbb{P}(V) = 0\,.$$

Let $\mu := 1 + \sum_{n\ge 1}\prod_{i=0}^{n-1}\mathbb{P}(\rho \le i)$.

Theorem 5: (Gallo, *et al.* (2014)) For the homogeneous fireworks process

$$\mathbb{P}(V) = \frac{1}{\mu}\,.$$

Esmaeeli and Sajadi (2021) extended this result for the propogation of rumour among sceptics. Suppose that at the beginning, only two individuals $\{0,1\}$ are active and set $B_0 := \{0,1\}$. Define the sequence of events $(B_n)_{n\ge 1}$ as

$$B_n := \{i \ge 2 \ : \ \exists j_1 \ne j_2 \in \cup_{i=0}^{n-1} B_i \quad \text{such that} \quad i \in [j_1, j_1+\rho_{j_1}] \cap [j_2, j_2+\rho_{j_2}] \cap \mathbb{N}\}.$$

Let $B := \bigcup_{n\ge 1} B_n$. B is the set of all sceptic individuals who have heard the rumour.Let $\bar{\mathcal{A}}$ be the event that the rumour survives among sceptic individuals. We have the following result.

Theorem 6: (Esmaeeli and Sajadi (2021)) $\mathbb{P}(\bar{\mathcal{A}}) = \dfrac{1}{\bar{\mu}}$, where

$$\bar{\mu} = 2 + \sum_{k=2}^{\infty}\prod_{i=2}^{k}\bar{\alpha}_i \text{ and } \bar{\alpha}_i = \sum_{l=1}^{i}(-1)^{l-1}\sum_{I\subset\{1,\dots,i\},|I|=l}\prod_{r\in I}\prod_{k=1,k\ne r}^{i}\mathbb{P}(\rho \le k-1), \quad i \ge 2. \tag{8}$$

They also showed that $\bar{\mu} < \infty$ if and only if $\mu < \infty$ and from that they concluded the rumour dies out among sceptics under the same conditions presented in Gallo, *et al.* (2014) for non-sceptics.

Theorem 7: (Esmaeeli and Sajad (2021))

$$\mathbb{P}(\bar{\mathcal{A}}) = 0 \Longleftrightarrow \mathbb{P}(\mathcal{A}) = 0.$$

6. Further Questions

Bertacchi and Zucca (2013) studied the spread of rumour in a random environment on $\mathbb{N}$ and on Galton-Watson trees. Also, as in Mukhopadhyay, *et al.* (2020), a natural question is to ask for the rate of the spread of a rumour in a complete graph of N individuals, when every individual samples a fixed k number of individuals. The mean field limit of this model may suggest the rate of spread. Also if there are competing rumours, then a majority rule mechanism may also be used to find which rumour survives and which does not. This approach may provide rigorous answers to the simulation based observations of Zanette (2001, 2002).

Acknowledgements

Rahul Roy acknowledges support from the Matrics Grant MTR/2017/000141.

References

Athreya, S., Roy, R. and Sarkar, A. (2004). On the coverage of space by random sets. *Advances in Applied Probability (SGSA)*, **36**, 1–18.

Bertacchi, D. and Zucca, F. (2013). Rumour processes in random environment on $\mathbb{N}$ and on Galton-Watson trees. *Journal of Statistical Physics*, **153**, 486–511.

Chiu, S. N., Stoyan, D. N., Kendall, W. S. V. and Mecke, J. (2013). *Stochastic Geometry and Its Applications*. 3rd Ed., John Wiley, Chichester.

Esmaeeli, N. and Sajadi, F. A. (2020). Rumour propagation among sceptics: the Markovian case. *Indian Journal of Pure and Applied Mathematics*, **51**, 1661–1671.

Esmaeeli, N. and Sajadi, F. A. (2021). Study of the sceptical rumour processes on $\mathbb{N}$ using discrete renewal processes. *Preprint*.

Feller, W. (1971). *Introduction to Probability Theory and Its Applications*. Wiley Eastern Limited.

Franceschetti, M. and Meester, R. (2007). *Random Networks for Communication: From Statistical Physics to Information Systems*, Cambridge University Press, Cambridge.

Gallo, S., Garcia, N., Junior, V. and Rodríguez, P. (2014). Rumour processes on $\mathbb{N}$ and discrete renewal processes. *Journal of Statistical Physics*, **155**, 591–602 .

Gallo, S. and Garcia, N. L. (2018). Discrete one-dimensional coverage process on a renewal process.*Journal of Statistical Physics*, **173**, 381–397.

Gilbert, E. N. (1961). Random plane networks, *Journal of the Society for Industrial and Applied Mathematics.*, **22**, 89–103.

Gupta, P., and Kumar, P. R. (1998). Critical power for asymptotic connectivity in wireless networks. *Stochastic Analysis, Control, Optimization and Applications: A Volume in honor of W. H. Fleming*, 547–566.

Hall, P. (1988). *Introduction to The Theory of Coverage Processes*, John Wiley, New York.

Junior, V., Machado, F. and Zuluaga, M. (2011). Rumour processes on $\mathbb{N}$. *Journal of Applied Probability*, **48**, 624–636.

Kertesz, J. and Vicsek, T. (1982). Monte Carlo renormalization group study of the percolation problem of discs with a distribution of radii. *Zeitschrift für Physik B* , **45**, 345–350.

Maki, D. P. and Thompson, M. (1973). *Mathematical Models and Applications, With Emphasis on Social, Life, and Management Sciences*. Prentice-Hall, Englewood Cliffs, NJ.

Meester, R. and Roy, R. (1996). *Continuum Percolation* . Cambridge: Cambridge University Press.

Matheron, G. (1968). Modèle séquential de partition aléatorie. *Techical Report*, Centre de Morpholgie Mathématique, Fontainbleau.

Mukhopadhyay, A., Mazumdar, R. R. and Roy, R.(2020). Voter and majority dynamics with biased and stubborn agents. *Journal of Statistical Physics*, **181**, 239–1265.

Newman, M. E. J. and Watts, D. J. (1999). Scaling and percolation in the small-world network model. *Physical Review E*, **60(6)**, 7332–7342.

Penrose, M. (2003). *Random Geometric Graphs*, Oxford University Press, Oxford.

Sajadi, F. A. and Roy, R. (2019). On rumour propagation among sceptics. *Journal of Statistical Physics*, **174**, 935–952 .

Sudbury, A. (1985). The proportion of population never hearing a rumour. *Journal of Applied Probability.*, **22(2)**, 443–446.

Zanette, D. H. (2001). Critical behavior of propagation on small-world networks. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, **64**, 050901.

Zanette, D. H. (2002). Dynamics of rumour propagation on small-world networks. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, **65**, 041908.

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 453–466

Analysis of Replicated Order-of-Addition Experiments

Jianbin Chen[1,2], Xue-Ru Zhang[1] and Dennis K.J. Lin[2]
[1] *School of Statistics and Data Science, LPMC & KLMDASR*
Nankai University, Tianjin 300071, China
[2] *Department of Statistics, Purdue University, West Lafayette, IN 47907, U.S.A*

Received: 7 February 2021; Revised: 18 April 2021; Accepted: 19 April 2021

Abstract

The objective of the order-of-addition (OofA) problem is to find the optimal (addition) order. Existing literature concentrated on the responses of different orders with homoscedasticity. Study was made here for the cases of heteroscedasticity, where the dispersion effects for replicated OofA experiments should be considered. This paper proposes some approaches to speculate optimal orders for the replicated OofA experiment. Based on the pair-wise-order (PWO) model, the obtained orders from the proposed methodologies not only achieve the goal of OofA experiment, but also minimize the standard deviation within the OofA framework. Theoretical support is given under the specific setups. Simulation studies are used to illustrate these methodologies. It is shown that the proposed methods perform well for replicated OofA experiments.

Key words: Constrained optimization; Dual response; Mean square error; Pair-wise-order model.

AMS Subject Classifications: 62K05

1. Introduction

The order-of-addition (OofA) experiment has been found to have wide applications in many areas, such as, bio-chemistry, nutritional science and scheduling problems. The goal of the OofA experiment is to find the optimal order. Suppose the OofA experiment involves m (≥ 2) components. There will be $m!$ intrinsic different orders of adding sequences in the system. Each order π is a permutation of $\{1, \ldots, m\}$. It is not affordable to test all the $m!$ orders, especially when m is large (*e.g.*, when $m = 10, m! = 10! \approx 3.6$ millions). A (relatively small) subset is desirable to explore the optimal order. Empirical studies show that a random selection is rather inefficient (Zhao, *et al.*, 2020). Thus the design problem arises to choose a subset of all possible orders for searching the optimal order.

For any pair of components i and j, Van Nostrand (1995) proposed the pair-wise-order (PWO) factor. Define

$$I_{ij} = \begin{cases} 1, & \text{if } i \text{ precedes } j; \\ -1, & \text{if } j \text{ precedes } i, \end{cases} \tag{1}$$

Corresponding Author: Xue-Ru Zhang
Email: zhangxueru2019@gmail.com

as the PWO factor, where i and j are different components. Clearly, there are $q = \binom{m}{2}$ PWO factors, corresponding to all pairs of component orders, and such factors are arranged according to the lexicographic ordering of the components' indices. Denote β_{ij} as the effect to response caused by I_{ij}, then the PWO model is a linear model of

$$y(\pi) = \beta_0 + \sum_{i<j} \beta_{ij} I_{ij}(\pi) + \epsilon, \qquad \pi \in \Pi, \tag{2}$$

where y is the response of interest, ϵ is a random error from independent normal distribution $N(0, \sigma^2)$, and Π is the set of all of $m!$ possible orders. There are $p = q + 1$ parameters to be estimated. The PWO model is used in most recent literature. For a comprehensive discussion on OofA experiments, one may refer to Lin and Peng (2019), Peng, *et al.* (2019b), Voelkel (2019), Chen, *et al.* (2020a, 2020b, 2020c), Mee (2020), Winker, *et al.* (2020), Yang, *et al.* (2020), Zhao, *et al.* (2020) and the references therein.

In general, a good solution (optimal order) should be reproducible under various (mostly uncontrollable) environments. It is thus critical to study the stability of the optimal order. Ding, *et al.* (2015) investigated the sequence of drug administration that can impact clinical outcomes. They also showed that the setting of the time interval influences the final cell livability. Hence, the sequence of drug and time interval should be considered simultaneously in the experiment. However, time intervals are not easy to be controlled in practice. An alternative way is treating such factors as noise factors. Assume the relationship between factors (variables) and response y is formulated by $y = f(\pi, \boldsymbol{z})$, where π is the ordering factor and $\boldsymbol{z}$ is the vector of noise factors. Specifically, the variation of y stems from both the random error ϵ and the variation of $\boldsymbol{z}$. The purpose of this problem is then to find the optimal addition order for achieving the accepted targets, while simultaneously minimizing its standard deviation. The existing analytical OofA methods are not appropriate for this case. The dispersion effect for each OofA experiment should be considered.

For a replicated OofA experiment, the location and the dispersion effects of the OofA experiment are the two interested responses. Solving the replicated OofA experiment can then be considered as solving a dual response surface problem. For such a dual response problem, a good solution is to achieve at some compromise involving these two responses. The dual response surface approach employs two separate models to the mean and standard deviation of response. Vining and Myers (1990) utilized the dual response approach by Myers and Carter (1973), and demonstrated to optimize one response with an acceptable constraint on the value of another response. For more details please refer to Lin and Tu (1995), Copeland and Nelson (1996) and Kim and Lin (1998).

This paper proposes two modified dual response approaches to speculate optimal orders of OofA experiment for balancing two objective functions: (a) achieve the goal of location effect and (b) minimize the dispersion effect. The rest of the paper is organized as follows. Section 2 discusses the dual response surface approach for OofA experiments. A case study for job scheduling problems is discussed in Section 3. The corresponding theoretical support is given in Section 4. Section 5 introduces a simulation study with $m = 10$. The conclusion and discussion are given in Section 6.

2. Proposed Methods

2.1. Review of the analyzing approach for OofA experiment

The PWO model (2) is used as the assumed model. The sign of the true parameter β_{ij} shows the order of components i and j, *i.e.*, the positive sign of β_{ij} represents $i \longrightarrow j$ (component i shall proceed component j), when the goal of the objective is "the larger the better", while the negative sign of β_{ij} represents $j \longrightarrow i$. The hypothesis $H_0 : \beta_{ij} = 0$ is used to identify the significant parameters. Each component is regarded as a node in the directed graph. The arranged order "$i \longrightarrow j$" indicates an edge from i to j, and an insignificant $\hat{\beta}_{ij}$ (namely, $\beta_{ij} = 0$) implies no edge between i and j. One can generate the corresponding directed graph by sequentially connecting the nodes according to the all significant parameters. Finally, all feasible paths (a small candidate pool of optimal orders) can be obtained from the specifically directed graph. The details of finding optimal order(s) can be found in Chen, *et al.* (2020c).

2.2. Dispersion effects for OofA experiments

We next consider the OofA experiments involves heterogeneous standard deviations. Suppose each OofA experiment with $t \geq 2$ replicates. Let y_{kl} denote the lth replicated experiment (or response) at the kth treatment (or design point), where $l = 1, \ldots, t$ and $k = 1, \ldots, n$. Define

$$m_k = \frac{1}{t}\sum_{l=1}^{t} y_{kl}, \qquad k = 1, \ldots, n,$$

and

$$s_k = \sqrt{\frac{1}{t-1}\sum_{l=1}^{t}(y_{kl} - m_k)^2}, \qquad k = 1, \ldots, n.$$

Let $\hat{y}_\mu = (m_1, \ldots, m_n)$ and $\hat{y}_\sigma = (s_1, \ldots, s_n)$ be the estimator of the local effect (y_μ) and the dispersion effect (y_σ). Here, we apply a dual response surface approach to solve the replicated OofA problem under the PWO model (2). Suppose the fitted location response function is

$$y_\mu = \hat{\alpha}_0 + \sum_{i<j}(\hat{\alpha}^c_{ij} I^c_{ij} + \hat{\alpha}^\mu_{ij} I^\mu_{ij}) + \epsilon_\mu, \tag{3}$$

and the fitted dispersion response is

$$y_\sigma = \hat{\beta}_0 + \sum_{i<j}(\hat{\beta}^c_{ij} I^c_{ij} + \hat{\beta}^\sigma_{ij} I^\sigma_{ij}) + \epsilon_\sigma, \tag{4}$$

where I^c_{ij} denote the variables in both y_μ and y_σ; I^μ_{ij} and I^σ_{ij} denote the variables only in y_μ and y_σ, respectively. Therefore $I_{ij} = (I^c_{ij}, I^\mu_{ij}, I^\sigma_{ij})^T$. For the location function, there are three basic situations to be considered: (a) "the target is the best"; (b) "the larger the better"; and (c) "the smaller the better". For simplicity, we only focus on the case (a), the other two cases can be conducted in a similar manner. For the dispersion function, only the situation for the smaller the better needs to be considered.

2.3. Proposed methods

Here, we propose two methods to tackle the replicated OofA problem: (a) the two-step approach: first minimize the dispersion model, and next achieve the location model closer to the target T; and (b) the mean square error (MSE) approach.

The PWO model (2) is commonly used as the assumed model. The location model y_μ and dispersion model y_σ are fitted based on the obtained data from the OofA design. For the two-step approach, we choose the appropriate levels for some factors (I_{ij}^{μ}, not in y_σ) to make y_μ closer to the target value T, and select the level of factor ($I_{ij}^{c}, I_{ij}^{\sigma}$) to minimize y_σ. Combining these factors setting together, the optimal order(s) is finally obtained. (One could also first select the level of factor ($I_{ij}^{c}, I_{ij}^{\sigma}$) to minimize y_σ), and then choose the appropriate levels for factor (I_{ij}^{μ}, not in y_σ) to make y_μ closer to the target value T).

The two-step approach is powerful when the location model y_μ does not have any common factors in the dispersion model y_σ. Otherwise, adjusting the factor level in y_μ may make y_σ undesirable, *i.e.*, the two-step approach may obtain undesirable results when two interesting responses have common factors (see Section 5). In this case, the MSE approach (see below) should be considered.

The MSE criterion allows the location effect closer to the target T, while keeps the minimum standard deviation. The MSE criterion is defined as

$$\text{MSE} = \hat{y}_\sigma^2 + (\hat{y}_\mu - T)^2, \tag{5}$$

where T is the target value, $\hat{y}_\mu$ and $\hat{y}_\sigma$ are the estimates of y_μ and y_σ, respectively. The MSE approach are formally presented in Algorithm 1.

Algorithm 1 The MSE approach

Step 1 Based on the PWO model, select the best OofA design and conduct OofA experiment with $t \geq 2$ replicates;

Step 2 Based upon the obtained data, the location model y_μ and dispersion model y_σ are fitted;

Step 3 Evaluate the MSE value (5) and regard it as the response to be optimized;

Step 4 Construct the corresponding directed graph according to the active factors in the MSE (5).

Step 5 Obtain the optimal orders as the output.

In Step 3 of Algorithm 1, the MSE value for OofA experiment is considered as the response. In Step 4, we construct the corresponding directed graph according to the significant parameters in MSE (5). Based on the constructed directed graph, the optimal sequences are thus obtained.

3. An Illustrative Example

3.1. Problem formulation

Inspired by the drug experiment of Ding, *et al.* (2015), an illustrative example of OofA problems with heteroscedasticity is considered in this section. The noise factor z is subject to a uniform distribution over $[-1, 1]$. Let $m = 3$, the true relationship between the two kind variables and response $y \in [0, 1]$ be

$$y = 0.5 - 0.1(4I_{12}(\pi) + I_{13}(\pi))z + 0.1I_{23}(\pi) + \epsilon,$$

where $\epsilon \sim N(0, 0.09^2)$. The optimal order (unknown to us) is $2 \longrightarrow 1 \longrightarrow 3$.

The main purpose of this problem is finding the optimal (stable) order to make y_μ as close to $T = 1$ as possible. For each order, we conduct three replicated experiments and denote the obtained response values by Y_1, Y_2, Y_3. The order can be converted into PWO factor I_{ij}. Take the first run (row) of Table 1 as an example, $3 \longrightarrow 2 \longrightarrow 1$, the component 2 precedes component 1, then $I_{12} = -1$. Similarly, we have $I_{13} = -1, I_{23} = -1$. All possible $(3! = 6)$ experiments, their resulting replications Y_1, Y_2, Y_3, the corresponding values of I_{12}, I_{13} and I_{23} for each order, as well as the estimated $\hat{y}_\mu$ and $\hat{y}_\sigma$ are displayed in Table 1.

Table 1: The design and responses of the drug problem

Run	Order	I_{12}	I_{13}	I_{23}	Y_1	Y_2	Y_3	$\hat{y}_\mu$	$\hat{y}_\sigma$
1	$3 \longrightarrow 2 \longrightarrow 1$	-1	-1	-1	0.199	0.010	0.508	0.239	0.251
2	$3 \longrightarrow 1 \longrightarrow 2$	1	-1	-1	0.731	0.137	0.307	0.392	0.306
3	$2 \longrightarrow 3 \longrightarrow 1$	-1	-1	1	0.223	0.605	0.385	0.404	0.192
4	$2 \longrightarrow 1 \longrightarrow 3$	-1	1	1	0.363	0.391	0.423	0.392	0.030
5	$1 \longrightarrow 2 \longrightarrow 3$	1	1	1	0.573	0.350	0.709	0.544	0.181
6	$1 \longrightarrow 3 \longrightarrow 2$	1	1	-1	0.332	0.085	0.053	0.157	0.153

3.2. Conventional approach

In the conventional method, the dispersion response y_σ is considered to be a constant; and the local response y_μ is considered as the only response. Without consideration of y_σ, the existing conventional OofA methods aim to find the optimal order based on y_μ. The fitting model for the location effect (y_μ) is,

$$y_\mu = 0.355 + 0.092I_{23}(\pi) + \epsilon_\mu. \tag{6}$$

This problem is essentially an unconstrained optimization problem that aims to make y_μ in (6) as close to $T = 1$ as possible. For the fitting model (6), the sign of the significant parameter β_{23} is positive ("+"), hence, the possible optimal orders are subject to $2 \longrightarrow 3$. Model (6) provides no information on the order relative to the component 1. Consequently, three possible orders are: (1) $1 \longrightarrow 2 \longrightarrow 3$; (2) $2 \longrightarrow 1 \longrightarrow 3$; and (3) $2 \longrightarrow 3 \longrightarrow 1$. All orders have the prediction $\hat{y}_\mu = 0.447$ via (6). With the conventional approach, any of these three orders can be considered as the optimal order. Note that these three orders have the same $\hat{y}_\mu$, but different standard deviations $\hat{y}_\sigma$ (as will be shown next). To obtain the stable order(s), new approach should be employed.

3.3. Proposed approaches

According the data in Table 1, the fitting models for y_μ and y_σ of the replicated response are

$$y_\mu = 0.355 + 0.092I_{23}(\pi) + \epsilon_\mu, \quad \text{and}$$
$$y_\sigma = 0.185 + 0.055I_{12}(\pi) - 0.082I_{13}(\pi) + \epsilon_\sigma. \tag{7}$$

For the two-step approach, the first step sets the level of the adjusting variables $I_{2,3}(\pi)$ as "+1" for closing $\hat{y}_\mu$ to $T = 1$. Thus, the possible orders should satisfy the arranged order "$2 \longrightarrow 3$". The second step is to determine the arranged orders of significant parameters to minimize $\hat{y}_\sigma$ in (7). Given the constraint $\hat{y}_\sigma \geq 0$, the recommend levels of PWO factors are set as $I_{1,2}(\pi) = -1$ and $I_{1,3}(\pi) = 1$ for minimizing $\hat{y}_\sigma$. Hence, the arranged orders are "$2 \longrightarrow 1$" and "$1 \longrightarrow 3$". Combining those arranged orders together, the optimal order $2 \longrightarrow 1 \longrightarrow 3$ is resulted. This is indeed the true optimal order.

Next, the MSE approach in Algorithm 1 is used to solve the problem. The MSE criterion $(\hat{y}_\mu - T)^2 + \hat{y}_\sigma^2$ is used to find the optimal order. In Step 3, the objective function of this example becomes

$$\min \quad (0.092I_{23}(\pi) - 0.645)^2 + (0.185 + 0.055I_{12}(\pi) - 0.082I_{13}(\pi))^2$$

For Step 4 of Algorithm 1, the level of PWO factors are found to be $I_{1,2}(\pi) = -1, I_{1,3}(\pi) = 1$ and $I_{2,3}(\pi) = 1$ for minimizing MSE. Thus, the possible orders should achieve "$2 \longrightarrow 1$", "$1 \longrightarrow 3$" and "$2 \longrightarrow 3$". Based on the generated directed graph, the optimal order $2 \longrightarrow 1 \longrightarrow 3$ is obtained. That is identical to the true optimal order.

3.4. Discussion

Using the conventional approach, three orders are resulted. For any OofA problem, y_σ should be taken into account to explore an optimal order. The dispersion effect y_σ mainly stems from the noise factor $\boldsymbol{z}$ (hard to be controlled accurately in practice). Therefore, the location model and dispersion model are respectively built based on both the control factors and ordering factors (such as PWO factors). The experimental goal is to find an optimal order such that (a) location effect y_μ is closer to the target T; and (b) dispersion effect y_σ is minimized. The proposed methods (both the two-step method and the MSE approach) yield the same optimal order $2 \longrightarrow 1 \longrightarrow 3$. This may not be the case in general.

Compared with the conventional approach which only considers y_μ (assuming y_σ is a constant), the proposed methods not only focus on the location model y_μ, but also consider the dispersion model y_σ. The dispersion model y_σ adds more restrictions on the possible orders decided by the location model y_μ. For this case, the resulting orders of the MSE approach achieve the target of location model(y_μ), while perform well on dispersion model (y_σ).

4. Theoretical Supports

Denote S_1 and S_2 are the candidate pool of optimal orders for $\hat{y}_\mu$ (3) and $\hat{y}_\sigma$ (4), respectively. Let S be the candidate pool of optimal orders for $\hat{y}_\mu$ (3) and $\hat{y}_\sigma$ (4). Suppose there exist orders $\pi_1 \in S_1$, $\pi_2 \in S_2$ and $\pi \in S$, we have the following theorem.

Theorem 1: For replicated OofA experiments, if we have orders $\pi_1 \in S_1$, $\pi_2 \in S_2$ and $\pi \in S$, then $MSE(\pi) \leq MSE(\pi_1)$ and $MSE(\pi) \leq MSE(\pi_2)$.

Proof: For simplicity, we only discuss the case of "the smaller, the better". The proofs of the remaining two cases ("the larger, the better" and "the target, the better") are similar and thus omitted. Recall that the fitted location model is

$$y_\mu = \hat{\alpha}_0 + \sum_{i<j}(\hat{\alpha}^c_{ij} I^c_{ij} + \hat{\alpha}^\mu_{ij} I^\mu_{ij}) + \epsilon_\mu,$$

and the fitted dispersion model is

$$y_\sigma = \hat{\beta}_0 + \sum_{i<j}(\hat{\beta}^c_{ij} I^c_{ij} + \hat{\beta}^\sigma_{ij} I^\sigma_{ij}) + \epsilon_\sigma,$$

where I^c_{ij} denote the variables in both y_μ and y_σ; I^μ_{ij} and I^σ_{ij} denote the variables only in y_μ and y_σ, respectively.

Case 1. Suppose location model y_μ (3) and dispersion model y_σ (4) have common factors, *i.e.*, $I^\mu_{ij} = I^\sigma_{ij} = 0$ for all $i, j \in \{1, \ldots, m\}$. In this case, the optimal order π_1 of y_μ (3) is also the optimal order for y_σ (4). Hence, for $\pi_1 \in S_1$ and $\pi \in S$, we have $MSE(\pi) = MSE(\pi_1)$.

Case 2. Suppose there exist an factor making $I^\mu_{kl}(\pi_1) = -I^\sigma_{kl}(\pi_2) \neq 0$. Note that

$$\begin{aligned} MSE(\pi_1) &= \hat{y}^2_\mu(\pi_1) + \hat{y}^2_\sigma(\pi_1) \\ &= \left[\hat{\alpha}_0 + \sum_{i<j, i\neq k,l; j\neq k,l}(\hat{\alpha}^c_{ij} I^c_{ij} + \hat{\alpha}^\mu_{ij} I^\mu_{ij}) + \hat{\alpha}^\mu_{kl} I^\mu_{kl}\right]^2 (\pi_1) + \\ &\quad \left[\hat{\beta}_0 + \sum_{i<j, i\neq k,l; j\neq k,l}(\hat{\beta}^c_{ij} I^c_{ij} + \hat{\beta}^\sigma_{ij} I^\sigma_{ij}) + \hat{\beta}^\sigma_{kl} I^\sigma_{kl}\right]^2 (\pi_1) \\ &= C + (\hat{\alpha}^\mu_{kl} + \hat{\beta}^\sigma_{kl}) B, \end{aligned}$$

where B and C are constant with other active standard deviations. Similarly, we have

$$\begin{aligned} MSE(\pi) &= \hat{y}^2_\mu(\pi) + \hat{y}^2_\sigma(\pi) \\ &= C + (\hat{\alpha}^\mu_{kl} - \hat{\beta}^\sigma_{kl}) B. \end{aligned}$$

Obviously, we have $MSE(\pi) < MSE(\pi_1)$. In this case, one seeks the optimal order π_1 making the target optimal regardless of standard deviation. Theses π_1 making the standard deviation large.

Case 3. For $\{k, l, s, t\} \neq \{i, j\}$, we have two active standard deviations $I^p_{kl}(\pi_1) \neq 0$ and $I^s_{st}(\pi_2) \neq 0$. Note that

$$\begin{aligned} MSE(\pi_1) &= \hat{y}^2_\mu(\pi_1) + \hat{y}^2_\sigma(\pi_1) \\ &= \left[\hat{\alpha}_0 + \sum_{i<j, i\neq k,l; j\neq k,l}(\hat{\alpha}^c_{ij} I^c_{ij} + \hat{\alpha}^\mu_{ij} I^\mu_{ij}) + \hat{\alpha}^\mu_{kl} I^\mu_{kl}\right]^2 (\pi_1) + \\ &\quad \left[\hat{\beta}_0 + \sum_{i<j, i\neq k,l; j\neq k,l}(\hat{\beta}^c_{ij} I^c_{ij} + \hat{\beta}^\sigma_{ij} I^\sigma_{ij})\right]^2 (\pi_1). \end{aligned}$$

Similarly, we have

$$\begin{aligned} MSE(\pi) &= \hat{y}_\mu^2(\pi) + \hat{y}_\sigma^2(\pi) \\ &= \left[\hat{\alpha}_0 + \sum_{i<j,i\neq k,l;j\neq k,l} (\hat{\alpha}_{ij}^c I_{ij}^c + \hat{\alpha}_{ij}^\mu I_{ij}^\mu) + \hat{\alpha}_{kl}^\mu I_{kl}^\mu\right]^2 (\pi) + \\ &\quad \left[\hat{\beta}_0 + \sum_{i<j,i\neq k,l;j\neq k,l} (\hat{\beta}_{ij}^c I_{ij}^c + \hat{\beta}_{ij}^\sigma I_{ij}^\sigma) + \hat{\beta}_{kl}^\sigma I_{kl}^\sigma\right]^2 (\pi). \end{aligned}$$

The solution π want to keep $\hat{y}_\sigma^2(\pi)$ smaller. Obviously, $\hat{y}_\mu^2(\pi_1) = \hat{y}_\mu^2(\pi)$ and $\hat{y}_\sigma^2(\pi) = \hat{y}_\sigma^2(\pi_1)$, thus $MSE(\pi) < MSE(\pi_1)$.

The other cases can be similarly proved. Hence, the proof is completed. □

Theorem 1 shows that the obtained order by the proposed approach is optimal compared to the conventional method. This indicates that the proposed approaches have a smaller MSE value in the replicated experiments.

5. A Numerical Simulation

Here, we provide a distinct example with $m = 10$ for illustrating the effectiveness of the proposed method for OofA experiments with heteroscedasticity. The underlying true model is

$$\begin{aligned} y &= 50 + 2I_{13}(\pi) - 5I_{14}(\pi) - 2I_{26}(\pi) + 3I_{2(10)}(\pi) + 3I_{35}(\pi)z_1 + 3I_{39}(\pi) - I_{45}(\pi) \\ &\quad -2I_{5(10)}(\pi) - 2I_{67}(\pi) + 4I_{68}(\pi)z_2 + 7I_{78}(\pi) + 5I_{7(10)}(\pi)z_1 + 4I_{8(10)}(\pi)z_2 + \epsilon, \quad (8) \end{aligned}$$

where z_1 and z_2 are two noise factors with $z_1 \sim N(0, 0.5^2)$, $z_2 \sim N(0, 1)$, and $\epsilon \sim N(0, 0.05^2)$. Note that $I_{2(10)}$ (for example) is the PWO variable between components 2 and 10. The purpose of this experiment is to find optimal order making y_μ close to $T = 23$, while minimizing y_σ. The optimal order is in fact $9 \longrightarrow 3 \longrightarrow 1 \longrightarrow 5 \longrightarrow 4 \longrightarrow 10 \longrightarrow 2 \longrightarrow 6 \longrightarrow 8 \longrightarrow 7$, whose resulting expectation is 23 with standard deviation 1.001.

A D-optimal OofA design (from Winker *et al.*, 2020), with 46 $\left(= 1 + \binom{10}{2}\right)$ runs, is chosen the OofA design. For each run, five replicated experiments are conducted and their responses $Y_1, \ldots, Y_5$ are obtained. Two responses ($\hat{y}_\mu$ and $\hat{y}_\sigma$) are evaluated by those five replicated responses. This is displayed in the Appendix (Table A.1). Via stepwise regression method, the location model and the dispersion model are respectively fitted as

$$\begin{aligned} y_\mu &= 49.583 + 2.231I_{13}(\pi) - 4.513I_{14}(\pi) - 1.286I_{18}(\pi) - 0.843I_{26}(\pi) - 1.085I_{27}(\pi) + 3.318 \\ &\quad \times I_{2(10)}(\pi) + 3.849I_{39}(\pi) + 0.959I_{3(10)}(\pi) - 1.145I_{45}(\pi) - 1.354I_{49}(\pi) - 3.406I_{5(10)}(\pi) \\ &\quad +7.843I_{78}(\pi) + \epsilon_\mu \quad (9) \end{aligned}$$

and

$$\begin{aligned} y_\sigma &= 5.051 + 1.697I_{14}(\pi) + 0.764I_{29}(\pi) - 2.423I_{35}(\pi) + 2.735I_{45}(\pi) + 1.618I_{57}(\pi) \\ &\quad -1.226I_{67}(\pi) + \epsilon_\sigma. \quad (10) \end{aligned}$$

We first employ the two-step approach to solve this example. To make y_μ close to $T = 23$, the recommend levels of PWO factors in y_μ (9) are $I_{13}(\pi) = I_{2(10)}(\pi) = I_{39}(\pi) = I_{49}(\pi) = I_{78}(\pi) = -1$ and $I_{14}(\pi) = I_{18}(\pi) = I_{26}(\pi) = I_{27}(\pi) = I_{3(10)}(\pi) = I_{45}(\pi) = I_{5(10)}(\pi) = 1$. According to those active factors, the orders of "3 $\longrightarrow$ 1, 10 $\longrightarrow$ 2, 9 $\longrightarrow$ 3, 9 $\longrightarrow$ 4, 8 $\longrightarrow$ 7,1 $\longrightarrow$ 4, 1 $\longrightarrow$ 8, 2 $\longrightarrow$ 6, 2 $\longrightarrow$ 7, 3 $\longrightarrow$ 10, 4 $\longrightarrow$ 5, 5 $\longrightarrow$ 10" are obtained. Given the constraint $y_\sigma \geq 0$, the adjusted factors are set as $I_{35}(\pi) = I_{57}(\pi) = I_{67}(\pi) = 1$ and $I_{29}(\pi) = -1$. From those active factors, the optimal orders satisfy "3 $\longrightarrow$ 5, 5 $\longrightarrow$ 7, 6 $\longrightarrow$ 7, 9 $\longrightarrow$ 2 ". Based on those arranged orders, all optimal orders can be found in the Appendix (Table A.2). As an example, one of the orders is 9 $\longrightarrow$ 3 $\longrightarrow$ 1 $\longrightarrow$ 4 $\longrightarrow$ 5 $\longrightarrow$ 10 $\longrightarrow$ 2 $\longrightarrow$ 6 $\longrightarrow$ 8 $\longrightarrow$ 7 ($\hat{y}_\mu = 22.377$, $\hat{y}_\sigma = 6.688$, $MSE = 45.118$ via (9), (10) and (5), respectively).

To avoid the above drawback of two-step method, the MSE method (Algorithm 1) is used here to find the optimal order. The objective function of MSE method in Step 3 is Equation (5) $(\hat{y}_\mu - T)^2 + \hat{y}_\sigma^2$, where now $\hat{y}_\mu$ and $\hat{y}_\sigma$ represent the location model (9) and dispersion model (10), respectively. Via Step 4 of Algorithm 1, the level of PWO factors are set as $I_{14}(\pi) = I_{18}(\pi) = I_{26}(\pi) = I_{27}(\pi) = I_{35}(\pi) = I_{3(10)}(\pi) = I_{57}(\pi) = I_{5(10)}(\pi) = I_{67}(\pi) = 1$ and $I_{13}(\pi) = I_{29}(\pi) = I_{45}(\pi) = I_{49}(\pi) = I_{2(10)}(\pi) = I_{39}(\pi) = I_{78}(\pi) = -1$.

Thus, the possible orders are "1 $\longrightarrow$ 4, 1 $\longrightarrow$ 8, 2 $\longrightarrow$ 6, 2 $\longrightarrow$ 7, 3 $\longrightarrow$ 5, 3 $\longrightarrow$ 10, 5 $\longrightarrow$ 7, 5 $\longrightarrow$ 10, 6 $\longrightarrow$ 7, 3 $\longrightarrow$ 1, 9 $\longrightarrow$ 2, 5 $\longrightarrow$ 4, 10 $\longrightarrow$ 2, 9 $\longrightarrow$ 4, 9 $\longrightarrow$ 3, 8 $\longrightarrow$ 7". Each component is regarded as a node, the possible order "$i \longrightarrow j$" implies a directed edge from i to j. One can generate the directed graph by connecting all directed edges (see Figure 1). According to Figure 1, all optimal orders can be found in Table 2. For example, one possible order is 9 $\longrightarrow$ 3 $\longrightarrow$ 1 $\longrightarrow$ 5 $\longrightarrow$ 4 $\longrightarrow$ 10 $\longrightarrow$ 2 $\longrightarrow$ 6 $\longrightarrow$ 8 $\longrightarrow$ 7 ($\hat{y}_\mu = 24.667$, $\hat{y}_\sigma = 1.218$ and $MSE = 4.262$ via (9), (10) and (5), respectively). This is to make $\hat{y}_\mu$ close to the target $T = 23$ while keeping $\hat{y}_\sigma$ relatively small. As compared to the solution from the two-step approach, the order obtained by the MSE approach has a much smaller $\hat{y}_\sigma$ while $\hat{y}_\mu$ is also close to the target. For confirmation, those two orders obtained by two-step as well MSE approaches were evaluated via the true model (8). It is shown that the expectations are 21 (for two-step order), and 23 (for MSE order), respectively; with identical standard deviations of 1.001.

Table 2: The optimal orders by MSE approach

Run	Order
1	9 $\longrightarrow$ 3 $\longrightarrow$ 1 $\longrightarrow$ 8 $\longrightarrow$ 5 $\longrightarrow$ 4 $\longrightarrow$ 10 $\longrightarrow$ 2 $\longrightarrow$ 6 $\longrightarrow$ 7
2	9 $\longrightarrow$ 3 $\longrightarrow$ 1 $\longrightarrow$ 5 $\longrightarrow$ 8 $\longrightarrow$ 4 $\longrightarrow$ 10 $\longrightarrow$ 2 $\longrightarrow$ 6 $\longrightarrow$ 7
3	9 $\longrightarrow$ 3 $\longrightarrow$ 1 $\longrightarrow$ 5 $\longrightarrow$ 4 $\longrightarrow$ 8 $\longrightarrow$ 10 $\longrightarrow$ 2 $\longrightarrow$ 6 $\longrightarrow$ 7
4	9 $\longrightarrow$ 3 $\longrightarrow$ 1 $\longrightarrow$ 5 $\longrightarrow$ 4 $\longrightarrow$ 10 $\longrightarrow$ 8 $\longrightarrow$ 2 $\longrightarrow$ 6 $\longrightarrow$ 7
5	9 $\longrightarrow$ 3 $\longrightarrow$ 1 $\longrightarrow$ 5 $\longrightarrow$ 4 $\longrightarrow$ 10 $\longrightarrow$ 2 $\longrightarrow$ 8 $\longrightarrow$ 6 $\longrightarrow$ 7
6	9 $\longrightarrow$ 3 $\longrightarrow$ 1 $\longrightarrow$ 5 $\longrightarrow$ 4 $\longrightarrow$ 10 $\longrightarrow$ 2 $\longrightarrow$ 6 $\longrightarrow$ 8 $\longrightarrow$ 7
7	9 $\longrightarrow$ 3 $\longrightarrow$ 5 $\longrightarrow$ 1 $\longrightarrow$ 8 $\longrightarrow$ 4 $\longrightarrow$ 10 $\longrightarrow$ 2 $\longrightarrow$ 6 $\longrightarrow$ 7
8	9 $\longrightarrow$ 3 $\longrightarrow$ 5 $\longrightarrow$ 1 $\longrightarrow$ 4 $\longrightarrow$ 8 $\longrightarrow$ 10 $\longrightarrow$ 2 $\longrightarrow$ 6 $\longrightarrow$ 7
10	9 $\longrightarrow$ 3 $\longrightarrow$ 5 $\longrightarrow$ 1 $\longrightarrow$ 4 $\longrightarrow$ 10 $\longrightarrow$ 8 $\longrightarrow$ 2 $\longrightarrow$ 6 $\longrightarrow$ 7
10	9 $\longrightarrow$ 3 $\longrightarrow$ 5 $\longrightarrow$ 1 $\longrightarrow$ 4 $\longrightarrow$ 10 $\longrightarrow$ 2 $\longrightarrow$ 8 $\longrightarrow$ 6 $\longrightarrow$ 7
11	9 $\longrightarrow$ 3 $\longrightarrow$ 5 $\longrightarrow$ 1 $\longrightarrow$ 4 $\longrightarrow$ 10 $\longrightarrow$ 2 $\longrightarrow$ 6 $\longrightarrow$ 8 $\longrightarrow$ 7

Note: All orders have $\hat{y}_\mu = 24.667$ and $\hat{y}_\sigma = 1.218$ and MSE=4.262.

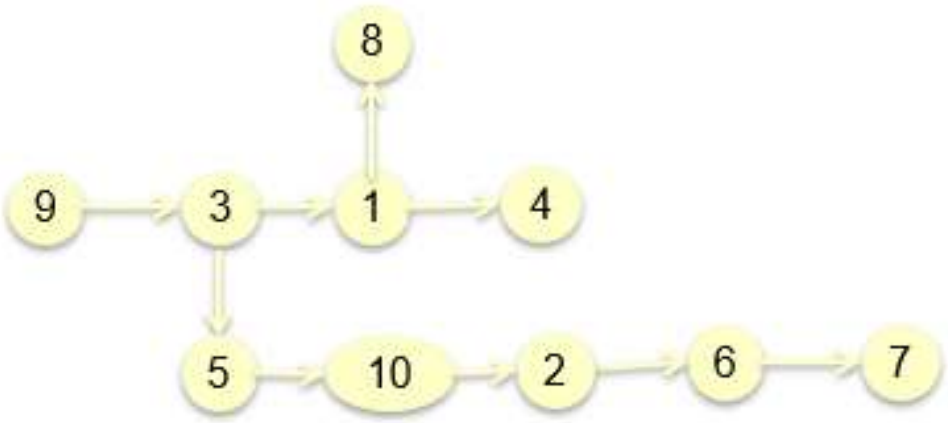


Figure 1: Directed graph

6. Conclusion and Discussion

The goal for an OofA experiment is to find the optimal order. The current research focuses on the case where the dispersion effect for each order is a constant ($y_\sigma \equiv c$). For the OofA experiments with heteroscedasticity, the dual response approach is employed. The location and the dispersion effects are the two interested responses. The two-step and MSE approaches are proposed to speculate optimal orders. When the location model and the dispersion model do not share any common factors, the commonly used two-step method, is able to find optimal order. The location model and dispersion model typically share some common active factors. The two-step approach may be misleading in this case. Lin and Tu (1995) showed that the optimization problem based on MSE is more appropriate to solve the dual response problem. Motivated by their idea, this paper proposes an MSE approach to find the optimal orders for the OofA experiments. Based on the MSE, the obtained orders not only achieve the goal of the OofA experiment, but also minimize the standard deviation within the OofA framework. Some theoretical supports are given (Section 4) to illustrate the effectiveness of the proposed method for OofA experiments with heteroscedasticity. Simulation studies confirmed that the proposed approaches perform well for searching the optimal order(s) in replicated OofA experiments.

The proposed method can be easily extended to the OofA problem with some prior information, the objective function of MSE criterion can be rewritten as

$$\text{MSE} = \omega\hat{y}_\sigma^2 + (1-\omega)(\hat{y}_\mu - T)^2, \tag{11}$$

where ω measures the relative importance of $\hat{y}_\sigma$ and $\hat{y}_\mu$ with $0 \leq \omega \leq 1$. Especially, for $\omega > 0.5$, the experimenter is inclined to "risk lover"; for $\omega = 0.5$, the experimenter tends to "risk-neutral"; and for $\omega < 0.5$, the experimenter is more likely to "risk averter". Naturally, if there is no prior information, we suggest setting $\omega = 0.5$. The MSE (11) can be used in Step 3 for Algorithm 1 to tackle the replicated OofA experiment with prior information.

Acknowledgements

This article is in memory of our best friend Aloke Dey. Aloke was a great researcher. His research is deeply important to us. The first two authors contributed equally to the work. The work of Lin is supported by the *National Science Foundation* via Grant DMS-18102925. The work of the Chen is supported by the National Natural Science Foundation of China (Grant No. 11771220), National Ten-Thousand Talents Program, Tianjin Development Program for Innovation and Entrepreneurship (Grant No.C02055). The work of Zhang was supported by China Scholarship Council.

References

Box, G. E. and Wilson, K. B. (1951). On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society: Series B (Methodological)*, 1–45.

Chen, J. B., Han, X. X., Yang, L. Q., Lin, D. K. J. and Zhou, Y. D. (2020a). Robust designs for ordering problem. Submitted manuscript.

Chen, J. B., Mukerjee, R. and Lin, D. K. J. (2020b). Construction of optimal fractional Order-of-Addition designs via block designs. *Statistics & Probability Letters*, forthcoming.

Chen, J. B., Peng, J. Y. and Lin, D. K. J. (2020c). A statistical perspective on NP-Hard problems: making uses of design for order-of-addition experiment. Submitted manuscript.

Copeland, K. A. and Nelson P. R. (1996) Dual response optimization via direct function minimization. *Journal of Quality Technology*, **28**, 331–336.

Ding, X., Matsuo, K., Xu, L., Yang, J. and Zheng, L. (2015). Optimized combinations of bortezomib, camptothecin, and doxorubicin show increased efficacy and reduced toxicity in treating oral cancer. *Anti-Cancer Drugs*, **26**, 547–554.

Kim, K. and Lin, D. K. J. (1998). Dual response surface optimization: A fuzzy modeling approach. *Journal of Quality Technology*, **30**, 1–10.

Lin, D. K. J. and Peng, J. Y. (2019). The order-of-addition experiments: a review and some new thoughts (with discussion). *Quality Engineering*, **31**(1), 49–59.

Lin, D. K. J. and Tu, W. Z. (1995). Dual response surface optimization. *Journal of Quality Technology*, **27**(1), 34–39.

Mee, R. W. (2020). Order of addition modeling. *Statistica Sinica*, **30**, 1543–1559..

Myers, R. H. and Carter, W. H. (1973). Response surface techniques for dual response systems. *Technometrics*, **15**, 301–317.

Peng, J. Y., Barton, R. and Lin, D. K. J. (2019a). Swap-r algorithm for the construction of order-of-addition designs. Submitted manuscript.

Peng, J. Y., Mukerjee, R. and Lin, D. K. J. (2019b). Design of order-of-addition experiments. *Biometrika*, **106** (3), 683—694.

Van Nostrand, R. C. (1995). Design of experiments where the order of addition is important. *ASA Proceedings of the Section on Physical and Engineering Sciences*. **155–160**. American Statistical Association, Alexandria, Virginia.

Vining, G. G., and Myers, R. H. (1990). Combining Taguchi and response surface philosophies: a dual response approach. *Journal of Quality Technology*, **22**(1), 38–45.

Voelkel, J. G. (2019). The design of order-of-addition experiments. *Journal of Quality Technology*, 51, 230–241.

Winker, P., Chen, J. B. and Lin, D. K. J . (2020). The construction of optimal design for order-of-addition experiment via threshold accepting. In *Contemporary Design of Experiments, Multivariate Analysis and Data mining*—In Celebration of Prof. Kai-Tai Fang's 80th Birthday (pp. 93–109). Springer.

Yang, J. F., Sun, F. S. and Xu, H. Q. (2020). A component-position model, analysis and design for order-of-addition experiments. *Technometrics*, forthcoming.

Zhao, Y. N., Lin, D. K. J. and Liu, M. Q. (2020). Designs for order of addition experiments. *Journal of Applied Statistics*, forthcoming.

ANNEXURE

Table A.1: The design and responses for $m = 10$ in Section 5

Run	Order	Y_1	Y_2	Y_3	Y_4	Y_5	$\hat{y}_\mu$	$\hat{y}_\sigma$
1	$4 \longrightarrow 5 \longrightarrow 10 \longrightarrow 2 \longrightarrow 3 \longrightarrow 8 \longrightarrow 9 \longrightarrow 6 \longrightarrow 7 \longrightarrow 1$	21.730	51.429	37.559	37.766	20.475	33.792	12.883
2	$8 \longrightarrow 3 \longrightarrow 4 \longrightarrow 10 \longrightarrow 6 \longrightarrow 5 \longrightarrow 9 \longrightarrow 2 \longrightarrow 7 \longrightarrow 1$	46.328	47.872	47.869	47.793	48.517	47.676	0.808
3	$3 \longrightarrow 1 \longrightarrow 8 \longrightarrow 10 \longrightarrow 6 \longrightarrow 2 \longrightarrow 9 \longrightarrow 5 \longrightarrow 7 \longrightarrow 4$	40.144	38.301	37.612	40.160	40.079	39.259	1.214
4	$9 \longrightarrow 6 \longrightarrow 4 \longrightarrow 10 \longrightarrow 8 \longrightarrow 7 \longrightarrow 3 \longrightarrow 5 \longrightarrow 2 \longrightarrow 1$	39.611	42.334	39.546	40.570	41.057	40.624	1.152
5	$8 \longrightarrow 2 \longrightarrow 10 \longrightarrow 6 \longrightarrow 4 \longrightarrow 7 \longrightarrow 9 \longrightarrow 1 \longrightarrow 3 \longrightarrow 5$	44.863	46.482	46.901	46.364	45.377	45.997	0.845
6	$9 \longrightarrow 2 \longrightarrow 7 \longrightarrow 3 \longrightarrow 6 \longrightarrow 5 \longrightarrow 10 \longrightarrow 1 \longrightarrow 8 \longrightarrow 4$	48.396	48.322	48.050	51.828	51.925	49.704	1.988
7	$3 \longrightarrow 9 \longrightarrow 1 \longrightarrow 7 \longrightarrow 8 \longrightarrow 5 \longrightarrow 4 \longrightarrow 10 \longrightarrow 2 \longrightarrow 6$	48.768	50.391	50.193	49.121	50.767	49.848	0.859
8	$1 \longrightarrow 5 \longrightarrow 6 \longrightarrow 2 \longrightarrow 9 \longrightarrow 3 \longrightarrow 10 \longrightarrow 4 \longrightarrow 7 \longrightarrow 8$	52.615	55.621	59.494	53.982	58.514	56.045	2.924
9	$6 \longrightarrow 10 \longrightarrow 2 \longrightarrow 4 \longrightarrow 3 \longrightarrow 9 \longrightarrow 1 \longrightarrow 8 \longrightarrow 7 \longrightarrow 5$	47.376	48.119	46.674	46.319	46.005	46.899	0.852
10	$10 \longrightarrow 6 \longrightarrow 7 \longrightarrow 2 \longrightarrow 1 \longrightarrow 8 \longrightarrow 5 \longrightarrow 4 \longrightarrow 3 \longrightarrow 9$	55.107	58.215	62.524	60.011	52.709	57.713	3.890
11	$8 \longrightarrow 9 \longrightarrow 7 \longrightarrow 5 \longrightarrow 1 \longrightarrow 10 \longrightarrow 3 \longrightarrow 4 \longrightarrow 2 \longrightarrow 6$	33.430	33.170	32.657	33.080	31.654	32.798	0.697
12	$10 \longrightarrow 8 \longrightarrow 2 \longrightarrow 6 \longrightarrow 4 \longrightarrow 3 \longrightarrow 5 \longrightarrow 7 \longrightarrow 1 \longrightarrow 9$	44.626	35.094	58.733	37.707	39.556	43.143	9.387
13	$10 \longrightarrow 2 \longrightarrow 1 \longrightarrow 7 \longrightarrow 8 \longrightarrow 5 \longrightarrow 9 \longrightarrow 3 \longrightarrow 6 \longrightarrow 4$	61.054	37.452	40.934	44.154	47.954	46.310	9.112
14	$7 \longrightarrow 4 \longrightarrow 9 \longrightarrow 1 \longrightarrow 3 \longrightarrow 10 \longrightarrow 6 \longrightarrow 5 \longrightarrow 8 \longrightarrow 2$	66.904	60.267	56.888	60.988	57.257	60.461	4.027
15	$5 \longrightarrow 6 \longrightarrow 8 \longrightarrow 2 \longrightarrow 1 \longrightarrow 4 \longrightarrow 10 \longrightarrow 9 \longrightarrow 3 \longrightarrow 7$	39.443	42.287	26.276	39.532	19.579	33.423	9.936
16	$9 \longrightarrow 3 \longrightarrow 7 \longrightarrow 8 \longrightarrow 6 \longrightarrow 4 \longrightarrow 2 \longrightarrow 1 \longrightarrow 5 \longrightarrow 10$	59.972	62.076	63.205	55.378	64.840	61.094	3.653
17	$7 \longrightarrow 1 \longrightarrow 6 \longrightarrow 4 \longrightarrow 2 \longrightarrow 8 \longrightarrow 3 \longrightarrow 5 \longrightarrow 10 \longrightarrow 9$	64.383	63.005	62.900	47.747	48.806	57.368	8.328
18	$8 \longrightarrow 5 \longrightarrow 9 \longrightarrow 2 \longrightarrow 10 \longrightarrow 1 \longrightarrow 3 \longrightarrow 6 \longrightarrow 7 \longrightarrow 4$	30.842	36.799	40.545	36.269	30.181	34.927	4.361
19	$8 \longrightarrow 6 \longrightarrow 4 \longrightarrow 9 \longrightarrow 3 \longrightarrow 10 \longrightarrow 1 \longrightarrow 2 \longrightarrow 7 \longrightarrow 5$	41.884	42.204	40.577	41.517	39.555	41.147	1.079
20	$3 \longrightarrow 8 \longrightarrow 2 \longrightarrow 7 \longrightarrow 1 \longrightarrow 4 \longrightarrow 5 \longrightarrow 6 \longrightarrow 9 \longrightarrow 10$	41.364	40.458	43.810	30.455	42.614	39.740	5.343
21	$10 \longrightarrow 4 \longrightarrow 1 \longrightarrow 3 \longrightarrow 2 \longrightarrow 6 \longrightarrow 9 \longrightarrow 7 \longrightarrow 5 \longrightarrow 8$	62.047	61.259	60.860	60.301	60.776	61.049	0.654
22	$4 \longrightarrow 2 \longrightarrow 8 \longrightarrow 1 \longrightarrow 7 \longrightarrow 3 \longrightarrow 9 \longrightarrow 10 \longrightarrow 5 \longrightarrow 6$	57.020	59.270	53.590	57.238	57.690	56.961	2.080
23	$4 \longrightarrow 6 \longrightarrow 7 \longrightarrow 8 \longrightarrow 2 \longrightarrow 9 \longrightarrow 5 \longrightarrow 3 \longrightarrow 10 \longrightarrow 1$	45.022	63.261	46.962	63.306	60.828	55.876	9.104
24	$10 \longrightarrow 3 \longrightarrow 7 \longrightarrow 6 \longrightarrow 1 \longrightarrow 9 \longrightarrow 5 \longrightarrow 4 \longrightarrow 2 \longrightarrow 8$	56.225	56.508	57.817	57.147	57.051	56.950	0.617
25	$5 \longrightarrow 2 \longrightarrow 9 \longrightarrow 8 \longrightarrow 4 \longrightarrow 3 \longrightarrow 1 \longrightarrow 6 \longrightarrow 10 \longrightarrow 7$	45.928	37.625	47.268	42.823	44.084	43.546	3.721
26	$3 \longrightarrow 6 \longrightarrow 8 \longrightarrow 1 \longrightarrow 10 \longrightarrow 7 \longrightarrow 4 \longrightarrow 5 \longrightarrow 2 \longrightarrow 9$	48.220	41.063	31.275	33.908	40.185	38.930	6.638
27	$3 \longrightarrow 5 \longrightarrow 8 \longrightarrow 9 \longrightarrow 4 \longrightarrow 10 \longrightarrow 1 \longrightarrow 7 \longrightarrow 6 \longrightarrow 2$	48.260	50.106	48.836	48.423	50.625	49.250	1.056
28	$10 \longrightarrow 9 \longrightarrow 2 \longrightarrow 7 \longrightarrow 4 \longrightarrow 6 \longrightarrow 3 \longrightarrow 8 \longrightarrow 5 \longrightarrow 1$	54.791	54.607	55.921	54.604	55.423	55.069	0.583
29	$7 \longrightarrow 10 \longrightarrow 5 \longrightarrow 8 \longrightarrow 3 \longrightarrow 9 \longrightarrow 1 \longrightarrow 2 \longrightarrow 6 \longrightarrow 4$	66.056	51.830	67.087	61.377	64.349	62.140	6.156
30	$1 \longrightarrow 6 \longrightarrow 3 \longrightarrow 5 \longrightarrow 4 \longrightarrow 2 \longrightarrow 7 \longrightarrow 10 \longrightarrow 9 \longrightarrow 8$	56.312	58.569	51.123	65.361	56.110	57.495	5.171
31	$10 \longrightarrow 3 \longrightarrow 7 \longrightarrow 4 \longrightarrow 1 \longrightarrow 8 \longrightarrow 9 \longrightarrow 5 \longrightarrow 6 \longrightarrow 2$	61.557	59.821	53.028	53.603	59.097	57.421	3.858
32	$4 \longrightarrow 6 \longrightarrow 1 \longrightarrow 3 \longrightarrow 10 \longrightarrow 8 \longrightarrow 9 \longrightarrow 5 \longrightarrow 7 \longrightarrow 2$	50.623	52.384	51.387	50.144	52.042	51.316	0.939
33	$3 \longrightarrow 9 \longrightarrow 4 \longrightarrow 1 \longrightarrow 2 \longrightarrow 5 \longrightarrow 10 \longrightarrow 6 \longrightarrow 8 \longrightarrow 7$	43.954	44.504	44.216	44.102	45.845	44.524	0.766
34	$6 \longrightarrow 1 \longrightarrow 5 \longrightarrow 7 \longrightarrow 8 \longrightarrow 9 \longrightarrow 10 \longrightarrow 3 \longrightarrow 2 \longrightarrow 4$	45.547	50.965	38.488	56.200	52.814	48.803	6.937
35	$1 \longrightarrow 3 \longrightarrow 10 \longrightarrow 7 \longrightarrow 2 \longrightarrow 4 \longrightarrow 9 \longrightarrow 5 \longrightarrow 8 \longrightarrow 6$	66.057	60.607	57.330	44.683	64.239	58.583	8.466
36	$1 \longrightarrow 3 \longrightarrow 7 \longrightarrow 2 \longrightarrow 6 \longrightarrow 8 \longrightarrow 4 \longrightarrow 9 \longrightarrow 10 \longrightarrow 5$	56.218	59.921	50.645	70.229	60.365	59.475	7.163
37	$8 \longrightarrow 2 \longrightarrow 9 \longrightarrow 3 \longrightarrow 7 \longrightarrow 4 \longrightarrow 10 \longrightarrow 6 \longrightarrow 1 \longrightarrow 5$	48.917	41.185	47.054	54.306	47.560	47.805	4.689
38	$4 \longrightarrow 9 \longrightarrow 6 \longrightarrow 8 \longrightarrow 10 \longrightarrow 7 \longrightarrow 1 \longrightarrow 2 \longrightarrow 5 \longrightarrow 3$	57.758	48.841	41.355	44.317	48.609	48.176	6.202
39	$7 \longrightarrow 10 \longrightarrow 2 \longrightarrow 8 \longrightarrow 3 \longrightarrow 1 \longrightarrow 6 \longrightarrow 4 \longrightarrow 9 \longrightarrow 5$	51.210	58.677	50.375	57.681	49.846	53.558	4.261
40	$10 \longrightarrow 5 \longrightarrow 7 \longrightarrow 8 \longrightarrow 6 \longrightarrow 4 \longrightarrow 3 \longrightarrow 2 \longrightarrow 9 \longrightarrow 1$	72.109	63.371	74.888	57.994	62.421	66.157	7.071
41	$9 \longrightarrow 5 \longrightarrow 6 \longrightarrow 3 \longrightarrow 8 \longrightarrow 7 \longrightarrow 10 \longrightarrow 4 \longrightarrow 2 \longrightarrow 1$	41.493	34.877	37.153	36.896	38.706	37.825	2.462
42	$8 \longrightarrow 6 \longrightarrow 7 \longrightarrow 4 \longrightarrow 5 \longrightarrow 10 \longrightarrow 1 \longrightarrow 2 \longrightarrow 3 \longrightarrow 9$	47.518	46.557	46.132	46.314	49.209	47.146	1.271
43	$1 \longrightarrow 4 \longrightarrow 8 \longrightarrow 6 \longrightarrow 9 \longrightarrow 7 \longrightarrow 2 \longrightarrow 10 \longrightarrow 3 \longrightarrow 5$	43.905	46.309	43.034	33.763	38.586	41.119	4.972
44	$7 \longrightarrow 8 \longrightarrow 9 \longrightarrow 1 \longrightarrow 4 \longrightarrow 6 \longrightarrow 10 \longrightarrow 2 \longrightarrow 3 \longrightarrow 5$	56.602	49.770	47.541	49.172	50.307	50.678	3.470
45	$9 \longrightarrow 3 \longrightarrow 8 \longrightarrow 5 \longrightarrow 6 \longrightarrow 7 \longrightarrow 10 \longrightarrow 1 \longrightarrow 2 \longrightarrow 4$	29.556	28.567	24.450	26.749	31.075	28.079	2.566
46	$10 \longrightarrow 8 \longrightarrow 1 \longrightarrow 9 \longrightarrow 4 \longrightarrow 7 \longrightarrow 6 \longrightarrow 5 \longrightarrow 3 \longrightarrow 2$	23.215	37.266	61.518	42.248	34.875	39.824	13.994

Table A.2: The optimal orders by two-step approach in Section 5

Run	Order
1	$9 \longrightarrow 3 \longrightarrow 1 \longrightarrow 8 \longrightarrow 4 \longrightarrow 5 \longrightarrow 10 \longrightarrow 2 \longrightarrow 6 \longrightarrow 7$
2	$9 \longrightarrow 3 \longrightarrow 1 \longrightarrow 4 \longrightarrow 8 \longrightarrow 5 \longrightarrow 10 \longrightarrow 2 \longrightarrow 6 \longrightarrow 7$
3	$9 \longrightarrow 3 \longrightarrow 1 \longrightarrow 4 \longrightarrow 5 \longrightarrow 8 \longrightarrow 10 \longrightarrow 2 \longrightarrow 6 \longrightarrow 7$
4	$9 \longrightarrow 3 \longrightarrow 1 \longrightarrow 4 \longrightarrow 5 \longrightarrow 10 \longrightarrow 8 \longrightarrow 2 \longrightarrow 6 \longrightarrow 7$
5	$9 \longrightarrow 3 \longrightarrow 1 \longrightarrow 4 \longrightarrow 5 \longrightarrow 10 \longrightarrow 2 \longrightarrow 8 \longrightarrow 6 \longrightarrow 7$
6	$9 \longrightarrow 3 \longrightarrow 1 \longrightarrow 4 \longrightarrow 5 \longrightarrow 10 \longrightarrow 2 \longrightarrow 6 \longrightarrow 8 \longrightarrow 7$

Note: All orders have $\hat{y}_\mu = 22.377$ and $\hat{y}_\sigma = 6.688$ with MSE=45.118.

Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 467–481

Investigation into The Robustness of Balanced Incomplete Block Designs

Janet Godolphin[1] and Helen R. Warren[2]
[1]*Department of Mathematics, University of Surrey, Surrey, UK*
[2]*Clinical Pharmacology Department, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK*

Received: 05 April 2021; Revised: 23 May 2021; Accepted: 25 May 2021

Abstract

A set of measures is developed which indicate the robustness of a Balanced Incomplete Block Design (BIBD) against yielding a disconnected eventual design in the event of observation loss. The measures have uses as a pilot procedure and as a tool to aid in design selection in situations in which significant observation loss is thought possible. The measures enable non-isomorphic BIBDs with the same parameters to be ranked. Investigation of a class of BIBDs suggests there is some correspondence between robustness against becoming disconnected and rankings associated with A-efficiency.

Key words: Connected; Efficiency; Observation loss; Optimality.

AMS Subject Classifications: 62D10, 62K05, 62K10

1. Introduction

Consider D, a binary connected incomplete block design. During experimentation, some observations may be lost and the properties of the eventual design, D_e, will be different from those of D. The eventual design may be far less efficient than the original design. In an extreme situation, D_e may be disconnected, resulting in serious damage to the aims of the experiment. When selecting a design for experimentation, it is prudent to assess the potential for observation loss to result in a disconnected design.

The universal optimality properties of BIBDs make such designs appealing when available within the practical constraints of an experiment. Where it is non-empty, the class of non-isomorphic BIBDs with v treatments arranged in b blocks of size k is denoted $\mathcal{D}(v, b, k)$. See Mathon and Rosa (1996) for sets of relatively small non-isomorphic designs. All designs in a $\mathcal{D}(v, b, k)$ have treatments replicated $r = bk/v$ times and each pair of treatments occurs together in $\lambda = r(k-1)/(v-1)$ blocks. Designs in $\mathcal{D}(v, b, k)$ are usually considered as having equal merit. In particular, optimality criteria cannot be used to distinguish between designs in a class. However, in the event of observation loss from a BIBD, the property of balance is destroyed and the A-efficiency of the eventual design may be small.

Corresponding Author: Janet Godolphin
Email: j.godolphin@surrey.ac.uk

Many authors investigate observation loss in binary connected incomplete block designs, which are not necessarily balanced. A design is said to be Criterion-1 robust against a specific pattern of observation loss, if a connected eventual design is guaranteed in the event of such observation loss. The Criterion-1 robustness of designs against the loss of t observations is investigated in Ghosh (1979), and results are given on the maximum number of blocks and the maximum number of observations that can be lost whilst ensuring a connected eventual design in Ghosh (1982) (see Kageyama (1990), for a review of other related work up to 1988). Baksalary and Tabis (1987) and Godolphin and Warren (2011) give sufficient conditions for Criterion-1 robustness against the loss of a subset of blocks. Bailey *et al.* (2013) investigate the Criterion-1 robustness of classes of universally optimal and D-optimal designs. Tsai and Liao (2013) look into the Criterion-1 robustness of designs with blocks of size two. Conditions on the number of individual observations and on the number of whole blocks that a design is Criterion-1 robust against losing, are given in terms of the E-value of the design and of the design support, in Godolphin (2016, 2019).

The A-efficiency of eventual designs following a specified pattern of observation loss, for which it is known that D_e will be connected, provides a second measure of design robustness. A design is Criterion-2 robust against a pattern of observation loss if the A-efficiency for any potential D_e is *not too small*. Dey (1993) investigates the Criterion-1 and Criterion-2 robustness of a design according to two patterns of loss: the loss of t observations on the same treatment; the loss of all observations in a single block. Lal *et al.* (2001) develop conditions for Criterion-1 robustness against the loss of any t observations, and give expressions for the A-efficiencies of eventual designs resulting from the loss of some configurations of observation pairs, any pair of blocks and for sets of disjoint blocks. Related work by Bhar (2014) advocates the advantages of the E-efficiencies of potential eventual designs as an alternative criterion to assess design robustness in the event of observation loss.

For results specific to Criterion-1 robustness of BIBDs see, for example, Ghosh (1982), where it is established that a BIBD is Criterion-1 robust against the loss of any $r-1$ observations and against the loss of any $r-1$ blocks. Key work associated with the Criterion-2 robustness of BIBDs includes Bhaumik and Whittinghill (1991), who consider the loss of complete blocks, and Whittinghill (1995) who considers the effect of losing any two observations on optimality criteria. Das and Kegeyama (1992) investigate the Criterion-2 robustness of a BIBD against observation loss in one block. Results of Lal *et al.* (2001) on observation loss in BIBDs mirror those of Whittinghill (1995), with Whittinghill's case 3 omitted. Prescott and Mansson (2001) investigate properties of eventual designs arising from the loss of observation pairs, with reference to a design in $\mathcal{D}(8, 14, 4)$. The Intersection Aberration criterion of Morgan and Parvu (2008) ranks members of $\mathcal{D}(v, b, k)$ according to efficiency properties of eventual designs arising from the loss of two blocks.

A Rank Reducing Observation Set (RROS) in D is a set of observations, the removal of which yields D_e with $\text{Rank}(X) > \text{Rank}(X_e)$, where X and X_e are the design matrices of D and D_e, respectively. In this work, the focus is on identifying the sizes and numbers of RROSs for designs in $\mathcal{D}(v, b, k)$ that are most damaging to the aims of the experiment. The work is closely aligned to the concept of Criterion-1 robustness: if D is not Criterion-1 robust against a specific pattern of observation loss, then there will be at least one set of observations corresponding to this pattern that comprise a RROS. Using an approach closely aligned to the treatment partitioning processes of Godolphin and Warren (2011), the smallest RROSs

for different treatment partitions are determined. Expressions for the measure (S_u, T_u) are developed, where S_u is the smallest number of observations in a RROS of specific type and T_u is the number of such RROSs. Observation loss is assumed to be random, that is, each observation has the same probability of being lost, independent of any other. For designs in a $\mathcal{D}(v, b, k)$ the (S_u, T_u) measure has uses:

(i) as a pilot procedure to provide information on the robustness of a design;

(ii) to aid selection of a design from a $\mathcal{D}(v, b, k)$ having cardinality greater than one.

In §2, the different types of RROS are defined and illustrated via an example. Formulae for the (S_u, T_u) measures are developed in §3. These are of two types: (S_u, T_u) depending only on v, b, k, u, which are fixed for all designs in a $\mathcal{D}(v, b, k)$; (S_u, T_u) that can vary within a $\mathcal{D}(v, b, k)$. The former can be used to assess the general robustness of designs in the class against giving rise to a disconnected D_e. The latter provide a means of design comparison and aid in design selection. In §4, results are illustrated by reference to $\mathcal{D}(8, 14, 4)$. The design ranking obtained by the (S_u, T_u) measure is found to be consistent with ranking according to worst A- and E-efficiencies according to the loss of between two and five observations, and to the Intersection Aberration criterion of Morgan and Parvu (2008).

2. Preliminaries

Consider D, a planned incomplete block design, that is both binary and connected, with n observations on v treatments in b blocks of size k. The observations are assumed to be uncorrelated each with variance σ^2, and the observation vector $\mathbf{Y}$ is assumed to follow the additive model.

$$E(\mathbf{Y}) = \mu\mathbf{1_n} + X_1\boldsymbol{\tau} + X_2\boldsymbol{\beta}.$$

Here, μ is a scalar constant, $\mathbf{1_n}$ is the vector of length n with all elements unity, and $\boldsymbol{\tau} = (\tau_1, \tau_2, \ldots, \tau_v)^T$ and $\boldsymbol{\beta}$ are vectors of the treatment and block effects. Matrices X_1 and X_2, of orders $n \times v$ and $n \times b$, relate to the treatment and block components of D, each row of X_i, $i = 1, 2$, having one element unity and remaining elements zero. The design has design matrix $X = (1_n\, X_1\, X_2)$ and $v \times v$ information matrix:

$$C = X_1^T X_1 - X_1^T X_2 (X_2^T X_2)^{-1} X_1.$$

Since D is connected, $\text{Rank}(C) = v - 1$ and the positive eigenvalues of C are expressed as:

$$0 < \mu_1 \leq \mu_2 \leq \cdots \leq \mu_{v-1}.$$

Any RROS of D can be categorised as being of Types I to III. These types are not mutually exclusive. Brief details are given below.

Type I: If observations comprising a Type I RROS are lost from D then $\mathcal{B}_e$, the set of blocks of D_e, can be partitioned into non-empty sets $\mathcal{B}_0$ and $\mathcal{B}_e \setminus \mathcal{B}_0$ with the treatments in $\mathcal{B}_0$ being distinct from those in $\mathcal{B}_e \setminus \mathcal{B}_0$.

Type II: A Type II RROS contains all observations from one or more blocks.

Type III: A Type III RROS contains all replicates of one or more treatments.

The most extreme consequence of observation loss arises if not all $v(v-1)/2$ pairwise treatment contrasts are estimable from D_e. The loss of a RROS that is of Type II only will not restrict the capacity to estimate treatment contrasts, but will affect the overall efficiency. Such RROSs are not the focus of this work. However, in the event of the loss of observations comprising a RROS that is Type I and/or Type III then $\text{Rank}(C_e) < v-1$, where C_e denotes the information matrix of D_e, and not all treatment contrasts will be estimable. A Type III RROS contains all replicates of a subset of v_0 treatments. Such RROSs are immediately evident from the treatment replications of D. In the event that a RROS that is Type III and not Type I is lost from D, then $\text{Rank}(C_e) = v - v_0 - 1$ and all contrasts in the $v - v_0$ treatments occurring in D_e will be estimable. If a Type I RROS is lost from D then no pairwise treatment contrast involving one treatment occurring in a block of $\mathcal{B}_0$ and one occurring in a block of $\mathcal{B}_e \setminus \mathcal{B}_0$ will be estimable. The available data comprise two observation sets which cannot be analysed as a single entity, although they can be analysed separately, see Searle (1971, §7.4), to gain limited information.

Whilst the aims of the experiment are seriously compromised by the loss of a Type I or a Type III RROS from D, the Type I RROSs are not easily identifiable from the planned design and these are the main focus of this work. For more extensive discussion of the types of RROS see Godolphin and Warren (2014). Type I and Type III RROSs are demonstrated in the following example.

Example: The design D has seven treatments, each with replication three, arranged in seven blocks of size three. The design is depicted below, with columns as blocks numbered 1 to 7.

$$D = \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline 1 & 1 & 1 & 2 & 2 & 3 & 4 \\ 2 & 3 & 3 & 5 & 5 & 4 & 5 \\ 7 & 4 & 6 & 7 & 7 & 6 & 6 \end{array}$$

The following six potential eventual designs, labelled $D_{e1}, \ldots, D_{e6}$, result from different configurations of observation loss, with each pattern of loss corresponding to a RROS of Type I and/or Type III.

$$D_{e1} = \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline 1 & 1 & 1 & * & * & 3 & 4 \\ * & 3 & 3 & 5 & 5 & 4 & 5 \\ 7 & * & 6 & 7 & 7 & 6 & 6 \end{array} \quad D_{e2} = \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline 1 & 1 & 1 & 2 & 2 & * & 4 \\ 2 & * & * & 5 & 5 & 4 & 5 \\ 7 & 4 & 6 & 7 & 7 & 6 & 6 \end{array} \quad D_{e3} = \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline * & 1 & 1 & 2 & 2 & 3 & 4 \\ 2 & 3 & 3 & 5 & 5 & 4 & * \\ 7 & 4 & 6 & 7 & 7 & 6 & 6 \end{array}$$

$$D_{e4} = \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline 1 & 1 & 1 & 2 & 2 & 3 & 4 \\ * & 3 & 3 & 5 & 5 & 4 & * \\ * & 4 & 6 & 7 & 7 & 6 & 6 \end{array} \quad D_{e5} = \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline * & 1 & 1 & 2 & 2 & 3 & 4 \\ 2 & * & 3 & 5 & 5 & 4 & * \\ 7 & 4 & 6 & 7 & 7 & 6 & 6 \end{array} \quad D_{e6} = \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline * & 1 & 1 & 2 & 2 & 3 & 4 \\ 2 & 3 & 3 & * & * & 4 & * \\ 7 & 4 & 6 & 7 & 7 & 6 & 6 \end{array}$$

The RROSs giving rise to D_{e1} and D_{e2} are Type III only. In both cases all replicates of one treatment have been lost. The eventual designs are connected designs in six treatments.

Thus $\text{Rank}(C_{e1}) = \text{Rank}(C_{e2}) = 5$. All pairwise treatment contrasts in the six treatments of D_{e1} and D_{e2} are estimable. The RROS giving rise to D_{e2} is minimal in the sense that, if any observation of the RROS is reinstated then the resulting design is a connected design in seven treatments. The RROS of D_{e2} is not minimal: if the replicate of treatment 4 in block 2 is reinstated then the observation loss incurred still corresponds to a Type III RROS.

The RROSs giving rise to D_{e3} to D_{e5} are Type I only. All three eventual designs are disconnected with the blocks partitioned so that each block contains treatments from exactly one set of $\{1, 3, 4, 6\}$ and $\{2, 5, 7\}$. The eventual designs have $\text{Rank}(C_{e3}) = \text{Rank}(C_{e4}) = \text{Rank}(C_{e5}) = 5$. Pairwise treatment contrasts are estimable within each treatment set, but the 12 pairwise treatment contrasts involving one treatment from each set, such as $\tau_1 - \tau_2$, are not estimable. The Type I RROSs leading to D_{e3} and D_{e4}, containing two and three observations respectively, are minimal. The Type I RROS leading to D_{e3} is of particular interest, since no smaller Type I RROS exists for D. The Type I RROS leading to D_{e5} is not minimal: reinstatement of the replicate of treatment 3 in block 2 gives D_{e3}.

Finally, the RROS giving rise to D_{e6} is both Type I and Type III. The eventual design D_{e6} is disconnected with block partition such that treatments in $\{2, 7\}$ are contained in $\mathcal{B}_0$, say, and treatments in $\{1, 3, 4, 6\}$ are contained in $\mathcal{B}_e \setminus \mathcal{B}_0$. Only seven pairwise treatment contrasts are estimable. The eventual design has $\text{Rank}(C_{e6}) = 4$.

Some basic properties can be established for minimal Type I RROSs, such as those RROSs leading to D_{e3} and D_{e4} in Example 1.

Theorem 1: Consider a Type I RROS for design D such that no subset is also a RROS. Then, the eventual design D_e arising due to the loss of the RROS from D has the properties:

(i) D_e, has b blocks;

(ii) D_e, has v treatments.

Proof: By the definition of a Type I RROS, the blocks of D_e, can be partitioned into non-empty sets $\mathcal{B}_0$ and $\mathcal{B}_e \setminus \mathcal{B}_0$ with the treatments in blocks of $\mathcal{B}_0$ being distinct from those in the blocks of $\mathcal{B}_e \setminus \mathcal{B}_0$. For (i): assume D_e has fewer than b blocks. Then the RROS contains all observations from at least one block of D. Reinstate any one observation in such a block to form $D_e^\dagger$. The additional block of $D_e^\dagger$, over those of D_e, can be allocated to one of $\mathcal{B}_0$ and $\mathcal{B}_e \setminus \mathcal{B}_0$ to form a partition in $D_e^\dagger$. Thus the observation loss resulting in $D_e^\dagger$ corresponds to a Type I RROS, but the missing observations are a subset of those lost to form D_e. This provides a contradiction and it follows that D_e has b blocks, as required. For (ii), assume D_e has fewer than v treatments. Then the RROS contains all replicates of one or more treatments of D. Reinstate one observation of such a treatment to form $D_e^{\dagger\dagger}$. By (i) the reinstated observation will be in a block in either $\mathcal{B}_0$ and $\mathcal{B}_e \setminus \mathcal{B}_0$. In either case, there is partition in $D_e^{\dagger\dagger}$. As with (i), the observations lost from D to form $D_e^{\dagger\dagger}$ comprise a Type I RROS, but are a subset of those lost to form D_e. This gives a contradiction. Hence, D_e has v treatments, as required.

Such RROSs are summarised in the following definition, where $\mathcal{V}$, $\mathcal{B}$ denote the sets of treatments and blocks of D, respectively:

Definition: A RROS(u) for D, for $1 \leq u \leq v/2$, is a Type I RROS with the following properties:

(i) no subset is also a RROS;

(ii) in D_e, the treatments are partitioned into sets $\mathcal{V}_u$, $\mathcal{V}_{v-u}$, with cardinalities u and $v-u$, and the blocks are partitioned into non-empty sets $\mathcal{B}_1$, $\mathcal{B}_2$, with treatments from $\mathcal{V}_u$ arranged exclusively in blocks of $\mathcal{B}_1$ and those from $\mathcal{V}_{v-u}$ exclusively in blocks of $\mathcal{B}_2$.

The partitioning of $\mathcal{V}$ and $\mathcal{B}$ induced by the loss of a Type I RROS is termed a *consistent treatment and block partition.*

3. Robustness Measures for BIBDs

Henceforth, any planned design, will be taken as being a BIBD, that is, $D \in \mathcal{D}(v, b, k)$. It is evident that D contains v Type III RROSs of size r. The aim is to add to this information by finding RROS(u)s of smallest size for $1 \leq u \leq v/2$. From Ghosh (1982), D is Criterion-1 robust against the loss of any $r-1$ observations. Thus, a RROS(u) must be of size at least r. A RROS(u) consists of all replicates of treatments in $\mathcal{V}_{v-u}$ contained in blocks of B_1 and all replicates of treatments in $\mathcal{V}_u$ contained in blocks of B_2. The consequence of losing all observations in the RROS(u) is that in $\mathcal{B}_1$ only treatments from $\mathcal{V}_u$ are preserved and it is precisely treatments in $\mathcal{V}_u$ that are lost from $\mathcal{B}_2$. For a given $\mathcal{V}_u$, denote the smallest number of observations in a RROS(u) by s_u. Further, define S_u to be $\min_{\mathcal{V}_u}\{s_u\}$, the minimisation being over all $v!/[u!(v-u)!]$ sets of u treatments, and define T_u to be the number of RROS(u)s of size S_u. The pair (S_u, T_u) forms the robustness measure. It gives the smallest number of observations that must be lost, and the number of observation sets of this size, for the possibility of an eventual design with a consistent treatment and block partition, with the treatment sets being of sizes u and $v-u$.

Relationships associated with the distribution of subsets of treatments in $\mathcal{V}_u$ amongst the blocks of D arise as a consequence of the properties of BIBDs. These relationships are given below without proof.

Theorem 2: For given $\mathcal{V}_u$, let b_j be the number of blocks in D containing exactly j elements from $\mathcal{V}_u$, for $0 \leq j \leq w$, where $w = \min\{u, k\}$. Then:

$$\sum_{j=0}^{w} b_j = b \tag{1}$$

$$\sum_{j=0}^{w} j b_j = ur \tag{2}$$

$$\sum_{j=0}^{w} \binom{j}{2} b_j = \binom{u}{2}\lambda. \tag{3}$$

Corollary 1: For any $D \in \mathcal{D}(v, b, k)$:

(i) Each of the v sets $\mathcal{V}_1$ has $(b_0, b_1) = (b-r, r)$;

(ii) Each of the $v(v-1)/2$ sets $\mathcal{V}_2$ has $(b_0, b_1, b_2) = (b - 2r + \lambda, 2r - 2\lambda, \lambda)$.

For any $D \in \mathcal{D}(v, b, 2)$:

(iii) Every $\mathcal{V}_u$ has $(b_0, b_1, b_2) = (b - ur + u(u-1)\lambda/2, ur - u(u-1)\lambda, u(u-1)\lambda/2)$.

Proof: (i) follows through application of (1) and (2) with $w = 1$. Similarly, (ii) and (iii) follow through use of (1) to (3), with $u = 2$, $k \geq 2$ for (ii), and $u \geq 2$, $k = 2$ for (iii).

From Corollary 1, the values of b_j are dependent only on the design parameters and u for $w = \min\{u, k\} \leq 2$. For many $D \in \mathcal{D}(v, b, k)$, the elements of $(b_0, b_1, \ldots, b_w)$ will depend on the particular $\mathcal{V}_u$, for $w \geq 3$. For example, some sets of three treatments may occur together in more blocks than other sets. Consider $D \in \mathcal{D}(v, b, k)$, with $k \geq 3$. Using (1), (2) and (3), a given $\mathcal{V}_3$ has:

$$(b_0, b_1, b_2, b_3) = (b - 3r + 3\lambda - b_3,\ 3r - 6\lambda + 3b_3,\ 3\lambda - 3b_3,\ b_3). \tag{4}$$

Further, for given u, the distributions of $(b_0, b_1, \ldots, b_w)$ may differ between designs within a $\mathcal{D}(v, b, k)$.

We now use the properties of $\mathcal{D}(v, b, k)$ design classes to obtain expressions for the (S_u, T_u) measures.

3.1. (S_u, T_u) measures for $\mathcal{D}(v, b, 2)$

Any non-empty $\mathcal{D}(v, b, 2)$ has cardinality one and thus the (S_u, T_u) measures provide a pilot process to check the Criterion-1 robustness properties of the design.

Let $D \in \mathcal{D}(v, b, 2)$. First consider the trivial case $u = 1$. For any $\mathcal{V}_1$, exactly r of the blocks of D contain the treatment in $\mathcal{V}_1$. There are $2^r - 1$ ways of selecting one observation from each of these blocks to form a RROS(1), that is, to yield a treatment disconnected eventual design in v treatments. There are v ways of selecting $\mathcal{V}_1$. Thus $S_1 = r$ and $T_1 = v(2^r - 1)$. Now consider any set $\mathcal{V}_u$ with $2 \leq u \leq v/2$. A RROS(u) is formed by selecting one observation from each of the b_1 blocks containing one element from $\mathcal{V}_u$. Using Corollary 1 (iii), $b_1 = ur - u(u-1)\lambda = ru(v-u)/(v-1)$. This is independent of the particular set of u treatments, indicating that $s_u = ur(v-u)/(v-1)$, for every $\mathcal{V}_u$. There are $v!/(u!(v-u)!)$ sets of u treatments. It follows that the robustness measures are:

$$(S_1, T_1) = (r, v(2^r - 1)) \tag{5}$$

$$(S_u, T_u) = \left(\frac{ru(v-u)}{v-1}, \frac{2^{S_u} v!}{u!(v-u)!} \right), \quad \text{for } 2 \leq u \leq v/2. \tag{6}$$

From (5) and (6) S_u increases monotonically with u. Thus a pilot procedure starts by evaluation of S_1. Hence, in addition to the v Type III RROSs of size r, by (5) there are $v(2^r - 1)$ Type I RROSs, also of size r. Then, by (6) there are many Type I RROSs of size $2r(v-2)/(v-1)$, and so on.

3.2. (S_u, T_u) measures for $\mathcal{D}(v, b, 3)$

Many $\mathcal{D}(v, b, 3)$ design classes have cardinality greater than one. For example, $\mathcal{D}(7, 14, 3)$, $\mathcal{D}(7, 21, 3)$ and $\mathcal{D}(7, 28, 3)$ have cardinalities 4, 10 and 35 respectively.

Consider $D \in \mathcal{D}(v, b, 3)$. For any set $\mathcal{V}_1$, exactly r blocks of D contain the treatment in $\mathcal{V}_1$. For a valid RROS(1), the eventual design must contain all v treatments, thus a smallest RROS(1) contains the replicate of the treatment of $\mathcal{V}_1$ from $r-1$ blocks and the replicates of the two treatments from $\mathcal{V}_{v-1}$ from the rth block. This gives $s_1 = r + 1$, independent of the particular $\mathcal{V}_1$. For given $\mathcal{V}_1$ there are r ways of selecting the replicate of the treatment in $\mathcal{V}_1$ that is preserved in D_e and there are v ways of selecting $\mathcal{V}_1$. Thus, $S_1 = r + 1$ and $T_1 = rv$. Now let $\mathcal{V}_u$ be any set with $2 \le u \le v/2$. For $\mathcal{V}_u$, a RROS(u) of smallest size comprises the observation of a treatment contained in $\mathcal{V}_u$ from each of the b_1 blocks containing one element of $\mathcal{V}_u$ and the observation of a treatment in $\mathcal{V}_{v-u}$ from each of the b_2 blocks containing two elements of $\mathcal{V}_u$. Using (2) and (3), $s_u = b_1 + b_2 = ur - \lambda u(u-1)/2 = ru(v-u)/(v-1)$. Again, the value of s_u is independent of the particular $\mathcal{V}_u$. Given $\mathcal{V}_u$, the RROS(u) of size s_u is unique. Thus, the robustness measures for $\mathcal{D}(v, b, 3)$ are:

$$(S_1, T_1) = (r + 1, rv) \tag{7}$$

$$(S_u, T_u) = \left(\frac{ru(v-u)}{v-1}, \frac{v!}{u!(v-u)!} \right), \quad \text{for } 2 \le u \le v/2. \tag{8}$$

As with $k = 2$, the value of S_u increases with u for $1 \le u \le v/2$.

Results for $k = 3$ merit special comment. In some $\mathcal{D}(v, b, 3)$, designs in the class differ in the number of repeated blocks. For example, the four designs in $\mathcal{D}(7, 14, 3)$ have support sizes (i.e. number of distinct blocks) of 7, 11, 13 and 14 respectively. Several authors, including Bhaumik and Whittinghill (1991), recommend avoiding BIBDs with repeated blocks when observation loss is possible. Also, see Raghavarao *et al.* for an investigation of designs in $\mathcal{D}(7, 21, 3)$ with emphasis on the relationship between the support size and the estimation of contrasts of the block effects. Conversely, Foody and Hedayat (1977) discuss some experimental situations in which deliberate use of designs with repeated blocks gives practical advantages. In assessing robustness within a $\mathcal{D}(v, b, 3)$ via (S_u, T_u) measures, no advantage is gained by the avoidance of designs with repeated blocks, since the formulae of (7) and (8) are the same for all designs in a class. Thus, all designs in a $\mathcal{D}(v, b, 3)$ are equally vulnerable to becoming disconnected through random observation loss.

3.3. (S_u, T_u) measures for $\mathcal{D}(v, b, k)$, with $k \ge 4$

Let $D \in \mathcal{D}(v, b, k)$, with $k \ge 4$.

For $1 \le u < k/2$, choose any u treatments from any one block for $\mathcal{V}_u$. From the same block, select the $k - u$ treatments in $\mathcal{V}_{v-u}$ for removal. From the $b - 1$ remaining blocks, select the treatments from $\mathcal{V}_u$ for removal. The $(k - u) + (r - 1)u = u(r - 2) + k$ selected observations comprise a RROS(u) and, by the process used, there is no smaller RROS(u) for that $\mathcal{V}_u$, giving $s_u = (k - u) + (r - 1)u = u(r - 2) + k$. The value of s_u does not depend on the chosen block or on the treatments used from the block for $\mathcal{V}_u$. Thus

$$(S_u, T_u) = \left(u(r-2) + k, \frac{k!b}{u!(k-u)!} \right), \quad \text{for } 1 \le u < k/2. \tag{9}$$

Now, for even k, consider $u = k/2$. For a $\mathcal{V}_{k/2}$ with $b_{k/2} > 0$, select the observations from $\mathcal{V}_{v-k/2}$ for removal from at least one of the $b_{k/2}$ blocks containing all $k/2$ treatments of $\mathcal{V}_{k/2}$. From all other blocks select the observations from $\mathcal{V}_{k/2}$ for removal. Every observation on a treatment in $\mathcal{V}_{v-k/2}$ occurs in a block with no more than $k/2$ treatments from $\mathcal{V}_{k/2}$, thus, the selected observations comprise a RROS($k/2$) of minimal size: $s_{k/2} = kr/2$. For the particular $\mathcal{V}_{k/2}$, there will be $2^{b_{k/2}} - 1$ RROS($k/2$)s of this size, giving

$$(S_{k/2}, T_{k/2}) = \left(\frac{kr}{2}, \sum_{\Psi_0} \left(2^{b_{k/2}} - 1 \right) \right), \tag{10}$$

where, the summation is over Ψ_0, the set of $\mathcal{V}_{k/2}$ sets with $b_{k/2} > 0$.

Now consider $k/2 < u \leq k$. For every $\mathcal{V}_u$, perform a scan of D in the following way. For blocks containing fewer than $k/2$ members of $\mathcal{V}_u$, select the members of $\mathcal{V}_u$ for removal. Conversely, for blocks containing more than $k/2$ members of $\mathcal{V}_u$, select the members of $\mathcal{V}_{v-u}$ for removal. For even k, for blocks containing exactly $k/2$ members of $\mathcal{V}_u$, select either treatment set for removal. Let the number of selected observations be N. Then

$$N = \sum_{i=1}^{[k/2]} i b_i + \sum_{i=[k/2]+1}^{k-1} (k-i) b_i,$$

where $[k/2]$ denotes the integer part of $k/2$. Using (2):

$$N = ur - \sum_{i=[k/2]+1}^{k-1} (2i-k) b_i. \tag{11}$$

The selected observations form a RROS of Type I and/or Type III. Any $\mathcal{V}_u$ with $b_j = 0$ for all $j > k/2$ has $N = ur$, and any $\mathcal{V}_u$ with $b_j > 0$ for at least one $j > k/2$ has $N < ur$. Thus, for any $\mathcal{V}_u$ yielding the minimum value for N, there is at least one block containing more than $k/2$ treatments from $\mathcal{V}_u$ in D. Suppose a RROS obtained by the scan for a $\mathcal{V}_u$ for which N is minimised comprises a Type III RROS. Call this $\mathcal{V}_u$ set $\mathcal{V}_u^*$. Then all replicates of a member of $\mathcal{V}_u^*$, say u_0, are selected by the scan and at least one treatment from $\mathcal{V} \setminus \mathcal{V}_u^*$, say u_1, is in a block containing more than $k/2$ treatments from $\mathcal{V}_u^*$. Now consider the set $\mathcal{V}_u^\dagger$ with u_1 replacing u_0 but with the other $u-1$ treatments common to those of $\mathcal{V}_u^*$. This has smaller N, providing a contradiction. Thus, $\mathcal{V}_u$ sets corresponding to the smallest value of N only yield RROS(u)s by the scan, and, by the process, no smaller RROS(u) exists for that $\mathcal{V}_u$. It follows that $S_u = \min_{\mathcal{V}_u}\{N\} < ur$. Let Ψ_1 be the set of $\mathcal{V}_u$ sets achieving S_u. Then $T_u = \sum_{\Psi_1} 2^{b_{k/2}}$, where $b_{k/2}$ is taken to be zero for odd k. Thus,

$$(S_u, T_u) = \left(ur - \max_{\mathcal{V}_u} \left\{ \sum_{i=[k/2]+1}^{k-1} (2i-k) b_i \right\}, \sum_{\Psi_1} 2^{b_{k/2}} \right), \quad \text{for} \quad k/2 < u \leq k.$$

Now consider $k < u \leq v/2$. As for $k/2 < u \leq k$, the approach is to conduct a scan for each $\mathcal{V}_u$ and to obtain N as given by (11). However, in this case the minimum value of N can arise for sets of selected observations corresponding to Type III RROSs. An additional step is required in order to identify RROS(u)s of smallest size.

Definition: A covering for $\mathcal{V}_u$ comprises two sets of blocks from D, denoted $\mathcal{B}_{1c}$ and $\mathcal{B}_{2c}$, such that together the blocks of $\mathcal{B}_{1c}$ contain all the treatments of $\mathcal{V}_u$, and together the blocks of $\mathcal{B}_{2c}$ contain all the treatments of $\mathcal{V}_{v-u}$. The weight of the covering is

$$\sum_{\substack{j=1 \\ \mathcal{B}_{1c}}}^{[k/2]-1} (k-2b_j) + \sum_{\substack{j=[k/2]+1 \\ \mathcal{B}_{2c}}}^{k-1} (2b_j - 1).$$

Consider a scan of D conducted with treatment set $\mathcal{V}_u$ in the usual way, but with an adjustment for the blocks of a covering. Treatments from $\mathcal{V}_{v-u}$ are selected from the blocks of $\mathcal{B}_{1c}$, and treatments from $\mathcal{V}_u$ are selected from the blocks of $\mathcal{B}_{2c}$, regardless of the numbers of treatments from $\mathcal{V}_u$ in blocks of either set. Then, the number of observations selected in total exceeds N by the weight of the covering. For a given $\mathcal{V}_u$, let W be the minimum weight of all coverings for $\mathcal{V}_u$. Then, the RROS(u) of smallest size for that $\mathcal{V}_u$ contains $W + N$ observations. These are: observations from $\mathcal{V}_{v-u}$ in blocks of $\mathcal{B}_{1c}$; observations from $\mathcal{V}_u$ in blocks of $\mathcal{B}_{2c}$; observations selected from the scan in the usual way for all other blocks. Thus $S_u = \min_{\mathcal{V}_u}\{W + N\}$. Let Ψ_2 be the set of $\mathcal{V}_u$ sets achieving S_u. Then

$$(S_u, T_u) = \left(\min_{\mathcal{V}_u} \{W + N\}, \sum_{\Psi_2} 2^{b_{k/2}} \right), \quad \text{for} \quad k < u \leq v/2. \tag{12}$$

3.4. A lower bound for S_u

For $k < u \leq v/2$, the process of obtaining minimal coverings for each $\mathcal{V}_u$, before running the scan, to obtain (S_u, T_u) via (12) can be computer intensive. The following result gives a lower bound for S_u. For u moderate in size, the magnitude of this bound might indicate that S_u is sufficiently large that the identification of the exact value is not of concern, given understanding of the expected level of observation loss for the particular experimental situation.

Theorem 3: For $D \in \mathcal{D}(v, b, k)$ and $1 \leq u \leq v/2$, a lower bound for S_u is given by:

$$\left\lceil \frac{u(v-u)r}{v-1} \right\rceil$$

Proof: For any set $\mathcal{V}_u$, the sum of concurrences between treatments in $\mathcal{V}_u$ and treatments in $\mathcal{V}_{v-u}$ is $u(v-u)\lambda$. To induce a Type I RROS through observation loss, the concurrence between any treatment in $\mathcal{V}_u$ and a treatment in $\mathcal{V}_{v-u}$ must be reduced to zero. The greatest reduction in the sum of the concurrences between treatments in $\mathcal{V}_u$ and $\mathcal{V}_{v-u}$ caused through the loss of a single observation occurs if the observation is in a block containing exactly one or $k-1$ treatments from $\mathcal{V}_u$. The loss of such an observation reduces the sum of the concurrences by $k-1$. Thus the number of observations in a RROS(u) is at least

$$\frac{u(v-u)\lambda}{k-1} = \frac{u(v-u)r}{v-1}$$

as required.

4. Investigation of Designs in $\mathcal{D}(8, 14, 4)$

We use the $\mathcal{D}(8, 14, 4)$ design class, which has cardinality four, to demonstrate the results of §3, and compare the design ranking produced with design comparisons focused on Criterion-2 robustness. A set of four non-isomorphic designs in $\mathcal{D}(8, 14, 4)$ is obtained by combining pairs of four base designs. Base designs D_a and D_b contain treatments $1, 2, \ldots, 7$ and base designs D_c and D_d contain treatments $1, 2, \ldots, 8$. Each base design is obtained via a cyclic construction, modulo 7: D_a and D_b are members of $\mathcal{D}(7, 7, 4)$ with initial blocks containing $1, 3, 4, 5$ and $1, 2, 3, 5$, respectively; D_c has initial block containing $1, 2, 4$ and each block is augmented with treatment 8; D_d is obtained from D_c with treatments 1 and 2 interchanged. The base designs are displayed below:

$$D_a = \begin{array}{|cccccccc|}\hline 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 4 & 5 & 6 & 7 & 1 & 2 \\ 4 & 5 & 6 & 7 & 1 & 2 & 3 \\ 5 & 6 & 7 & 1 & 2 & 3 & 4 \\ \hline \end{array} \qquad D_b = \begin{array}{|cccccccc|}\hline 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 3 & 4 & 5 & 6 & 7 & 1 \\ 3 & 4 & 5 & 6 & 7 & 1 & 2 \\ 5 & 6 & 7 & 1 & 2 & 3 & 4 \\ \hline \end{array}$$

$$D_c = \begin{array}{|cccccccc|}\hline 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 3 & 4 & 5 & 6 & 7 & 1 \\ 4 & 5 & 6 & 7 & 1 & 2 & 3 \\ 8 & 8 & 8 & 8 & 8 & 8 & 8 \\ \hline \end{array} \qquad D_d = \begin{array}{|cccccccc|}\hline 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline 2 & 1 & 3 & 4 & 5 & 6 & 7 \\ 1 & 3 & 4 & 5 & 6 & 7 & 2 \\ 4 & 5 & 6 & 7 & 2 & 1 & 3 \\ 8 & 8 & 8 & 8 & 8 & 8 & 8 \\ \hline \end{array}$$

Members of $\mathcal{D}(8, 14, 4)$, denoted by $D1, D2, D3$ and $D4$, comprise the base design pairs:

$D1$: D_a and D_d, $\quad D2$: D_b and D_d, $\quad D3$: D_b and D_c, $\quad D4$: D_a and D_c

The labelling of the designs as $D1$ to $D4$ is consistent with Morgan and Parvu (2008). Design $D3$ is the design given careful consideration in Prescott and Mansson (2001).

4.1. (S_u, T_u) measures for $\mathcal{D}(8, 14, 4)$

Robustness measures for $u = 1$ and $u = 2$ are common to all designs in the class. By (9), $(S_1, T_1) = (9, 56)$. Every $\mathcal{V}_2$ in each design has $b_2 = \lambda = 3$, giving $(S_2, T_2) = (14, 196)$, by (10). These measures indicate the extent of observation loss required from designs in $\mathcal{D}(8, 14, 4)$ to result in an eventual design in which the treatments are partitioned into sets of size one and seven, and into sets of size two and six, respectively. The lowest value of u enabling discrimination between the four designs is $u = 3$. The measure (S_3, T_3) is different for each design and ranks the designs in terms of robustness against incurring a RROS(3). For each design and each $\mathcal{V}_3$, (11) gives $s_3 = 3r - 2b_3 = 21 - 2b_3$. Designs $D1$ to $D3$ each have some $\mathcal{V}_3$ sets with $b_3 = 2$. For example $D1$ has $b_3 = 2$ for the sets $\{1, 3, 5\}$, $\{1, 6, 7\}$, $\{2, 3, 7\}$ and $\{2, 5, 6\}$. Thus, designs $D1$ to $D3$ each have $S_3 = 17$. By contrast, $D4$ has $b_3 = 1$ for every set $\mathcal{V}_3$, giving $S_3 = 19$. The values of T_3 depend on the numbers of $\mathcal{V}_3$ with maximum b_3. Using (4),

$$(b_0, b_1, b_2, b_3) = (2 - b_3,\ 3 + 3b_3,\ 9 - 3b_3,\ b_3). \tag{13}$$

For designs $D1$ to $D3$, the $\mathcal{V}_3$ sets with $b_3 = 2$ each have $b_2 = 3$, by (13) and contribute $2^3 = 8$ to T_3. For $D4$, each $\mathcal{V}_3$ set has $b_2 = 6$, and contributes 64 to T_3. The (S_3, T_3) measures

Table 1: (S_3, T_3) measures for designs in $\mathcal{D}(8, 14, 4)$

Design	$\max_{\mathcal{V}_3} b_3$	(S_3, T_3)	Rank
$D1$	2	$(17, 32)$	2
$D2$	2	$(17, 48)$	3
$D3$	2	$(17, 56)$	4
$D4$	1	$(19, 3584)$	1

are displayed in Table 1. Design $D4$ is ranked highest with $S_3 = 19$. The other designs have $S_3 = 17$ and are ranked according to T_3. To summarise, of designs in $\mathcal{D}(8, 14, 4)$, design $D4$ is the most robust against becoming disconnected through a consistent treatment and block partition with treatments separated into sets of sizes three and five. Two more observations need to be lost from $D4$ than from the other designs, before there is a possibility of a D_e with a consistent block and treatment partition, with the treatments partitioned into sets of cardinalities 3 and 5.

4.2. A- and E-efficiencies of eventual designs

It would be hoped that the loss of as many as 17 observations from a design in $\mathcal{D}(8, 14, 4)$ would be considered a remote possibility in most experimental situations. It is interesting to investigate the quality of eventual designs arising from the loss of much smaller numbers of observations from $D1$ to $D4$. To compare the designs with regards to Criterion-2 robustness, the A-efficiencies of eventual designs are considered and, in line with suggestions of Bhar (2014), the E-efficiencies are also obtained.

For $D \in \mathcal{D}(v, b, k)$, all non-zero eigenvalues of C are $v\lambda/k$. Let D_e be a connected eventual design arising from the loss of observations from D and let the eigenvalues of C_e be $0 < \mu_{1e} \leq \mu_{2e} \leq \cdots \leq \mu_{(v-1)e}$. The A- and E-efficiencies of D_e, denoted $E_A(D_e)$ and $E_E(D_e)$, have formulae:

$$E_A(D_e) = \frac{\sum_{i=1}^{v-1} \frac{1}{\mu_i}}{\sum_{i=1}^{v-1} \frac{1}{\mu_{ie}}} = \frac{(v-1)^2 k}{vr(k-1)\sum_{i=1}^{v-1} \frac{1}{\mu_{ie}}} \quad \text{and} \quad E_E(D_e) = \frac{\mu_{1e}}{\mu_1}.$$

Hence, for designs in $\mathcal{D}(8, 14, 4)$, the A- and E-efficiencies of D_e are $E_A(D_e) = 7/(6\sum_{i=1}^{v-1} \frac{1}{\mu_{ie}})$ and $E_E(D_e) = \mu_{1e}/6$. In Table 2 results are given on the lowest A- and E-efficiencies of D_e arising from the loss of up to five observations from designs in $\mathcal{D}(8, 14, 4)$. As established in the literature, for example see Whittinghill (1995), all designs are equivalent when only one observation is lost. For the loss of between 2 and 5 observations the ranking of $D1$ to $D4$, according to the lowest A- and E-efficiences of eventual designs, is consistent with the design ranking according to (S_3, T_3). Design $D4$ consistently demonstrates better performance. Designs $D1$ to $D3$ have the same values for the lowest A- and E-efficiences, but the number of eventual designs with worst properties is consistent with T_3 measure. It is notable that, in each case, the eventual designs with lowest A-efficiency are exactly those with lowest E-efficiency.

4.3. Intersection Aberration

The Intersection Aberration criterion of Morgan and Parvu (2008) extends results of Bhaumik and Whittinghill (1991) to enable the comparison of designs in a $\mathcal{D}(v, b, k)$

Table 2: Smallest A-efficiency values following the loss of up to five observations from designs in $\mathcal{D}(8, 14, 4)$

Design	No.of missing observations	min{A-efficiency}	min{E-efficiency}	No. of eventual designs
$D1$	1	0.9722	0.8333	36
$D2$	1	0.9722	0.8333	36
$D3$	1	0.9722	0.8333	36
$D4$	1	0.9722	0.8333	36
$D1$	2	0.9354	0.6806	12
$D2$	2	0.9354	0.6806	18
$D3$	2	0.9354	0.6806	21
$D4$	2	0.9373	0.6944	168
$D1$	3	0.8885	0.5462	36
$D2$	3	0.8885	0.5462	54
$D3$	3	0.8885	0.5462	63
$D4$	1	0.8909	0.5556	280
$D1$	4	0.8216	0.4096	36
$D2$	4	0.8216	0.4096	54
$D3$	4	0.8216	0.4096	63
$D4$	4	0.8249	0.4167	280
$D1$	5	0.7155	0.2722	12
$D2$	5	0.7155	0.2722	18
$D3$	5	0.7155	0.2722	21
$D4$	5	0.7206	0.2778	168

according to lowest A-efficiency on the loss of two blocks. This criterion enables the ranking of designs within a $\mathcal{D}(v, b, k)$ in order of their robustness against suffering the most damage on the loss of any two blocks. For $D \in \mathcal{D}(v, b, k)$, let $\eta_g(D)$ denote the number of pairs of blocks such that blocks in a pair have exactly g common treatments. These design properties can be summarised by the intersection aberration vector $\boldsymbol{\eta}(D) = (\eta_0(D), \eta_1(D), \cdots, \eta_k(D))$. Following Morgan and Parvu (2008):

Definition: Let designs $D\dagger$, $D\ddagger \in \mathcal{D}(v, b, k)$, and let p be the largest integer such that $\eta_p(D\dagger) \neq \eta_p(D\ddagger)$. Then $D\dagger$ is described as having less intersection aberration than $D\ddagger$ if $\eta_p(D\dagger) < \eta_p(D\ddagger)$.

A design with less intersection aberration has greater Criterion-2 robustness against the loss of two blocks than one with more intersection aberration.

Designs in $\mathcal{D}(8, 14, 4)$, investigated in Morgan and Parvu (2008), have intersection aberration vectors:

$$\begin{aligned}
\boldsymbol{\eta}(D1) &= (3, 12, 72, 4, 0) \\
\boldsymbol{\eta}(D2) &= (1, 18, 66, 6, 0) \\
\boldsymbol{\eta}(D3) &= (0, 21, 63, 7, 0) \\
\boldsymbol{\eta}(D4) &= (7, 0, 84, 0, 0)
\end{aligned}$$

The designs are ranked by Intersection Aberration from most to least robust in the order $D4$, $D1$, $D2$, $D3$, according to $\eta_3(.)$ values. This ranking is consistent with the ranking obtained by consideration of (S_3, T_3) in §4.1. Note that the $\eta_3(.)$ values are precisely the number of $\mathcal{V}_3$ sets having $b_3 = 2$ for each design. Thus within $\mathcal{D}(8, 14, 4)$ the robustness of a design against incurring a RROS(3) through the loss of random observations corresponds to its robustness against lowest A-efficiency from the loss of two blocks.

See Thornewell (2011) for further investigation into coincidence between rankings of designs according to Intersection Aberration and (S_3, T_3).

5. Conclusion

For $D \in \mathcal{D}(v, b, k)$, the (S_u, T_u) measures developed in §3 give the smallest number of observations that comprise a specific kind of RROS and the number of such observation sets. Loss of observations in such a RROS yields an eventual design in which the treatments are partitioned into sets of size u and $v - u$ respectively, and the usual analysis to compare the treatments cannot be conducted.

For $u \in \{1, 2\}$ and $u < [k/2]$, both S_u and T_u are fixed for all designs in $\mathcal{D}(v, b, k)$. Also, for $\mathcal{D}(v, b, 2)$ and $\mathcal{D}(v, b, 3)$ design classes, all S_u and T_u are functions of the basic design parameters. Information obtained from these measures complements information on the Type III RROSs to give a full picture of the vulnerability of a design to become disconnected through observation loss. Prior to experimentation, calculation of fixed measures, and knowledge of the potential level of observation loss, provide the experimenter with a pilot procedure to check that the eventual design is likely to be connected.

Other measures are dependent on properties of the particular design. For k even and at least six, $S_{k/2}$ is fixed but $T_{k/2}$ is design dependent. For $k \geq 4$ and $u > k/2$, both S_u and T_u can vary within a $\mathcal{D}(v, b, k)$ indicating that consequences of observation loss may vary within the design class. For a $\mathcal{D}(v, b, k)$ with cardinality greater than one, comparison of measures for the lowest value of u for which (S_u, T_u) vary, enables the designs to be ranked according to vulnerability.

Investigation of designs in $\mathcal{D}(8, 14, 4)$ indicates that designs which are ranked high according to (S_u, T_u) also perform well with regards to Criterion-2 robustness in the event of different patterns of observation loss.

Acknowledgement

Helen Warren (nee Thornewell)'s PhD research at the University of Surrey was supported by EPSRC grant KB9021.

References

Bailey, R. A., Schiffl, K. and Hilgers, R-D. (2013). A note on robustness of D-optimal block designs for two-colour microarray experiments. *Journal of Statistical Planning and Inference*, **143** 1195–1202.

Baksalary, J. K. and Tabis, Z. (1987). Conditions for the robustness of block designs against the unavailability of data. *Journal of Statistical Planning and Inference*, **16** 49–54.

Bhar, L. (2014). Robustness of variance balanced designs. *Sankhya B*, **76**, 305–316.

Bhaumik, K. and Whittinghill, D. C. (1991). Optimality and robustness to the unavailability of blocks in block designs. *Journal of the Royal Statistical Society*, **B53**, 399–407.

Das, A. and Kageyama, S. (1992). Robustness of BIB and extended BIB designs against the nonavailability of any number of observations in a block. *Computational Statistics and Data Analysis*, **14**, 343–358.

Dey, A. (1993). Robustness of block designs against missing data. *Statistica Sinica*, **3**, 219–231.

Foody, W. and Hedayat, A. (1977). On theory and applications of BIB designs with repeated blocks. *Annals of Statistics*, **5** 932–945.

Ghosh, S. (1979). On robustness of designs against incomplete data. *Sankhya B*, **40**, 204–208.

Ghosh, S. (1982). Robustness of BIBD against the unavailability of data. *Journal of Statistical Planning and Inference*, **6**, 29–32.

Godolphin, J. D. (2016). A link between the E-value and the robustness of block designs. *Journal of the American Statistical Association*, **111**, 1736–1745.

Godolphin, J. D. (2019). Conditions for connectivity of incomplete block designs. *Quality and Reliability Engineering International*, **35**, 1279–1287.

Godolphin, J. D. and Warren, H. R. (2011). Improved conditions for the robustness of binary block designs against the loss of whole blocks. *Journal of Statistical Planning and Inference*, **141**, 3498–3505.

Godolphin, J. D. and Warren, H. R. (2014). An efficient procedure for the avoidance of disconnected incomplete block designs. *Computational Statistics and Data Analysis*, **71**, 1134–1146.

Kageyama, S. (1990). Robustness of block designs. *Probability Statistics and Design of Experiments* (R. R. Bahadur, ed.). Wiley Eastern, New Delhi, 425–438.

Lal, K., Gupta, V. K. and Bhar, L. (2001). Robustness of designed experiments against missing data. *Journal of Applied Statistics*, **28**, 63–79.

Mathon, R. and Rosa, A. (1996). 2-(v, k, λ) designs of small order. *CRC Handbook of Combinatorial Designs*, Ed. (C. J. Colbourn and J. H. Dinitz, ed.). CRC Press: Boca Raton, 3–41.

Morgan, J. P. and Parvu V. (2008). Most robust BIBDs. *Statistica Sinica*, **18**, 689–707.

Prescott, P. and Mansson, R. (2001). Robustness of balanced incomplete block designs to randomly missing observations. *Journal of Statistical Planning and Inference*, **92**, 283–296.

Raghavarao, D., Federer, W. T. and Schwager, S. J. (1986). Characteristics for distinguishing among balanced incomplete blocks with repeated blocks. *Journal of Statistical Planning and Inference*, **13**, 151–163.

Searle, M. (1971). *Linear Models.* Wiley, New York.

Shah, K. R. and Sinha, B. K. (1989). *Theory of Optimal Designs.* Springer, New York.

Thornewell, H. (2011). *Robustness Criteria for Vulnerability of Block Designs in the Event of Observation Loss.* Ph.D. thesis, University of Surrey, Mathematics.

Tsai, S-F and Liao, C-T. (2013). Minimum breakdown designs in blocks of size two. *Journal of Statistical Planning and Inference*, **143**, 202–208.

Whittinghill, D. C. (1995). A note on the optimality of block designs resulting from the unavailability of scattered observations. *Utilitas Mathematica*, **47**, 21–31.

Publisher
Society of Statistics, Computer and Applications
B - 133, Ground Floor, C.R. Park, New Delhi - 110019
Tele: 011 - 40517662
https://ssca.org.in/
statapp1999@gmail.com
2021

Printed by : Galaxy Studio & Graphics
Mob: +91 9818 35 2203, +91 9582 94 1203