Special Proceeding of the 22nd Annual Conference of SSCA held at Savitribai Phule Pune University, Pune, during January 02–04, 2020; pp 39-51

Linear Model Perspectives of fMRI Studies

Bikas Kumar Sinha
Former Professor, Indian Statistical Institute, Kolkata, India

Received: 29 April 2020; Revised: 09 May 2020; Accepted: 19 May 2020

Abstract

Functional Magnetic Resonance Imaging (fMRI) is a technology for studying how our brains respond to mental stimuli. It is interesting to note the potential developments of linear models in the study of 'design sequences' employed for fMRI studies. At the design stage, one is interested in developing a sequence of mental stimuli for collecting data in order to render information about some 'unknown' yet 'meaningful' parameters under an assumed statistical model. The simplest such model incorporates linear relation between 'mean response' and the 'parameters' describing the effects of the stimuli, applied at regularly spaced time points during the study period. In this paper, we introduce the linear model and discuss estimation issues. In the process, we take up a study of relative performances of comparable design sequences.

Key words: fMRI, Linear model; h-Parameters; Estimability; Information matrix; Generalized variance; Average variance; Design issues.

1. Introduction

It is interesting to note that **Statistics and Applications** published, in as early as 2008, an article dealing with **"event- related functional magnetic resonance imaging . . . "**. In fMRI studies, the brain functions of the experimental subjects are captured through response profiles at a number of instances. Each subject experiences onset of a stimulus at an instant if the stimulus is 'active' [denoted by code '1'] at that instant; otherwise, the subject is at 'resting state' [denoted by code '0'] at that instant. Each instant is defined as a compact duration of '4 seconds'. At any instant, the brain voxel captures the cumulative effects of a fixed (but unknown) parameter θ and other model parameters, known as h-parameters at the current instant as well as at each of the immediate past ordered $(K-1)$ instants - for some K - whenever there has been an onset of active stimulus at any of these instances. The reader familiar with the concept of 'carry-over effects' in the context of Repeated Measurement Designs [RMDs] or Cross-Over Designs will find a similarity in the model description. [Vide Shah and Sinha (1989)]. We will also mention about 'circular models' and for that we refer to Kunert (1984).

An anonymous referee has aptly pointed out another related piece of work by Maus *et al*. (2010).

This talk was delivered in C. R. Rao Birth Centenary Celebration session on January 02, 2020.
Correponding Author: Bikas K. Sinha
Email:bikassinha1946@gmail.com

A more general scenario exhibits itself in terms of different stages of activation of the brain stimuli, rather than just being ‘active’ - as coded by ‘1’ in the above. We refer to Kao *et al.* (2008) for this and related considerations.

Below we introduce the linear (mean) model as has been described in the literature. There is a ‘design sequence’ in the form of a collection of 1s and 0s of, say length n. We denote it by D_n. As for example, for $n = 8$, the following describes an 8-point design: $D_8 = [0, 1, 1, 0, 1, 1, 0, 1]$. The implementation of the suggested design D_8 is described below. For any n, D_n is very much like D_8. The linear model to be described below is developed as a ’circular’ model - a well-known consideration in the context of RMDs or Cross-Over Designs. Vide Kunert (1984) or Shah and Sinha (1989). To visualize a circular model, the same sequence (describing D_8) is used as a ’dummy’ sequence and this is described as follows:

$< 0, 1, 1, 0, 1, 1, 0, 1 >$	$\rightarrow$	$[0, 1, 1, 0, 1, 1, 0, 1]$
Dummy Sequence	followed by	Data-generating Sequence

There are 8 data/time points and as such we observe y_1 to y_8 corresponding to the 8 time points in the data-generating sequence $[0, 1, 1, 0, 1, 1, 0, 1]$ - going from left to right. In the terminology of RMDs or Cross-Over Designs, for the first time point, the ‘direct effect’ [denoted by h_1] is to be captured along with the ’carry-over effects’ $[h_2, h_3, ...]$ of the preceding time points as described in the Dummy Sequence - from right to left. Althrough, at each data point, only if the stimulus is active [denoted by 1], the corresponding h-parameter will be present in the mean model. Moreover, for n data/time points, we can incorporate at the most n ‘parameters’- including the fixed parameter θ. This implies that we can incorporate in the model at the most $(n - 1)$ h-parameters. Otherwise/estimability issues creep in. In terms of K, it means that we assume - to start with - that $K \leq (n - 1)$.

We start with the following Table 1 describing the linear (mean) model underlying the design D_8. We assume $K = 7$. For clarity, we explain the derivation of the mean model for y_1. The co-efficients to be attached to the regression parameters i.e., h-parameters $[h_1$ to $h_7]$ in the expression for the mean model corresponding to y_1 are: $(0, 1, 0, 1, 1, 0)$. This is seen as follows. In the data-generating sequence, extreme left-hand coefficient (0) is attached to h_1; then the coefficients in the dummy sequence are taken successively from right to left for attachment to h_2 to h_7. There are 6 h-parameters (in addition to h_1), and hence 6 of the coefficients are selected in the order from right to left in the dummy sequence. That gives the coefficients for h_2 to h_7 in the order $(1, 0, 1, 1, 0, 1)$. Hence the mean model for y_1 is given by $\theta + h_2 + h_4 + h_5 + h_7$. Likewise, for y_2, the coefficients start from the second member from the left of the data-generating sequence and proceeds along the left direction, cutting across the dummy-sequence and covers a total of 7 coefficients. The coefficients are thus $(1, 0, 1, 0, 1, 1, 0)$. All these are displayed in Table 1. Note that in Table 1, the h-parameters are listed in the reverse order.

Remark 1: It may be noted that the linear mean model developed above has similarity with one in the set-up of ’biased spring balance weighing designs’. Vide Raghavarao (1971) or Shah and Sinha (1989). It follows that θ-parameter represents the bias component in spring balance weighing design context. The co-efficient matrix $\mathbf{X} = ((x_{ij}))$ consists of 0s and 1s. However, the $\mathbf{X}$- matrix is shown in the reverse order. Multiplication by a permutation

Table 1: Linear Model with positional carry-over effects in terms of h-parameters

S1. No.	h_7	h_6	h_5	h_4	h_3	h_2	h_1	y	Mean Model
1	1	0	1	1	0	1	0	y_1	$\theta + h_2 + h_4 + h_5 + h_7$
2	0	1	1	0	1	0	1	y_2	$\theta + h_1 + h_3 + h_5 + h_6$
3	1	1	0	1	0	1	1	y_3	$\theta + h_1 + h_2 + h_4 + h_6 + h_7$
4	1	0	1	0	1	1	0	y_4	$\theta + h_2 + h_3 + h_5 + h_7$
5	0	1	0	1	1	0	1	y_5	$\theta + h_1 + h_3 + h_4 + h_6$
6	1	0	1	1	0	1	1	y_6	$\theta + h_1 + h_2 + h_4 + h_5 + h_7$
7	0	1	1	0	1	1	0	y_7	$\theta + h_2 + h_3 + h_5 + h_6$
8	1	1	0	1	1	0	1	y_8	$\theta + h_1 + h_3 + h_4 + h_6 + h_7$

matrix $\mathbf{P}$ will bring it to the right/standard order. Finally, the linear model $(\mathbf{Y}, \mathbf{X}^{(*)}\beta, \sigma^{\mathbf{2}}\mathbf{I})$ is obtained as usual where $\mathbf{X}^{(*)} = (\mathbf{1}, \mathbf{PX})$ and β=(θ, h_1, h_2, ...)'. It is assumed that the errors are, as usual, uncorrelated with zero means and equal variances.

Remark 2: We must note that a 'circular model' has been explicitly used in Table 1. The dummy - sequence is derived from the data - generating sequence on which the circular model is built. Another implication is that the columns $\mathbf{h}_1, \mathbf{h}_2$, ... are circular in nature. That is, the columns of the matrix $\mathbf{X}$ are circular in nature. For a non- circular design/model, the carry - over effects are dependent on the nature of 1s and 0s - for each incoming unit/patient-at the two ends of the design sequence.

At times, the number of h-parameters may be specified and it may happen that there are $K^*[< K]$ h-parameters in the model. In that case, the understanding is that the initial set of K^* h-parameters viz. $h_1, h_2, \ldots, h_{K^*}$ are important and the rest can be ignored from the mean model. For $K^* = 4$, the model expectations of successive responses corresponding to the above design would be:

$$\theta + h_2 + h_4, \theta + h_1 + h_3, \theta + h_1 + h_2 + h_4, \theta + h_2 + h_3, \theta + h_1 + h_3 + h_4,$$

$$\theta + h_1 + h_2 + h_4, \theta + h_2 + h_3, \theta + h_1 + h_3 + h_4.$$

Note that the above design with $n = 8$ instances [for experimentation] generates more number of observations when only $K^* = 4$ h-parameters are assumed to be present. In such a situation, we might curtail the experiment from D_8 to D_5 since there are 5 parameters, including the common/fixed parameter θ. Use of $D_5 : [0, 1, 1, 0, 1]$ provides for the mean model the expressions:

$$\theta + h_2 + h_4, \theta + h_1 + h_3, \theta + h_1 + h_2 + h_4, \theta + h_2 + h_3, \theta + h_1 + h_3 + h_4.$$

On the other hand, use of $D_{alt.5} : [1, 1, 0, 1, 1]$ provides for the mean model the expressions:

$$\theta + h_1 + h_2 + h_3, \theta + h_1 + h_2 + h_3 + h_4, \theta + h_2 + h_3 + h_4, \theta + h_1 + h_3 + h_4, \theta + h_1 + h_2 + h_4.$$

Note that in both the cases, we have taken due consideration of circular nature of the sequence in working out the mean models. A natural question would be to search out the difference, if any, between the two D_5 designs. Popular optimality criteria rest on the computation of the 'information matrix' for the h-parameters - based on the Gauss-Markov Model, assuming homoscedastic errors with mean 0 and variance σ^2. Minimization of the generalized variance [computed as reciprocal of the determinant of the information matrix] is an acceptable criterion for choice of the best design. This is the so-called D-optimality Criterion [Vide Shah and Sinha (1989)]. We will take up this comparative study in the next section.

2. Linear Estimation of Model Parameters

Since the linear model involves a fixed parameter (θ), for a given number of observations n, we can incorporate a maximal set of $(n-1)$ h-parameters. That is, we can develop the full model with θ and additional $(n-1)$ h-parameters. Naturally, the response vector $\mathbf{Y}$ of dimension $n \times 1$ will come under the standard Gauss-Markov Linear Model mentioned earlier. However, estimability of the h-parameters or of θ are not necessarily guaranteed for all choices of the design sequence.

We have already introduced the 'design matrix' $\mathbf{X}^{(*)} = (\mathbf{1}, \mathbf{PX})$ and the underlying parameters β=(θ, h_1, h_2, ...)'. For a given design D_n, when there are $K[\leq (n-1)]$ h-parameters viz., $h_1, h_2, \ldots, h_K$ in the model, the h-parameters are all estimable iff Rank $(\mathbf{X}^{(*)}) = 1+K$ where $\mathbf{X}^{(*)}$ is based on K column vectors corresponding to the K h-parameters, in addition to the column vector $\mathbf{1}$. The 'if' part is easy to see. On the other hand, if all the h-parameters are estimable, θ is trivially so based on any single observation and hence the rank condition is satisfied.

In the above example, for $K = 7$, it can be seen that the design D_8 ensures estimability of all the model parameters. Explicit expressions for the estimates of h-parameters are shown below. For θ, expression for its estimator follows readily.

$$h_1 : y_6 - y_1;\ h_2 : -y_1 - y_2 + y_6 + y_7;\ h_3 : -y_1 - y_2 - y_3 + y_6 + y_7 + y_8;$$

$$h_4 : -y_2-y_3-y_4+y_6+y_7+y_8;\ h_5 : -y_3-y_4-y_5+y_6+y_7+y_8;\ h_6 : -y_4-y_5+y_7+y_8; h_7 : -y_5+y_8.$$

This suggests that $\mathbf{X}^{(*)}$ is a full rank square matrix of order 8. Hence, all its column vectors are linearly independent. Therefore, for all values of K^*, the number of non-negligible h-parameters, the above design sequence D_8 provides unbiased estimates for each one of them. This holds for all $1 \leq K^* \leq K = 7$.

At this stage, we may as well resolve two more cases. For $K^* = 4$, we may check the acceptability of the two D_5 design sequences listed above: $D_5 : [0, 1, 1, 0, 1]$ and $D_{alt.5}$: $[1, 1, 0, 1, 1]$. It turns out that both are acceptable from estimability point of view. It would be interesting to make a comparison of their performances with respect to, say, D-optimality criterion. Necessary computations are shown below.

$$I(\beta) = [(5, 3, 3, 3, 3), (3, 3, 1, 2, 2), (3, 1, 3, 1, 2), (3, 2, 1, 3, 1), (3, 2, 2, 1, 3)].$$

$$I(h) = [(6, -4, 1, 1), (-4, 6, -4, 1), (1, -4, 6, -4), (1, 1, -4, 6)], \quad Det(I) = 125.$$

$$I_{alt}(\beta) = [(5, 4, 4, 4, 4), (4, 4, 3, 3, 3), (4, 3, 4, 3, 3), (4, 3, 3, 4, 3), (4, 3, 3, 3, 4)].$$

$$I_{alt}(h) = [(4, -1, -1, -1), (-1, 4, -1, -1), (-1, -1, 4, -1), (-1, -1, -1, 4)], \quad Det(I_{alt}) = 125.$$

It thus turns out that the two design sequences provide identical generalized variance of the estimates of h-parameters. We will return to this comparison later again in Remark 3.

3. Choice of D_n for given n and K^*

In the context of fMRI study, assume that it is a priori known that, for some K^*, h_1,h_2, ..., h_{K^*} are the only h-parameters present in the mean model. Therefore, we need $n \geq (1 + K^*)$ design points and the choice of D_n must be such that the formation of $\mathbf{X}$ enables one to ensure rank condition. For a chosen n, it is obvious that there are a large number of design sequences of length n-comprising of 1s and 0s. This count is 2^n. It is easy to note that the two extreme sequences $(1, 1, \ldots, 1)$ and $(0, 0, \ldots, 0)$ are inadmissible. In other words, no patient can be in resting phase or in active phase althrough the time duration of the experiment for collection of data. Generally, a mixture of the two phases is called for.

Below we examine the status of a special "Design Sequence [DS]" of length n. Consider the design sequence D_n: [1, 1, 0,. . . , 0, 0] which gives rise to [1, 1, 0, . . . , 0, 0] dummy sequence followed by [1, 1, 0, . . . , 0, 0] data-gathering sequence.

Therefore, model expectations of the resulting responses ys are given by: $[\theta + h_1, \theta + h_1 + h_2, \theta + h_2 + h_3, \ldots, \theta + h_{(n-2)} + h_{(n-1)}, \theta + h_{(n-1)}]$, assuming that there are $(n - 1)$ h-parameters in the model. It is interesting to note the following:

(i) For $n = 4,\ K = 3$, the joint information matrix is singular. (ii) For $n = 5,\ K = 4$, the joint information matrix is non - singular. (iii) For $n = 6,\ K = 5$, the joint information matrix is singular. (iv) For $n = 7,\ K = 6$, the joint information matrix is again non - singular

It turns out that for all n $(even) \geq 4,\ K = (n - 1)$, the joint information matrix is singular while for all n $(odd)\ \geq 5,\ K = (n - 1)$, the joint information matrix is non-singular. Let us fix $n = 8,\ K = 7$ so that DS_8 = [1,1, 0, 0, 0, 0, 0, 0] is not admissible. What if we replace the extreme right- end code 0 by 1? We are asking about the status of $DS^*_8 = [1, 1, 0, 0, 0, 0, 0, 1]$. It follows that the 8×8 joint information matrix is given by

$$[(8, 3, 3, 3, 3, 3, 3, 3);\ (3, 3, 2, 1, 0, 0, 0, 1);\ (3, 1, 3, 2, 1, 0, 0, 0);\ \ldots,\ (3, 1, 0, 0, 0, 1, 2, 3)].$$

and it is of full rank.

Therefore, it pays off to change exactly one code in the above.

For n odd, each member of the above series of design sequences provides estimates of all the relevant h-parameters. For $n = 7, K = 6$, it follows that

$$V(\hat{h}_2) = 2\sigma^2,\ V(\hat{h}_4) = 4\sigma^2,\ V(\hat{h}_6) = 6\sigma^2,$$

while

$$V(\hat{h}_1) = 6\sigma^2,\ V(\hat{h}_3) = 4\sigma^2,\ V(\hat{h}_5) = 2\sigma^2.$$

Again, for $n = 9,\ K = 8$, we obtain

$$V(\hat{h}_2) = 2\sigma^2,\ V(\hat{h}_4) = 4\sigma^2,\ V(\hat{h}_6) = 6\sigma^2,\ V(\hat{h}_8) = 8\sigma^2,$$

while

$$V(\hat{h}_1) = 8\sigma^2,\ V(\hat{h}_3) = 6\sigma^2,\ V(\hat{h}_5) = 4\sigma^2,\ V(\hat{h}_7) = 2\sigma^2.$$

These expressions suggest general form of the variances of estimates of the h-parameters. For specified (n, K), we can also work out the variance-covariance matrix of the estimates of the h-parameters. For the choice $n = 7, K = 6$, we derive the form of the variance-covariance matrix as given below.

$$[(6, 1, 4, 3, 2, 5), (-, 2, 0, 2, 0, 2), (-, -, 4, 1, 2, 3),$$

$$(-, -, -, 4, 0, 4), (-, -, -, -, 2, 1), (-, -, -, -, -, 6)].$$

4. Comparison of Design Sequences

When we address this problem for design sequences of the same length n, there are effectively $2^n - 2$ such comparable sequences - barring the two extremes [all 0s and all 1's]. Actual number of admissible sequences may be much smaller - depending on the number K of non-negligible h-parameters. Anyway, such a comparison of two admissible sequences may rest on, say the criterion of smaller average variance or smaller generalized variance of the estimated h-parameters. Below we take up the case of a saturated model with $n = 7, K = 6$ and compare all available admissible design sequences. Note that we have already studied one such admissible design sequence in the above. In this case there are $2^7 - 2 = 126$ possible design sequences of length 7 each-barring the two inadmissible extreme allocations (viz., all 1's and all 0's). These design sequences can be classified into distinct types as follows.

$TypeI : (i)\ \ [1, 0, 0, 0, 0, 0, 0];\ (ii)\ \ [1, 1, 0, 0, 0, 0, 0];\ (iii)\ \ [1, 1, 1, 0, 0, 0, 0];\ (iv)\ \ [1, 1, 1, 1, 0, 0, 0];$

$$TypeI\ continued : (v)\ \ [1, 1, 1, 1, 1, 0, 0];\ (vi)\ \ 1, 1, 1, 1, 1, 1, 0]$$

and all their cyclic permutations-covering 42 design sequences;

$$TypeII : (i)\ \ [1, 0, 1, 0, 0, 0, 0];\ (ii)\ \ [1, 0, 0, 1, 0, 0, 0]$$

and all their cyclic permutations involving 2 non-consecutive 1's-covering 14 design sequences;

$TypeIII : (i)\ \ [1, 1, 0, 1, 0, 0, 0] - replicated\,twice;\ (ii)\ \ [1, 1, 0, 0, 1, 0, 0];\ (iii)\ \ [1, 0, 1, 0, 1, 0, 0]$

and all their cyclic permutations involving 3 non-consecutive 1's-covering 28 design sequences;

$TypeIV : (i)\ \ [1, 1, 1, 0, 1, 0, 0] -\ replicated\,twice; (ii)\ \ [1, 1, 0, 1, 1, 0, 0];\ (iii)\ \ [1, 1, 0, 1, 0, 1, 0]$

and all their cyclic permutations involving 4 non-consecutive 1's-covering 28 design sequences;

$$TypeV : (i) \quad [1, 1, 1, 1, 0, 1, 0]; \;\; (ii) \quad [1, 1, 1, 0, 1, 1, 0]$$

and all their cyclic permutations involving 5 non-consecutive 1's-covering 14 design sequences.

Routine computations can be done to ascertain respective status of each of the design sequences listed above for any specified value of K-the number of non-negligible h-parameters. Below we show the detailed analysis of the design sequences of Type I.

Table 2: Type I(i): Coefficients of estimates of h-parameters and their variances

parameter / coefficient	y_1	y_2	y_3	y_4	y_5	y_6	y_7	Variance coefficient
θ	0	0	0	0	0	0	1	1
h_1	1	0	0	0	0	0	-1	2
h_2	0	1	0	0	0	0	-1	2
h_3	0	0	1	0	0	0	-1	2
h_4	0	0	0	1	0	0	-1	2
h_5	0	0	0	0	1	0	-1	2
h_6	0	0	0	0	0	1	-1	2

Table 3: Type I(ii) : Coefficients of estimates of h-parameters and their variances

parameter / coefficient	y_1	y_2	y_3	y_4	y_5	y_6	y_7	Variance coefficient
θ	1	-1	1	-1	1	-1	1	7
h_1	0	1	-1	1	-1	1	-1	6
h_2	-1	1	0	0	0	0	0	2
h_3	0	0	0	1	-1	1	-1	4
h_4	-1	1	-1	1	0	0	0	4
h_5	0	0	0	0	0	1	-1	2
h_6	-1	1	-1	1	-1	1	0	6

Table 4: Type I(iii): Coefficient of estimates of h-parameters and their variances

parameter / coefficient	y_1	y_2	y_3	y_4	y_5	y_6	y_7	Variance coefficient
θ	1	1	-2	1	1	-2	1	13
h_1	0	0	1	-1	0	1	-1	4
h_2	-1	0	1	0	-1	1	0	4
h_3	0	-1	1	0	0	0	0	2
h_4	0	0	0	0	0	1	-1	2
h_5	-1	0	1	-1	0	1	0	4
h_6	0	-1	1	0	-1	1	0	4

Table 5: Type I(iv): Coefficient of estimates of h-parameters and their variances

parameter / coefficient	y_1	y_2	y_3	y_4	y_5	y_6	y_7	Variance coefficient
θ	-1	-1	3	-1	-1	-1	3	23
h_1	1	0	-1	1	0	0	-1	4
h_2	0	1	-1	0	1	0	-1	4
h_3	0	0	0	0	0	1	-1	2
h_4	0	0	-1	1	0	0	0	2
h_5	1	0	-1	0	1	0	-1	4
h_6	0	1	-1	0	0	1	-1	4

Table 6: Type I(v): Coefficient of estimates of h-parameters and their variances

parameter / coefficient	y_1	y_2	y_3	y_4	y_5	y_6	y_7	Variance coefficient
θ	-2	3	-2	3	-2	-2	3	43
h_1	0	0	1	-1	1	-1	0	4
h_2	0	0	0	0	0	1	-1	2
h_3	0	-1	1	-1	1	0	0	4
h_4	1	-1	0	0	0	1	-1	4
h_5	0	0	0	-1	1	0	0	2
h_6	1	-1	1	-1	0	1	-1	6

Table 7: Type I(vi): Coefficient of estimates of h-parameters and their variances

parameter / coefficient	y_1	y_2	y_3	y_4	y_5	y_6	y_7	Variance coefficient
θ	1	1	1	1	1	-5	1	31
h_1	0	0	0	0	0	1	-1	2
h_2	-1	0	0	0	0	1	0	2
h_3	0	-1	0	0	0	1	0	2
h_4	0	0	-1	0	0	1	0	2
h_5	0	0	0	-1	0	1	0	2
h_6	0	0	0	0	-1	1	0	2

To summarize the performances of the above design sequences of Type I, we find that in terms of average variance of the estimates of the h-parameters,

$$(i) = (vi) < (iii) = (iv) < (v) < (ii).$$

Table 8: Type II(i): Coefficients of estimates of h-parameters and their variances

parameter / coefficient	y_1	y_2	y_3	y_4	y_5	y_6	y_7	Variance coefficient
θ	-1	1	1	-1	-1	1	1	7
h_1	1	-1	0	1	0	-1	0	4
h_2	1	0	-1	1	1	-1	-1	6
h_3	0	0	0	0	1	0	-1	2
h_4	0	-1	0	1	0	0	0	2
h_5	1	-1	-1	1	1	-1	0	6
h_6	1	0	-1	0	1	0	-1	4

Table 9: Type II(ii): Coefficients of estimates of h-parameters and their variances

parameter / coefficient	y_1	y_2	y_3	y_4	y_5	y_6	y_7	Variance coefficient
θ	1	1	1	-1	-1	-1	1	7
h_1	0	0	0	1	0	0	-1	2
h_2	-1	0	0	1	1	0	-1	4
h_3	-1	-1	0	1	1	1	-1	6
h_4	-1	-1	-1	1	1	1	0	6
h_5	0	-1	-1	0	1	1	0	4
h_6	0	0	-1	0	0	1	0	2

Table 10: Type III(i): Coefficients of estimates of h-parameters and their variances

parameter / coefficient	y_1	y_2	y_3	y_4	y_5	y_6	y_7	Variance coefficient
2θ	2	-1	2	-1	-1	-1	2	4
$2h_1$	0	1	-1	1	0	0	-1	1
$2h_2$	-1	1	0	0	1	0	-1	1
$2h_3$	-1	0	0	1	0	1	-1	1
$2h_4$	-1	0	-1	1	1	0	0	1
$2h_5$	0	0	-1	0	1	1	-1	1
$2h_6$	-1	1	-1	0	0	1	0	1

We have completed computations of $\hat{h}$s along with their variances for all the effectively sixteen (16) competing design sequences. We may now display the totals of variances across all competitors.

$TypeI(i)$ 12; $TypeI(ii)$ 24; $TypeI(iii)$ 20; $TypeI(iv)$ 20; $TypeI(v)$ 22; $TypeI(vi)$ 12

$TypeII(i)$ 24; $TypeII(ii)$ 24

$TypeIII(i)$ 6; $TypeIII(ii)$ 20; $TypeIII(iii)$ 20

$TypeIV(i)$ 6; $TypeIV(ii)$ 20; $TypeIV(iii)$ 20

$TypeV(i)$ 24; $TypeV(ii)$ 24

Table 11: Type III(ii): Coefficients of estimates of h-parameters and their variances

parameter / coefficient	y_1	y_2	y_3	y_4	y_5	y_6	y_7	Variance coefficient
θ	1	1	-2	1	-2	1	1	13
h_1	0	0	1	-1	1	-1	0	4
h_2	0	0	1	0	0	0	-1	2
h_3	-1	0	1	0	1	-1	0	4
h_4	0	-1	1	0	1	0	-1	4
h_5	-1	0	0	0	1	0	0	2
h_6	0	-1	1	-1	1	0	0	4

Table 12: Type III(iii): Coefficients of estimates of h-parameters and their variances

parameter / coefficient	y_1	y_2	y_3	y_4	y_5	y_6	y_7	Variance coefficient
θ	1	1	1	1	-1	-1	1	13
h_1	0	0	0	0	1	0	-1	2
h_2	-1	0	0	0	1	1	-1	4
h_3	-1	0	1	0	1	-1	0	4
h_4	0	-1	-1	0	1	1	0	4
h_5	0	0	-1	-1	1	1	0	4
h_6	0	0	0	-1	0	1	0	2

Table 13: Type IV(i): Coefficients of estimates of h-parameters and their variances

parameter / coefficient	y_1	y_2	y_3	y_4	y_5	y_6	y_7	Variance coefficient
θ	1	1	-1	1	-1	-1	1	7
$2h_1$	0	0	1	-1	1	0	-1	1
$2h_2$	-1	0	1	0	0	1	-1	1
$2h_3$	-1	-1	1	0	1	0	0	1
$2h_4$	0	-1	0	0	1	1	-1	1
$2h_5$	-1	0	0	-1	1	1	0	1
$2h_6$	0	-1	1	-1	0	1	0	1

In conclusion, we find that the design sequences $TypeIII\ (i) : [1, 1, 0, 1, 0, 0, 0]$ and $TypeIV\ (i) : [1, 1, 1, 0, 1, 0, 0]$ are, together with their cyclic permutations, most efficient with respect to the average variance criterion. It is again readily observed that for both these designs, pair- wise covariance terms of the estimates of the h-parameters are all equal and it is the same for both. Therefore, as such, the two competing sequences are information-equivalent !

Our task will not be complete unless we discuss one more pertinent observation in this context. The above comparison may not be 'fair' since the design sequences are based on

Table 14: Type IV(ii): Coefficients of estimates of h-parameters and their variances

parameter / coefficient	y_1	y_2	y_3	y_4	y_5	y_6	y_7	Variance coefficient
θ	3	-1	3	-1	-1	-1	-1	23
h_1	-1	1	-1	1	0	0	0	4
h_2	-1	0	0	0	1	0	0	2
h_3	-1	0	-1	1	0	1	0	4
h_4	-1	0	-1	0	1	1	0	4
h_5	0	0	-1	0	0	1	0	2
h_6	-1	1	-1	0	0	0	1	4

Table 15: Type IV(iii): Coefficients of estimates of h-parameters and their variances

parameter / coefficient	y_1	y_2	y_3	y_4	y_5	y_6	y_7	Variance coefficient
θ	3	-1	-1	-1	-1	-1	3	23
h_1	0	1	0	0	0	0	-1	2
h_2	-1	1	1	0	0	0	-1	4
h_3	-1	0	1	1	0	0	-1	4
h_4	-1	0	0	1	1	0	-1	4
h_5	-1	0	0	0	1	1	-1	4
h_6	-1	0	0	0	0	1	0	2

Table 16: Type V(i): Coefficients of estimates of h-parameters and their variances

parameter / coefficient	y_1	y_2	y_3	y_4	y_5	y_6	y_7	Variance coefficient
θ	3	3	-2	-2	3	-2	-2	43
h_1	-1	0	1	0	-1	1	0	4
h_2	-1	-1	1	1	-1	0	1	6
h_3	0	-1	0	1	0	0	0	2
h_4	-1	0	0	0	0	1	0	2
h_5	-1	-1	1	0	-1	1	1	6
h_6	0	-1	0	1	-1	0	1	4

unequal number of 1s. Note that every sequence comprises of 1s and 0s and the understanding is that a 0-phase corresponds to 'idle' phase while a 1-phase is 'active'. So the number of active phases should also be considered while examining relative performances. We may apply the usual concept of "Efficiency" and work out "Efficiency per active phase". For a single parameter, efficiency is directly related to and measured by [Fisher] Information. For $K = 6$ h-parameters, we can compute the average variance of the estimates and multiply it by the number of 1s and minimize this quantity. If we are guided by this consideration, we find that the design sequence $TypeIII$ $(i) : [1, 1, 0, 1, 0, 0, 0]$ is the best of all! We can argue that this is also the best with respect to generalized variance criterion as well.

Table 17: Type V(ii): Coefficients of estimates of h-parameters and their variances

parameter / coefficient	y_1	y_2	y_3	y_4	y_5	y_6	y_7	Variance coefficient
θ	3	3	-2	-2	-2	-2	3	43
h_1	0	0	1	0	0	0	-1	2
h_2	-1	0	1	1	0	0	-1	4
h_3	-1	-1	1	1	1	0	-1	6
h_4	-1	-1	0	1	1	1	-1	6
h_5	-1	-1	0	0	-1	-1	0	4
h_6	0	-1	0	0	0	1	0	2

Remark 3: At the end of Section 2, we had introduced two design sequences D_5 and $D_{alt.\ 5}$ for the case of $n = 5,\ K^* = 4$. We also observed that the two sequences possess the same generalized variance. However, it can be seen that the alternative sequence provides smaller average variance. Now we note that whereas in D_5 the number of active phases used was 3, in the alternative design this number was 4. As in the above, we borrow the concept of "Efficiency per observation" while this time we define the "Efficiency" as reciprocal of the generalized variance, raised to the power 1/4 since there are 4 h-parameters. Otherwise, we can also use the reciprocal of the average variance. Adjusting for the difference in the number of active phases, we conclude that (i) D_5 is better than $D_{alt.\ 5}$ under the generalized variance criterion, while (ii) alternative sequence is better under average variance criterion.

Acknowledgments

The author is thankful to Professor Rajender Parsad of IASRI, New Delhi for bringing to his notice the publication by Kao *et al*. (2008) cited in the list of references. He also expresses his thanks to Professors Nripes K Mandal and Manisha Pal, Department of Statistics, Calcutta University, for taking keen interest in this topic of research and for fruitful discussions from time to time. The author also acknowledges citation of a related study in Maus *et al*. (2010) by an anonymous referee. One research collaborator, Dr Sobita Sapam, has kindly helped the author in formatting of the latex version following the template of instructions.

References

Cheng, C. S. and Kao, M. H. (2015). Optimal experimental designs for fMRI via circulant biased weighing designs. *Annals of Statistics*, **43(6)**, 2565–2587.

Kunert, J. (1984). Optimality of balanced uniform repeated measurements designs. *Annals of Statistics,* **12(4)**, 1006–1017.

Maus, B., Van Breukelen, G. J. P., Goebel, R. and Berger, M. P. F.(2010). Optimization of blocked designs in fMRI studies. *Psychometrika*, **75**, 373–390.

Kao, M.-H., Mandal, A. and Stufken, J. (2008: New series). Optimal design for event-related functional magnetic resonance imaging considering both individual stimulus effects and pairwise contrasts. *Statistics and Applications*, **6(1,2)**, 235–256.

Raghavarao, D. (1971). *Construction and Combinatorial Problems in Design of Experiments.* John Wiley and Sons Inc. New York.

Shah, K. R. and Sinha, Bikas K. (1989). *Theory of Optimal Designs.* Springer Lecture Notes in Statistics Series No. 54.