Statistics and Applications {ISSN 2452-7395(online)}
Volume 18, No. 2, 2020 (New Series), pp 97–119

The Simultaneous Assessment of Normality and Homoscedasticity in One-Way Random Effects Models

Ye Yang and Thomas Mathew
Department of Mathematics and Statistics
University of Maryland Baltimore County, Maryland, USA

Received: 23 May 2020; Revised: 26 June 2020; Accepted: 28 June 2020

Abstract

The article investigates the simultaneous assessment of normality and homoscedasticity in a one-way random effects model. Test procedures are developed assuming a smooth alternative to the normal distribution, specified using Legendre polynomials and Hermite polynomials. Score statistics are derived under both classes of alternatives, and a data driven approach is used to determine the order of the polynomials. Numerical results are reported in order to assess the accuracy of the chisquare distribution as the null distribution of the score statistics. Estimated powers are reported in order to compare the score tests derived under the alternatives based on Legendre polynomials and Hermite polynomials. An example and the corresponding data analysis are reported in order to illustrate the results. Possible extensions to other models involving random effects are briefly indicated.

Key words: Hermite polynomials; Legendre polynomials; Score test; Smooth alternative.

AMS Subject Classifications: 62F03, 62J10

1. Introduction

Mixed and random effects models are among the most widely used tools in applied work. While analyzing data using such models, the standard assumptions include normality of the random effects and the experimental error terms, as well as homoscedasticity of the error distribution. The present work is focused on testing these assumptions under the simplest random effects model, namely, the one-way random effects model. The usual practice is to test these assumptions separately. For example, if normality can be assumed, homoscedasticity can be assessed using a formal test such as Bartlett's test. One can also use a test that is less sensitive to the normality assumption, for example the Levene test and the modified Levene test; see the article by Chang, Pal, and Lin (2017), and Section 3.4 in the book by Montgomery (2020). On the other hand, normality is often assessed using formal tests or using a graphical method such as the normal probability plot, after assuming homoscedasticity. It is certainly desirable to have test procedures that will permit us to simultaneously assess homoscedasticity and normality. The present work aims to develop such test procedures.

Corresponding Author: Thomas Mathew
Email: mathew@umbc.edu

In order to formally introduce the relevant hypothesis, consider data falling into a groups (for example, corresponding to a treatments in a designed experiment). Suppose we have n_l observations available from the lth group, with y_{lj} denoting the jth observation; $j = 1, 2, \cdots, n_l$, $l = 1, 2, \cdots, a$. We are allowing the n_l's to be unequal, so that we can have unbalanced data. The one-way random model for the y_{lj}, along with the normality assumptions, is given by

$$y_{lj} = \mu + \alpha_l + \epsilon_{lj}, \quad \alpha_l \sim N(0, \sigma_\alpha^2),\ \epsilon_{lj} \sim N(0, \sigma_l^2), \tag{1}$$

where all the random variables are assumed to be independent. Note that the error terms (*i.e.*, the ϵ_{lj}'s) have variance σ_l^2, which could differ across the a groups. The α_l's in (1) denote the random effects, $l = 1, 2, \cdots, a$. If we write $\boldsymbol{y}_l = (y_{l1}, y_{l2}, \cdots, y_{ln_l})'$, then the above assumptions imply

$$\boldsymbol{y}_l \sim N\left(\mu \mathbf{1}_{n_l}, \sigma_l^2 I_{n_l} + \sigma_\alpha^2 \mathbf{1}_{n_l}\mathbf{1}_{n_l}'\right), \tag{2}$$

where $\mathbf{1}_r$ is an $r \times 1$ vector of ones. We note that the $\boldsymbol{y}_l$'s are independent for $l = 1, 2, \cdots, a$. Clearly, data analysis based on the one-way random model under the standard assumptions of normality and homoscedasticity amounts to analyzing the data under the multivariate normal model (2), having the structured covariance matrix as specified, and having σ_l^2s all equal. Thus testing homoscedasticity and normality under the one-way random effects model is equivalent to testing the equality of the σ_l^2 along with the multivariate normality of the $\boldsymbol{y}_l$, $l = 1, 2, \cdots, a$, where the covariance matrix has the structure specified in (2). Consequently, our normality assessment is for the multivariate normality of the $\boldsymbol{y}_l$, $l = 1, 2, \cdots, a$, and not for the univariate normality of the random effects and the error terms, even though the latter implies the multivariate normal distribution in (2).

Our development relies on the specification of alternatives to normality to be the class of *smooth alternatives* proposed by Neyman (1937). In general, suppose the problem is to test if a continuous random variable Y follows a specified distribution having density, say $f(y, \boldsymbol{\beta})$, depending on an unknown parameter vector $\boldsymbol{\beta}$. The alternative hypothesis is specified in terms of a smooth alternative involving orthonormal polynomials, say $\{p_i(y; \boldsymbol{\beta})\}$, $i = 1, 2, \cdots$, that are orthonormal with respect to $f(y; \boldsymbol{\beta})$. An order k smooth alternative, say $g_k(y; \boldsymbol{\theta}, \boldsymbol{\beta})$, is given by

$$g_k(y; \boldsymbol{\theta}, \boldsymbol{\beta}) = C(\boldsymbol{\theta}, \boldsymbol{\beta}) \exp\left\{\sum_{i=1}^{k} \theta_i p_i(y; \boldsymbol{\beta})\right\} f(y; \boldsymbol{\beta}). \tag{3}$$

In (3), $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_k)'$ is a vector of unknown parameters, and $C(\boldsymbol{\theta}, \boldsymbol{\beta})$ is a normalizing constant. As already noted, the $\{p_i(y; \boldsymbol{\beta})\}$ are orthonormal polynomials, orthonormal with respect to the null density $f(y; \boldsymbol{\beta})$. It should be clear that if $\boldsymbol{\theta}$ is the null vector, then $g_k(y; \boldsymbol{\theta}, \boldsymbol{\beta})$ in (3) reduces to the null density $f(y; \boldsymbol{\beta})$. In other words, the null density is embedded in the class of alternatives specified in (3), and testing for the goodness-of-fit of the null density is equivalent to testing if the $k-$dimensional vector $\boldsymbol{\theta}$ is the null vector. That is, the goodness-of-fit problem is now reduced to that of testing a hypothesis concerning a finite number of parameters.

Several authors have derived score tests for testing goodness-of-fit under the Neyman (1937) framework, and have investigated the theoretical properties of such tests for a variety of goodness-of-fit problems: see Ledwina (1994), Kallenberg and Ledwina (1995, 1997a, 1997b), Inglot and Ledwina (1996), Inglot, Kallenberg and Ledwina (1997), Kallenberg, Ledwina and Rafajlowicz (1997) and Janic and Ledwina (2009). Most of these articles address testing only the goodness of fit of a particular distribution; however, Kallenberg, Ledwina and Rafajlowicz (1997) address the simultaneous testing of normality and independence in a bivariate scenario. The simultaneous assessment of the various assumptions in a standard linear model set up is taken up in Peña and Slate (2006), and the authors consider the Neyman (1937) framework in order to specify alternatives to normality. More recently, Yang and Mathew (2018) have addressed the simultaneous testing of normality and homoscedasticity in a fixed effects model when we have grouped data, similar to those in an ANOVA context with fixed effects. A book-length discussion of smooth tests is available in Rayner, Thas and Best (2009).

In the next section, we shall derive score tests for testing normality and homoscedasticity in the set up of the model (2), assuming smooth alternatives of the type (3). We shall consider two specifications for the smooth alternative based on two choices for the orthonormal polynomials $\{p_i(y;\boldsymbol{\beta})\}$ in (3), namely, Legendre polynomials and Hermite polynomials. Thus we have two score tests corresponding to the two specifications for the alternatives. While specifying the alternatives, there is obvious arbitrariness in the choice of the order of the Legendre polynomials and Hermite polynomials, i.e., the quantity k in (3). We shall follow a *data driven* approach for choosing the order; an idea developed in Inglot, Kallenberg and Ledwina (1994), and pursued in some of the later papers by the authors, cited earlier. Our tests being based on score statistics, we can think of approximating the null distribution with a chisquare distribution. Thus we shall report numerical results in order to assess the accuracy of the chisquare distribution as the null distribution. The tests will be compared using estimated powers. Data analysis based on an example will be reported in order to illustrate the results.

2. Smooth Alternatives and Score Tests

Before we formally specify the alternative hypothesis, we shall consider an orthogonal transformation of each of the data vectors $\boldsymbol{y}_l$ in (2), $l = 1, 2, \cdots, a$. Let $Q_l = \left(\frac{1}{\sqrt{n_l}}1_{n_l}, Q_l^*\right)$ be an $n_l \times n_l$ Helmert matrix, and consider the transformation

$$\boldsymbol{v}_l = Q_l'\boldsymbol{y}_l,\ l = 1, 2, \cdots, a, \tag{4}$$

$$\text{so that } E(\boldsymbol{v}_l) = (\mu\sqrt{n_l}, 0, \cdots, 0)', \quad V(\boldsymbol{v}_l) = \sigma_l^2 I_{n_l} + \sigma_\alpha^2 \text{diag}(n_l, 0, 0, \cdots, 0).$$

Clearly, testing multivariate normality of the $\boldsymbol{y}_l$'s is equivalent to testing the same for the $\boldsymbol{v}_l$'s. Writing $\boldsymbol{v}_l = (v_{l1}, v_{l2}, \cdots, v_{ln_l})'$, we note that multivariate normality for the $\boldsymbol{y}_l$'s implies

$$v_{l1} \sim N(\mu\sqrt{n_l}, \sigma_l^2 + n_l\sigma_\alpha^2),\ v_{lj} \sim N(0, \sigma_l^2),\ j = 2, 3, \cdots, n_l, \tag{5}$$

$l = 1, 2, \cdots, a$, where the v_{lj}'s for $j = 1, 2, \cdots, n_l$, are also independent (in view of the diagonal covariance matrix of $\boldsymbol{v}_l$, noted above). We also note that the v_{lj}'s for $j = 1, 2, \cdots,$

n_l, are uncorrelated random variables, even if multivariate normality of the $\boldsymbol{y}_l$'s does not hold.

In the remainder of this section we shall test the equality of the σ_l^2's and the univariate normality of the v_{lj}'s against smooth alternatives defined through Legendre polynomials and Hermite polynomials. We shall do so assuming that the v_{lj}'s for $j = 1, 2, \cdots, n_l$, are all independent. That is, we are testing if the the v_{lj}'s for $j = 1, 2, \cdots, n_l$, are all independent normally distributed against the alternative that they are independent having a non-normal distribution defined through a smooth alternative. In other words, the class of smooth alternatives that we are considering is somewhat restricted in view of the independence assumption of the v_{lj}'s under the alternative. The advantage of the independence assumption is that the normality testing is now reduced to testing the univariate normality specified in (5). It should be noted that in an article on normality testing in a two-way random model with and without interaction, Xu, Li and Song (2013) reduced the problem to that of testing univariate normality, after transforming the data to uncorrelated components based on a transformation that depends on the unknown variance components. The authors then replaced the unknown variance components with estimates, and applied standard univariate normality tests, proceeding under the assumption that the transformed univariate components are independent even under the alternative. The transformation that we have used, based on the Helmert matrices Q_l, is of course parameter free.

2.1. Smooth alternatives based on Legendre polynomials

Let

$$z_{l1} = \frac{v_{l1} - \mu\sqrt{n_l}}{\sqrt{\sigma_l^2 + n_l\sigma_\alpha^2}}, \; u_{l1} = \Phi\left(z_{l1}\right), \; z_{lj} = v_{lj}/\sigma_l, \quad \text{and} \quad u_{lj} = \Phi\left(z_{lj}\right), \; j = 2, 3, \cdots, n_l, \tag{6}$$

where the z_{lj}'s are independent standard normal random variables for $j = 1, 2, \cdots, n_l$ and $l = 1, 2, \cdots, a$, and $\Phi(.)$ denotes the standard normal cdf. In this subsection, we shall specify smooth alternatives based on Legendre polynomials; we recall that these are polynomials that are orthonormal with respect to the uniform distribution in the interval (0, 1). Let $b_i(.)$, $i = 1, 2, \cdots$, denote the system of Legendre polynomials. It is easily verified that if $y \sim N(\mu, \sigma^2)$, then a system of orthonormal polynomials with respect to the $N(\mu, \sigma^2)$ density is obtained as $b_i\left(\Phi\left(\frac{y-\mu}{\sigma}\right)\right)$, $i = 1, 2, 3, \cdots$. In view of this, we conclude that for each fixed l and j, $b_i(u_{lj}) = b_i\left(\Phi(z_{lj})\right)$, $i = 1, 2, \cdots$, form a system of orthonormal polynomials with respect to the standard normal distribution, where $u_{lj} = \Phi\left(z_{lj}\right)$, as defined in (6). While specifying the smooth alternative, we will consider the case of only a common alternative across the a different groups. It is certainly possible to have different alternatives across the different groups, but we shall not consider this case.

In order to specify the likelihood function under a Legendre polynomial based smooth alternative, we note that the smooth alternative in (3) is specified in terms of a parameter vector $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_k)'$, and the null density corresponds to $\boldsymbol{\theta}$ being the null vector. Since we need to specify smooth alternatives for v_{l1}, $l = 1, 2, \cdots, a$, and for v_{lj}, $j = 2, 3, \cdots, n_l$, $l = 1, 2, \cdots, a$, where these quantities are defined in (4) and (5), we shall use two parameter vectors similar to $\boldsymbol{\theta}$. Thus let $\boldsymbol{\theta}_1 = (\theta_{11}, \theta_{12}, \cdots, \theta_{1k_1})'$, $\boldsymbol{\theta}_2 = (\theta_{21}, \theta_{22}, \cdots, \theta_{2k_2})'$, and $\boldsymbol{\theta} = (\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')'$. The vector $\boldsymbol{\theta}_1$ will be used to specify the smooth alternative for v_{l1}, $l =$ 1, 2, $\cdots$, a, and the vector $\boldsymbol{\theta}_2$ will be used to specify the smooth alternative for v_{lj}, $j = 2$, 3, $\cdots$, n_l, $l = 1$, 2, $\cdots$, a. Recall that we are assuming the independence of all the the v_{lj}'s, within and across the groups. The likelihood function under the assumption of a common alternative across the groups can be specified as

$$L = \prod_{l=1}^{a}\left[C(\boldsymbol{\theta},\mu,\boldsymbol{\sigma},\sigma_\alpha)\exp\left\{\sum_{r=1}^{k_1}\theta_{1r}b_r(u_{l1}) + \sum_{j=2}^{n_l}\sum_{s=1}^{k_2}\theta_{2s}b_s(u_{lj})\right\}f_{l1}(v_{l1},\mu,\sigma_l,\sigma_\alpha)\prod_{j=2}^{n_l}f_{l2}(v_{lj},\sigma_l)\right], \tag{7}$$

where $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \cdots, \sigma_a)'$, and for $l = 1$, 2, $\cdots$, a, $f_{l1}(v_{l1}, \mu, \sigma_l, \sigma_\alpha)$ and $f_{l2}(v_{lj}, \sigma_l)$, respectively, denote the normal density functions of v_{l1} and v_{lj} (j=2, 3, $\cdots$, n_l), when normality holds. The smooth alternatives that represent the departure from the normal distribution have the order k_1 for v_{l1} ($l = 1$, 2, $\cdots$, a), and order k_2 for v_{lj} ($l = 1$, 2, $\cdots$, a, j=2, 3, $\cdots$, n_l). The null hypothesis to be tested is that of normality and homoscedasticity. In terms of the parameters in (7) the hypothesis can be stated as

$$H_0:\ \boldsymbol{\theta} = (\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')' = \mathbf{0},\ \text{and}\ \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_a^2. \tag{8}$$

We shall find it convenient to transform the vector $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \cdots., \sigma_a)'$ using an $a \times a$ Helmert matrix $Q = ((q_{lc}))$ having the first column equal to $\frac{1}{\sqrt{a}}\mathbf{1}_a$, as done in Yang and Mathew (2018). We shall denote the transformed vector by $\boldsymbol{\eta}$. That is,

$$\boldsymbol{\eta} = Q'\boldsymbol{\sigma} = (\eta_1, \eta_2, \cdots, \eta_a)'. \tag{9}$$

It is easy to see that $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_a^2$ is equivalent to $\eta_2 = \eta_3 = \cdots = \eta_a = 0$. The log-likelihood function has the expression

$$\begin{aligned}\ln L &= \sum_{l=1}^{a}\Bigg[\ln\{C(\boldsymbol{\theta},\mu,\boldsymbol{\sigma},\sigma_\alpha)\} + \sum_{r=1}^{k_1}\theta_{1r}b_r(u_{l1}) + \sum_{j=2}^{n_l}\sum_{s=1}^{k_2}\theta_{2s}b_s(u_{lj}) + \ln\{f_{l1}(v_{l1},\mu,\sigma_l,\sigma_\alpha)\}\\ &\quad + \sum_{j=2}^{n_l}\ln\{f_{l2}(v_{lj},\sigma_l)\}\Bigg].\end{aligned} \tag{10}$$

In order to obtain the score statistic for testing the hypothesis in (8), we need the elements of the score vector and its variance-covariance matrix, evaluated under the null hypothesis. Explicit expressions can be obtained for these, and are given in the appendix. We note that some of the covariances are zeros. In the case of balanced data (ie., all the n_l, $l = 1$, 2, $\cdots$, a, are equal having a common value, say n), some additional covariances become zeros. These are also noted in the appendix.

While computing the score statistic, the unknown parameters are obviously replaced with their ML estimates under the null hypothesis. Thus we need the MLEs of μ, σ^2 and σ_α^2, where σ^2 is the common value of the σ_l^2's, under the null hypothesis. In the case of balanced data, we shall use the MLEs computed without imposing the nonnegativity constraint on the σ_α^2. If $\bar{y}_{..}$ denotes the average of all the y_{lj}s, and SS_e and SS_α, respectively, denote the sums of squares due to error and due to the α_i's under the model (1), the MLEs in the balanced case are given by

$$\hat{\mu} = \bar{y}_{..}, \ \hat{\sigma}^2 = \frac{SS_e}{a(n-1)} \ \text{ and } \ \hat{\sigma}_\alpha^2 = \frac{1}{n}\left[\frac{SS_\alpha}{a} - \frac{SS_e}{a(n-1)}\right], \tag{11}$$

where σ^2 denotes the common value of the σ_l^2s and n denotes the common value of the n_ls. As we shall see, the lack of the nonnegativity of $\hat{\sigma}_\alpha^2$ will not present any problems for us, since the estimator required in our application will be $\hat{\sigma}^2 + n\hat{\sigma}_\alpha^2$, which is always nonnegative (being equal to SS_α/a). Similar explicit estimates can also be obtained in the unbalanced case as follows. Define

$$\begin{aligned} \bar{y}_l &= \sum_{j=1}^{n_l} y_{lj}/n_l, \ \bar{\bar{y}} = \frac{1}{a}\sum_{l=1}^{a} \bar{y}_l, \ SS_e = \sum_{l=1}^{a}\sum_{j=1}^{n_l}(y_{lj} - \bar{y}_l)^2, \\ \tilde{n} &= a \times \left\{\frac{1}{n_1} + \frac{1}{n_2} + \cdots + \frac{1}{n_a}\right\}^{-1}, \ SS_\alpha = \tilde{n}\sum_{l=1}^{a}(\bar{y}_{l.} - \bar{\bar{y}})^2, \end{aligned} \tag{12}$$

where we note that $\tilde{n}$ is the harmonic mean of the n_l's. Such a formulation is due to Thomas and Hultquist (1978); see also Krishnamoorthy and Mathew (2009, Chapter 4). It can be verified that $E(\bar{\bar{y}}) = \mu$, $E(SS_e) = (N-a)\sigma^2$ and $E(SS_\alpha) = (a-1)(\tilde{n}\sigma_\alpha^2 + \sigma^2)$, where $N = \sum_{l=1}^{a} n_l$. The estimates of μ, σ^2 and σ_α^2 that we shall use are given by

$$\hat{\mu} = \bar{\bar{y}}, \ \hat{\sigma}^2 = SS_e/(N-a) \ \text{ and } \ \hat{\sigma}_\alpha^2 = \frac{1}{\tilde{n}}\left[\frac{SS_\alpha}{a} - \frac{SS_e}{N-a}\right]. \tag{13}$$

It should be noted that the estimates in (13) are not the MLEs. In fact the MLEs have no explicit expression and have to be numerically obtained in the unbalanced case. Nevertheless, for convenience we shall use the estimates in (13).

In order to give an expression for the score statistic, let us write the parameters in the order $(\boldsymbol{\theta}', \boldsymbol{\eta}^{*'}, \eta_1, \sigma_\alpha, \mu)'$, where $\boldsymbol{\eta}* = (\eta_2, \eta_3, \cdots, \eta_a)'$ (see (9)) and we recall the partitioning of $\boldsymbol{\theta}$ into the two components $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ of dimensions $k_1 \times 1$ and $k_2 \times 1$, respectively. Thus the null hypothesis in (8) is equivalent to

$$H_0: \ \boldsymbol{\theta} = (\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')' = \mathbf{0}, \ \text{and } \boldsymbol{\eta}* = 0. \tag{14}$$

The null hypothesis involves $(k_1 + k_2 + a - 1)$ parameters; in addition, we have three nuisance parameters, namely η_1, σ_α and μ. Now the the score vector has dimension $k_1 + k_2 + a + 2$, which is the total number of parameters under the model (7). Consequently, if V denotes the variance-covariance matrix of the score vector, whose elements are arranged according to the parameter order $(\boldsymbol{\theta}', \boldsymbol{\eta}^{*'}, \eta_1, \sigma_\alpha, \mu)'$, then clearly V has dimension $(k_1 + k_2 + a + 2) \times (k_1 + k_2 + a + 2)$. The elements of the score vector, and those of V are given in the appendix, where the expressions have been simplified assuming the null hypothesis (14).

Let us consider a partitioning of V as

$$V = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix},$$

where the dimension of V_{11} is $(k_1 + k_2 + a - 1) \times (k_1 + k_2 + a - 1)$, corresponding to the parameters $\boldsymbol{\theta}$ and $\boldsymbol{\eta}*$ in the null hypothesis (14), and the dimensions of the remaining blocks of V should be clear. Let $\boldsymbol{u}$ denote the score vector and $\boldsymbol{u}_1$ denote the first (k_1+k_2+a-1) elements of $\boldsymbol{u}$ corresponding to the parameters $\boldsymbol{\theta}$ and $\boldsymbol{\eta}*$, and define $V_{11.2} = V_{11} - V_{12}V_{22}^{-1}V_{21}$. Recall that $\boldsymbol{u}_1$ and $V_{11.2}$ are functions of the nuisance parameters η_1, σ_α and μ; equivalently σ, σ_α and μ (since, in view of (9), $\eta_1 = \sqrt{a} \times \sigma$ under the null hypothesis). The score statistic for testing the hypothesis (8), or equivalently (14), is given by

$$\hat{S} = \hat{\boldsymbol{u}}_1' \hat{V}_{11.2}^{-1} \hat{\boldsymbol{u}}_1, \tag{15}$$

where we have used the notations $\hat{\boldsymbol{u}}_1$ and $\hat{V}_{11.2}$ to emphasize that the unknown nuisance parameters σ, σ_α and μ have been replaced by estimates computed assuming the null hypothesis; we shall use the estimates of σ^2, σ_α^2 and μ, exhibited earlier in this section. For fixed k_1 and k_2, the score statistic $\hat{S}$ in (15) has an approximate chisquare distribution with df $= (k_1 + k_2 + a - 1)$ under the null hypothesis.

In order to implement the test based on the score statistic given in (15), it is necessary to choose the orders k_1 and k_2 in the likelihood function (7). For this, we shall follow a data-driven approach; *i.e.*, estimate the orders based on the data, as done in Inglot, Kallenberg and Ledwina (1994). Such an approach was also adopted in Yang and Mathew (2018) in a fixed effects linear model. Here we shall only present the relevant expressions that will facilitate the numerical computation of k_1 and k_2, referring to the original articles for details of the methodology and the associated theoretical results. A brief explanation of the methodology is also given in Yang (2016).

Referring to the quantities defined in (6), let

$$\hat{z}_{l1} = \frac{v_{l1} - \hat{\mu}\sqrt{n_l}}{\sqrt{\hat{\sigma}^2 + n_l\hat{\sigma}_\alpha^2}}, \ \hat{u}_{l1} = \Phi\left(\hat{z}_{l1}\right), \ \hat{z}_{lj} = \hat{v}_{lj}/\hat{\sigma}, \ \text{ and } \ \hat{u}_{lj} = \Phi\left(\hat{z}_{lj}\right), \ j = 2, 3, \cdots, n_l, \tag{16}$$

where $\hat{\sigma}$, $\hat{\sigma}_\alpha$ and $\hat{\mu}$ are the estimates obtained under the null hypothesis, and used in the computation of the score statistic $\hat{S}$ in (15). Let $\hat{H}_{1,k_1}(\hat{\sigma}, \hat{\sigma}_\alpha, \hat{\mu})$ and $\hat{H}_{2,k_2}(\hat{\sigma})$ be defined as

$$\begin{aligned} \hat{H}_{1,k_1}(\hat{\sigma}, \hat{\sigma}_\alpha, \hat{\mu}) &= \frac{1}{a}\sum_{i=1}^{k_1}\left\{\sum_{l=1}^{a} b_i\left(\hat{u}_{l1}\right)\right\}^2 \\ \hat{H}_{2,k_2}(\hat{\sigma}) &= \frac{1}{N-a}\sum_{i=1}^{k_2}\left\{\sum_{l=1}^{a}\sum_{j=2}^{n_l} b_i\left(\hat{u}_{lj}\right)\right\}^2, \end{aligned} \tag{17}$$

where $N = \sum_{l=1}^{a} n_l$, the $b_i(.)$'s are Legendre polynomials and the remaining quantities are defined earlier in this section. We note that the divisor a in the expression for $\hat{H}_{1,k_1}$ is the number of $\hat{z}_{l1}$s, $l = 1, 2, \cdots, a$, and the divisor $N - a$ in the expression for $\hat{H}_{2,k_2}$ is the number of $\hat{z}_{lj}$s, $j = 2, 3, \cdots, n_l$, $l = 1, 2, \cdots, a$. Assuming upper bounds d_1 for k_1 and d_2 for k_2, the orders k_1 and k_2, say $\hat{k}_1$ and $\hat{k}_2$, are determined as follows:

$$\begin{aligned} \hat{k}_1 &= \min\{k_1 : 1 \le k_1 \le d_1, \hat{H}_{1,k_1}(\hat{\sigma}, \hat{\sigma}_\alpha, \hat{\mu}) - k_1\ln(a) \ge \hat{H}_{1,r}(\hat{\sigma}, \hat{\sigma}_\alpha, \hat{\mu}) - r\ln(a), r = 1, \cdots, d_1\} \\ \hat{k}_2 &= \min\{k_2 : 1 \le k_2 \le d_2, \hat{H}_{2,k_2}(\hat{\sigma}) - k_2\ln(N-a) \ge \hat{H}_{2,s}(\hat{\sigma}) - s\ln(N-a), s = 1, \cdots, d_2\}. \end{aligned} \tag{18}$$

In order to determine $\hat{k}_1$ and $\hat{k}_2$ according to the expressions in (18), we need to select the bounds d_1 and d_2. For this, we shall make use of a result mentioned in Kallenberg and Ledwina (1997b), citing Inglot, Kallenberg and Ledwina (1994), which states that the data driven order converges to the value 1 in probability. Even though our set up is different from that under which this result is proved, we proceed under the assumption that the data driven approach is unlikely to yield values of $\hat{k}_1$ and $\hat{k}_2$ that are far removed from the value 1. In our simulations, we chose $d_1 = d_2 = 6$.

Algorithm 1 given below gives a summary of the steps necessary to implement our proposed test for testing the null hypothesis in (8) against a Legendre polynomial based smooth alternative.

Algorithm 1

1. Compute the estimates of $\hat{\mu}$, $\hat{\sigma}^2$ and $\hat{\sigma}^2_\alpha$ given in (13).
2. Compute the $n_l \times 1$ vectors $\boldsymbol{v}_l = (v_{l1}, v_{l2}, \cdots, v_{ln_l})'$ given in (7).
3. Compute the $\hat{u}_{lj}$ given in (16), which requires the quantities computed in the previous two steps.
4. Compute $\hat{k}_1$ and $\hat{k}_2$ in (18) where $\hat{H}_{1,k_1}(\hat{\sigma}, \hat{\sigma}_\alpha, \hat{\mu})$ and $\hat{H}_{2,k_2}(\hat{\sigma})$ are given in (17). For the quantities d_1 and d_2 in (18), we recommend the values $d_1 = d_2 = 6$.
5. Compute the score vector and its variance-covariance matrix using the expressions given in the appendix, and replace μ, σ^2 and σ^2_α with $\hat{\mu}$, $\hat{\sigma}^2$ and $\hat{\sigma}^2_\alpha$, respectively. Let the quantities so obtained be denoted by $\hat{\boldsymbol{u}}$ and $\hat{V}$, respectively.
6. Partition $\hat{\boldsymbol{u}}$ and $\hat{V}$ as

$$\hat{\boldsymbol{u}} = \left(\hat{\boldsymbol{u}}_1', \hat{\boldsymbol{u}}_2'\right)' \text{ and } \hat{V} = \begin{pmatrix} \hat{V}_{11} & \hat{V}_{12} \\ \hat{V}_{21} & \hat{V}_{22} \end{pmatrix},$$

   where $\hat{\boldsymbol{u}}_1$ is a $(\hat{k}_1 + \hat{k}_2 + a - 1) \times 1$ vector, and $\hat{V}_{11}$ is a $(\hat{k}_1 + \hat{k}_2 + a - 1) \times (\hat{k}_1 + \hat{k}_2 + a - 1)$ matrix.
7. Compute $\hat{V}_{11.2} = \hat{V}_{11} - \hat{V}_{12}\hat{V}_{22}^{-1}\hat{V}_{21}$ and the score statistic $\hat{S} = \hat{\boldsymbol{u}}_1'\hat{V}_{11.2}^{-1}\hat{\boldsymbol{u}}_1$. Reject H_0 in (8) if the value of the score statistic exceeds the appropriate percentile of the chisquare distribution with df $= (\hat{k}_1 + \hat{k}_2 + a - 1)$.

For the case of balanced data, the n_l's have to be replaced with their common value, say n, in all the expressions. As a result, some of the covariances will become zeros, and these are noted in the appendix. It can then be verified that in the balanced case the matrix $\hat{V}_{11.2}$ is a block-diagonal matrix, with three diagonal blocks having dimensions $\hat{k}_1 \times \hat{k}_1$, $\hat{k}_2 \times \hat{k}_2$ and $(a-1) \times (a-1)$, corresponding to $\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_2$ and $\boldsymbol{\eta}*$, where $\boldsymbol{\eta}* = (\eta_2, \eta_3, \cdots, \eta_a)'$. Accordingly, the score statistic splits into three components. Thus when the null hypothesis is rejected, it is possible to draw conclusions on which component/components contributed to the rejection:

the non-normality of v_{l1}'s ($l = 1, 2, \cdots, a$), the non-normality of v_{lj}'s ($j = 2, 3, \cdots, n_l$, $l = 1, 2, \cdots, a$), or the heteroscedasticity.

It turns out that the distributions of the score vector $\hat{\boldsymbol{u}}_1$ and the submatrix of $\hat{V}_{11.2}$ corresponding to $\boldsymbol{\eta}*$ depend on σ and σ_α. However, in the balanced case we have

$$\mathrm{Var}\left(\frac{\partial \ln L}{\partial \eta_c}\right) = 2\left[\frac{\sigma^2}{(\sigma^2+n\sigma_\alpha^2)^2} + \frac{(n-1)}{\sigma^2}\right] = 2\left[\frac{1}{\sigma^2(1+n\lambda)^2} + \frac{(n-1)}{\sigma^2}\right],$$

$c = 1, 2, \cdots, a$, where $\lambda = \frac{\sigma_\alpha^2}{\sigma^2}$. We note that the above variance varies between $(n-1)/\sigma^2$ and n/σ^2, as λ varies between 0 and ∞. Thus the impact of σ_α^2 on the score test appears to be fairly small in the balanced case, and we may anticipate this to be so in the unbalanced case as well. Later we shall examine this further through numerical results.

2.2. Smooth alternatives based on Hermite polynomials

We shall now consider the likelihood function under a smooth alternative based on Hermite polynomials. The likelihood is similar to (7) except that Hermite polynomials are used instead of Legendre polynomials. Thus let $h_i(z)$, $i = 1, 2, \cdots$, denote the system of Hermite polynomials. The log-likelihood is now given by

$$\begin{aligned}\ln L \;=\; & \sum_{l=1}^{a}\Bigg[\ln\{C(\boldsymbol{\theta},\mu,\boldsymbol{\sigma},\sigma_\alpha)\} + \sum_{r=1}^{k_1}\theta_{1r}h_r(z_{l1}) + \sum_{j=2}^{n_l}\sum_{s=1}^{k_2}\theta_{2s}h_s(z_{lj}) + \ln\{f_{l1}(v_{l1},\mu,\sigma_l,\sigma_\alpha)\} \\ & + \sum_{j=2}^{n_l}\ln\{f_{l2}(v_{lj},\sigma_l)\}\Bigg], \qquad (19)\end{aligned}$$

where the z_{l1} and z_{lj} are defined in (6). The components of the score vector (under the null hypothesis) can be derived similar to the Legendre case, and are given in the appendix. The elements of the variance covariance matrix of the score vector are also given in the appendix. The score statistic can be worked out similar to (15). It can also be verified that the matrix analogous to $V_{11.2}$ is a block diagonal matrix, having three diagonal blocks. Furthermore, in the balanced case, $V_{11.2}$ will reduce to a completely diagonal matrix. Recall that under the Legendre polynomial based alternative, $V_{11.2}$ simplified to a block diagonal matrix only under balanced data.

We can compute data driven choices of the orders k_1 and k_2 under the Hermite polynomial based smooth alternative also. For this we need to define quantities analogous to those in (17), with the Legendre polynomial terms replaced by the corresponding Hermite polynomial based terms. The orders can then be determined proceeding as in (18). However, when we define the analogous quantities in (17) for the Hermite polynomial case, the summations will be from $i = 3$ to k_1 and $i = 3$ to k_2 (instead of $i = 1$ to k_1 and $i = 1$ to k_2). For this, we need to show that the terms corresponding to $i = 1$ and $i = 2$ are zeros. Actually they are exactly zeros in the balanced case, and we shall choose to ignore them in the unbalanced case since they are likely to be quite small. Such zero-terms in the context of the Hermite polynomial based alternative have been noted, for example, in Rayner, Thas and Best (2009) in the context of testing univariate normality.

Similar to (17), let's define $\tilde{H}_{1,k_1}(\hat{\sigma},\hat{\sigma}_\alpha,\hat{\mu})$ and $\tilde{H}_{2,k_2}(\hat{\sigma})$ as

$$\begin{aligned}\tilde{H}_{1,k_1}(\hat{\sigma},\hat{\sigma}_\alpha,\hat{\mu}) &= \frac{1}{a}\sum_{i=3}^{k_1}\left\{\sum_{l=1}^{a} h_i\left(\hat{z}_{l1}\right)\right\}^2 \\ \tilde{H}_{2,k_2}(\hat{\sigma}) &= \frac{1}{N-a}\sum_{i=3}^{k_2}\left\{\sum_{l=1}^{a}\sum_{j=2}^{n_l} h_i\left(\hat{z}_{lj}\right)\right\}^2, \end{aligned} \tag{20}$$

where the $h_i(.)$'s are Hermite polynomials and $\hat{z}_{lj}$'s are defined in (16). Assuming upper bounds d_1 for k_1 and d_2 for k_2, the orders k_1 and k_2, say $\tilde{k}_1$ and $\tilde{k}_2$, are determined as follows:

$$\begin{aligned}\tilde{k}_1 &= \min\{k_1 : 3 \le k_1 \le d_1, \tilde{H}_{1,k_1}(\hat{\sigma},\hat{\sigma}_\alpha,\hat{\mu}) - k_1\ln(a) \ge \tilde{H}_{1,r}(\hat{\sigma},\hat{\sigma}_\alpha,\hat{\mu}) - r\ln(a), r = 3,\cdots,d_1\} \\ \tilde{k}_2 &= \min\{k_2 : 3 \le k_2 \le d_2, \tilde{H}_{2,k_2}(\hat{\sigma}) - k_2\ln(N-a) \ge \tilde{H}_{2,s}(\hat{\sigma}) - s\ln(N-a), s = 3,\cdots,d_2\}.\end{aligned} \tag{21}$$

Let's now consider the case of balanced data and show that the terms corresponding to $i=1$ and $i=2$ are zeros in $\tilde{H}_{1,k_1}(\hat{\sigma},\hat{\sigma}_\alpha,\hat{\mu})$ and $\tilde{H}_{2,k_2}(\hat{\sigma})$ in (20), so that in the definition of these quantities the summation can start from $i=3$, as done in (20). Using (4), (6) and (16), and recalling that the Helmert matrix is an orthogonal matrix with first column being a multiple of a vector of ones, we have the following simplifications in the balanced case, under the null hypothesis,

$$\begin{aligned}& v_{l1} = \sqrt{n}\bar{y}_l,\ \hat{z}_{l1} = \frac{\sqrt{n}(\bar{y}_l - \bar{y}_{..})}{\sqrt{aSS_\alpha}},\ \sum_{l=1}^{a}\hat{z}_{l1}^2 = \frac{1}{a} \\ & \sum_{j=1}^{n} y_{lj}^2 = \sum_{j=1}^{n} v_{lj}^2 = n\bar{y}_l^2 + \sum_{j=2}^{n} v_{lj}^2,\ \sum_{j=2}^{n} v_{lj}^2 = \sum_{j=1}^{n} y_{lj}^2 - n\bar{y}_l^2 \\ & \sum_{l=1}^{a}\sum_{j=2}^{n} v_{lj}^2 = \sum_{l=1}^{a}\left[\sum_{j=1}^{n} y_{lj}^2 - n\bar{y}_l^2\right] = SS_e \\ & \sum_{l=1}^{a}\sum_{j=2}^{n}\hat{z}_{lj}^2 = \frac{1}{a(n-1)},\end{aligned} \tag{22}$$

where we have used the expression $SS_\alpha = n\sum_{l=1}^{a}(\bar{y}_l - \bar{y}_{..})^2$. The first two Hermite polynomials are given by

$$h_1(z) = z, \quad \text{and} \quad h_2(z) = \frac{1}{\sqrt{2}}(z^2 - 1).$$

From the observations in (22), it now follows that for balanced data, $\sum_{l=1}^{a} h_i(z_{l1}) = 0$, $i = 1, 2$, and $\sum_{l=1}^{a}\sum_{j=2}^{n} h_i(z_{lj}) = 0$, $i = 1, 2$. It should be noted that this conclusion holds only for the case of balanced data. For unbalanced data, we anticipate that these terms will be small even though they may not be exactly equal to zero.

Once $\tilde{k}_1$ and $\tilde{k}_2$ are determined as in (21), we note that the number of Hermite polynomial terms in $\tilde{H}_{1,\tilde{k}_1}(\hat{\sigma},\hat{\sigma}_\alpha,\hat{\mu})$ and $\tilde{H}_{2,\tilde{k}_2}(\hat{\sigma})$ are $\tilde{k}_1 - 2$ and $\tilde{k}_2 - 2$, respectively. Hence the score statistic will have an approximate chisquare distribution with df $= (\tilde{k}_1 - 2) + (\tilde{k}_2 - 2) + (a-1)$.

3. Numerical Results

We shall now report some numerical results to assess the behavior of the score tests we have proposed in the previous sections. Our purpose here is several: to assess the dependence of the score tests on σ_α^2, to examine the accuracy of the asymptotic chisquare distribution (under the null), and to compare the powers of the score tests based on the Legendre polynomial alternative and based on the Hermite polynomial alternative.

In order to the examine the dependence of the tests on σ_α^2, we first estimated the 95th percentiles of the score statistic in the case of balanced data, for various values of a and n, and for $\sigma_\alpha^2 = 0.1$, 2 and 10. We also chose $\mu = 0$ and $\sigma^2 = 1$, where σ^2 denotes the common value of the σ_l^2's. The data driven approach explained in the previous section was used to obtain the order of the polynomials. Most of the time, the data driven choices $\hat{k}_1$ and $\hat{k}_2$ were equal to one in the Legendre polynomial case, and $\tilde{k}_1$ and $\tilde{k}_2$ were equal to three in the Hermite polynomial case, so that the score statistic has an approximate chisquare distribution with df $= a+1$ under the null hypothesis. Table 1 gives the estimated percentiles based on 10^4 simulated samples. We have also included the 95th percentile of the chisquare distribution with df $= a + 1$. We can draw the following conclusions from the numerical results in Table 1: (i) the chisquare distribution approximates the null distribution of the score statistic reasonable well; however, the actual percentiles are slightly larger than that of the chisquare distribution, and (ii) the null distribution is not sensitive to the value of σ_α^2. It appears that in order to have a more accurate test, one can use a Monte Carlo estimate of the corresponding percentile (instead of using the chisquare percentile) after simply assuming $\sigma_\alpha^2 = 1$, regardless of the true value of σ_α^2.

In order to further see the insensitivity of the null distribution with respect to the value of σ_α^2, we estimated the type I error probabilities of the score test when $\sigma_\alpha^2 = 0.1$ and 10, when the test is carried out using the estimated critical value (*i.e.*, 95th percentile) corresponding to $\sigma_\alpha^2 = 2$. The rest of the simulation set up is the same as that used to obtain the results in Table 1. The type I error probabilities are given in Table 2. The insensitivity of the type I error probabilities with respect to the value of σ_α^2 should be clear.

In addition, we looked at Type I error probabilities for unbalanced cases, using critical values estimated from the balanced case with a common value chosen as $\tilde{n}$, which is the harmonic mean of the n_l's; see (13). For this we used a= 10, 20, 30 and 50 groups. We also made three choices in terms of severity of the unbalancedness: mild, moderate and severe. For $a = 10$, our choice of the n_l's to represent mild unbalancedness is (4,4,5,5,5,5,5,5,5,10), which gives the harmonic mean $\tilde{n}$=5. For moderate unbalancedness, we chose the n_l's to be (2,3,4,5,6,7,8,9,10,14), resulting in $\tilde{n}$=4.999008 (we shall take $\tilde{n} = 5$ in this case). For the severe unbalanced case, we made the choice (3,3,4,4,5,5,5,5,40,120) for the n_l's, which also yields $\tilde{n}$=5. For $a = 20$, we adopted the above choices except that each n_l value was repeated twice. For $a = 30$ and 50, the same strategy was followed; *i.e.*, each n_l value chosen for $a =$ 10 was repeated three times and five times each.

The choices of the n_ls just described all resulted in $\tilde{n} = 5$, as already noted. We shall also consider choices that will give $\tilde{n} = 10$, 30 and 50. For this, we multiplied each of the n_ls in the earlier choices with 2, 6 and 10, so as to result in $\tilde{n} = 10$, 30 and 50, respectively.

Table 1: Monte Carlo estimates of the 95th percentiles of the score statistic for σ_α^2= 0.1, 2, 10; $\chi^2_{(a+1,0.95)}$ denotes the 95th percentile of the chisquare distribution with $a+1$ df

		Legendre			Hermite			
a	n	σ_α^2=0.1	σ_α^2=2	σ_α^2=10	σ_α^2=0.1	σ_α^2=2	σ_α^2=10	$\chi^2_{(a+1,0.95)}$
10	5	21.71	21.70	21.20	19.57	20.05	20.17	19.68
10	10	21.02	20.99	20.70	20.09	20.52	20.49	19.68
10	30	20.51	20.49	20.43	20.16	20.33	20.39	19.68
10	50	20.80	20.80	20.79	20.56	20.49	20.49	19.68
20	5	35.25	35.16	34.32	34.56	35.97	35.96	32.67
20	10	33.34	33.34	33.35	35.11	35.42	35.31	32.67
20	30	33.29	33.30	33.25	34.22	34.40	34.46	32.67
20	50	33.25	33.24	33.20	33.95	34.08	34.13	32.67
30	5	47.87	47.78	46.83	48.58	49.45	49.48	44.99
30	10	46.82	46.84	46.83	48.16	48.15	48.23	44.99
30	30	45.51	45.52	45.53	47.43	47.18	47.14	44.99
30	50	45.72	45.72	45.71	46.79	46.63	46.64	44.99
50	5	72.99	72.69	71.64	75.35	76.03	75.88	68.67
50	10	70.67	70.70	70.88	73.54	73.86	73.90	68.67
50	30	69.34	69.34	69.34	71.35	71.39	71.35	68.67
50	50	69.63	69.65	69.64	70.61	70.72	70.56	68.67

Table 2: Estimated type I error probabilities of the score test carried out using the estimated critical value corresponding to $\sigma_\alpha^2 = 2$, for a 5% significance level

		Legendre		Hermite	
a	n	σ_α^2=0.1	σ_α^2=10	σ_α^2=0.1	σ_α^2=10
10	5	0.053	0.050	0.046	0.051
10	10	0.050	0.050	0.046	0.050
10	30	0.049	0.050	0.049	0.051
10	50	0.052	0.050	0.051	0.050
20	5	0.058	0.051	0.044	0.050
20	10	0.045	0.050	0.048	0.050
20	30	0.049	0.050	0.049	0.051
20	50	0.051	0.050	0.049	0.051
30	5	0.051	0.051	0.046	0.050
30	10	0.055	0.050	0.050	0.050
30	30	0.048	0.050	0.052	0.050
30	50	0.053	0.050	0.051	0.050
50	5	0.051	0.051	0.048	0.050
50	10	0.047	0.050	0.049	0.050
50	30	0.049	0.050	0.050	0.050
50	50	0.055	0.050	0.050	0.049

We used the value σ_α^2=2 to estimate the type I error probabilities, and considered both Legendre polynomial-based and Hermite polynomial-based alternatives. The estimated type I error probabilities are given in Table 3 (under Legendre polynomial-based alternatives) and in Table 4 (under Hermite polynomial-based alternatives). The following conclusions can be drawn from the numerical results. Under the Legendre polynomial-based alternative, there is no Type I error inflation for the mild unbalanceness cases; we notice a somewhat mild type I error inflation in the case of moderate unbalancedness when $\tilde{n}$ is small, and a more pronounced type I error inflation in the case of severe unbalancedness when $\tilde{n}$ is small. However, when $\tilde{n}$ is 30 or more, the type I error probabilities are all close to the nominal level of 5%. Under the Hermite polynomial-based alternative, the type I error inflation appears to be a bit more severe, especially in the case of severe unbalancedness and small $\tilde{n}$. However, the results are once again quite satisfactory when $\tilde{n}$ is 30 or more. Overall, using the balanced set up critical value based on $\tilde{n}$ appears to be a satisfactory option, except in the severe unbalanced case and a small $\tilde{n}$.

Table 3: Type I error probabilities for unbalanced data under Legendre polynomial-based alternatives and using the balanced data critical value with $n = \tilde{n}$, for a 5% significance level

a	$\tilde{n}$	Unbalancedness: Severe	Unbalancedness: Moderate	Unbalancedness: Mild	Estimated critical value when $\sigma_\alpha^2 = 2$	Chisquare critical value
10	5	0.065	0.047	0.042	21.702	19.675
10	10	0.060	0.052	0.048	20.987	19.675
10	30	0.056	0.048	0.049	20.490	19.675
10	50	0.043	0.045	0.044	20.803	19.675
20	5	0.069	0.054	0.047	35.165	32.671
20	10	0.065	0.055	0.055	33.336	32.671
20	30	0.048	0.052	0.047	33.299	32.671
20	50	0.044	0.049	0.049	33.244	32.671
30	5	0.076	0.062	0.049	47.776	44.985
30	10	0.064	0.051	0.046	46.841	44.985
30	30	0.053	0.048	0.052	45.516	44.985
30	50	0.048	0.041	0.044	45.723	44.985
50	5	0.077	0.061	0.053	72.693	68.669
50	10	0.064	0.057	0.047	70.697	68.669
50	30	0.052	0.050	0.050	69.338	68.669
50	50	0.042	0.046	0.043	69.645	68.669

The type I error probabilities reported in Table 3 and Table 4 were computed when the test was carried out using the estimated critical value corresponding to $\sigma_\alpha^2 = 2$; these critical values are also given in the tables. The tables also give the chisquare critical values. We note that the chisquare critical values are smaller than the estimated critical values, as was noted in Table 1. Thus if the test is carried out using the asymptotic chisquare distribution, one should expect an inflated type I error probability.

Table 4: Type I error probabilities for unbalanced data under Hermite polynomial-based alternatives and using the balanced data critical value with $n = \tilde{n}$, for a 5% significance level

a	$\tilde{n}$	Unbalancedness Severe	Moderate	Mild	Estimated critical value when $\sigma_\alpha^2 = 2$	Chisquare critical value
10	5	0.093	0.063	0.053	20.052	19.675
10	10	0.066	0.059	0.052	20.516	19.675
10	30	0.058	0.053	0.052	20.327	19.675
10	50	0.046	0.049	0.049	20.495	19.675
20	5	0.082	0.060	0.051	35.972	32.671
20	10	0.065	0.053	0.051	35.420	32.671
20	30	0.051	0.054	0.050	34.400	32.671
20	50	0.046	0.052	0.052	34.080	32.671
30	5	0.080	0.066	0.052	49.451	44.985
30	10	0.066	0.059	0.054	48.149	44.985
30	30	0.054	0.049	0.054	47.175	44.985
30	50	0.050	0.051	0.051	46.626	44.985
50	5	0.075	0.062	0.053	76.031	68.669
50	10	0.058	0.056	0.048	73.862	68.669
50	30	0.053	0.051	0.054	71.387	68.669
50	50	0.046	0.056	0.051	70.715	68.669

Some limited numerical results on the power are reported in Table 5 in the case of balanced data for $a = 10$ and $n = 5$ and $a = 30$ and $n = 5$. We note that the null hypothesis can be violated by having a non-normal distribution for the errors and/or random effects, and/or by having heteroscedasticity of the error distribution. In Table 5, the powers are first reported when random effects are normally distributed and the error terms alone are non-normal and/or heteroscedastic (the first few rows of the table). The last few rows of the table correspond to the alternative scenario where the error terms are normally distributed and could be heteroscedastic, but the random effects are non-normal. The very last row of the table corresponds to the alternative where both the error terms and the random effects follow non-normal distributions, but the errors are homoscedastic. A 5% significance level and estimated critical values are used while computing the power. For introducing heteroscedasticity into the alternative, we proceeded as follows. We split the a groups into two sets, having $a/2$ groups in each set (we have chosen only an even value of a in our simulations). Data are generated from the same alternative error distribution, except that for the data in the second set, the randomly generated error term was multiplied by $\sqrt{2}$, which will result in twice the error variance for the data in the second set, compared to those in the first set. The results on the power show that most of the powers are comparable when the tests are derived using a Legendre polynomial-based alternative or a Hermite polynomial-based alternatives. However, the test derived under the Hermite polynomial-based alternatives appears to have a slight edge in terms of power. Perhaps more extensive simulation are necessary before we can draw firm conclusions.

Table 5: Estimated powers of the Legendre polynomial based and Hermite polynomial based tests using estimated critical values for a 5% significance level

		a=10, n=5		a=30, n=5	
Alternative		Legendre	Hermite	Legendre	Hermite
Error	N(0,1)	0.04	0.05	0.05	0.05
	N(0,1)*	0.11	0.13	0.22	0.21
	t(5)	0.24	0.29	0.56	0.57
	t(5)*	0.30	0.36	0.70	0.70
	Gamma (2,1)	0.28	0.35	0.65	0.68
	Gamma (2,1)*	0.36	0.44	0.78	0.80
Random Effect	t(2)	0.17	0.14	0.41	0.48
	t(2)*	0.23	0.21	0.54	0.59
	Gamma (2,1)	0.09	0.07	0.16	0.18
	Gamma (2,1)*	0.15	0.15	0.36	0.35
Error + Random Effect	Gamma (2,1) + t(2)	0.40	0.43	0.79	0.84

*Heteroscedasticity

4. An Example

We use a quality control data set from clinical chemistry on serum sodium measurements. The data are taken from Andrews and Herzberg (1985), and serum sodium measurements are given for 10 specimens tested by 100 labs. The specimens are from a large homogeneous pool of serum, and one specimen is sent to the labs every two weeks. Here we shall use only a subset of the data, and these data are from 24 labs that used the same analysis method (Method 5 mentioned in Andrews and Herzberg (1985)), and had all 10 specimens tested, so that we have balanced data with a=10 and n=24; the data we used are given in Yang (2016). The results of the data analysis are presented in Table 6. Data driven choices were made for the orders k_1 and k_2. For a 5% significance level, the estimated critical values necessary to carry out the test are given in Table 6. The data driven choices of the orders are also given in the table. The upper bounds d_1 and d_2 were chosen as $d_1 = d_2$=6. We noted earlier that that for one-way random model with balanced data, the matrix $V_{11.2}$ used to compute the score statistic is a block-diagonal matrix consisting of 3 blocks; if the null hypothesis of normality and homoscedasticity is rejected, it is possible to draw conclusions on which component/components contributed to the rejection: normality or homoscedasticity of the error distribution, or the normality of the random effect. We note from Table 6 that the null hypothesis is rejected by the tests based on both Legendre polynomial-based and Hermite-polynomial-based alternatives. In the table, we have also given the decomposition of the score statistic into the three components; the first component corresponds to normality of the random effects, the second and third components correspond, respectively, to normality and heteroscedasticity of the error distribution. The results indicate that there is evidence for both non-normality and heteroscedasticity for the error distribution, but there is no evidence of non-normality of the random effects.

Table 6: Analysis of the serum sodium data using a 5% significance level

Polynomial	Orders	Estimated critical value	Score statistic	Decision
Legendre	$\hat{k}_1 = 1$, $\hat{k}_2 = 2$	20.584	$55.505 = 0.066 + 32.776 + 22.663$	Reject H_0
Hermite	$\tilde{k}_1 = 3$, $\tilde{k}_2 = 6$	20.388	$251.439 = 0.024 + 228.752 + 22.663$	Reject H_0

5. Discussion: Possible Extensions and Limitations

Data analysis in a linear model framework relies on several assumptions: normality, homoscedasticity (especially when the data fall into different groups), independence, and the assumption that the mean vector belongs to a specified subspace. Simultaneous assessment of these assumptions is clearly of interest. An attempt in this direction has been made by Peña and Slate (2006). Their assessment of normality assumes a smooth alternative. The assessment of homoscedasticity assumes that departure from this assumption can be modeled as a function of the mean. However, the violation of the homoscedasticity need not imply that the variance changes with the mean, even though this is a possibility. In our work, we have explored the simultaneous assessment of both normality and homoscedasticity. It should be noted that the smooth tests available in the literature address only the problem of testing the adequacy of a parametric distribution, specified in terms of appropriate orthogonal polynomials. However, Kallenberg, Ledwina and Rafajlowicz (1997) did address the problem of simultaneously testing normality and independence for bivariate data. Such simultaneous testing has been facilitated by having a parametric form under the smooth alternative.

We want to point out several limitations, and some possible generalizations, of the work reported here. In the context of models that involve random effects, if we want to go beyond the one-way random model, difficulties do arise for the simultaenous assessment of normality of the random effects, normality of the error terms, and homoscedasticity. Unbalancd data will add further complications. However, we feel that our methodology in the context of the one-way random model can be generalized to general mixed or random effects models for the simultaneous assessment of normality of the random effects and the error terms, assuming that homoscedasticity holds, provided we have *balanced* data. We shall now illustrate this in the context of the two-way random effects model and the two-way mixed effects models, when the model includes interactions.

5.1. Testing normality in a two-way random model with balanced data

Consider two factors having a and b levels, randomly selected, and suppose we have n observations one each level combination; thus we have balanced data. Let y_{ijm} denote the mth observation corresponding to the ith and jth levels of the two factors; $m = 1, 2, \cdots, n$. The two-way model with interaction is given by

$$y_{ijm} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijm},$$

$i = 1, \cdots, a$, $j = 1, \cdots, b$ and $m = 1, \cdots, n$. We assume that all the effects are random and the random variables α_is, β_js, γ_{ij}s and the error terms ϵ_{ijm}s are all independent. The usual normality assumptions state that

$$\alpha_i \sim N(0, \sigma_\alpha^2), \quad \beta_j \sim N(0, \sigma_\beta^2), \quad \gamma_{ij} \sim N(0, \sigma_\gamma^2), \quad \text{and} \quad \epsilon_{ijm} \sim N(0, \sigma^2).$$

Note that we are now assuming homoscedasticity. The above normality assumptions imply a multivariate normal distribution with a structured variance-covariance matrix for the data vector consisting of all the y_{ijm}; the mean vector is simply $\mu \mathbf{1}_{abn}$. Our goal is to test if such a multivariate normality assumption holds, by using an appropriate smooth alternative. We recall that in the one-way random model, this was accomplished by transforming the data vector into two uncorrelated components, and simultaneously testing univariate normality for each component assuming independent smooth alternatives. We shall do the same under the two-way random model as well. However, now four sets of independent random variables are involved in the model (α_i's, β_j's, γ_{ij}'s and the ϵ_{ijm}'s). Consequently, we will suitably transform the data vector, and end up with four uncorrelated components. We can then simultaneously test univariate normality of the four components against the assumption of smooth alternatives that are also independent.

Consider the vector $\boldsymbol{y}_{ij} = (y_{ij1}, y_{ij2}, \cdots, y_{ijn})'$, and let $Q_n = \left(\frac{1}{\sqrt{n}} 1_n, Q_n^*\right)$ be an $n \times n$ Helmert matrix. Consider the transformation

$$\boldsymbol{v}_{ij} = Q_n' \boldsymbol{y}_{ij},$$

$i = 1, 2, \cdots, a$, and $j = 1, 2, \cdots, b$. Let $\boldsymbol{v}_{ij} = (v_{ij1}, \boldsymbol{v}_{ij0}')'$, so that $\boldsymbol{v}_{ij0}$ is an $(n-1) \times 1$ vector. It follows that $\boldsymbol{v}_{ij0}$ has the mean vector and covariance matrix given by

$$E(\boldsymbol{v}_{ij0}) = \mathbf{0}, \; Var(\boldsymbol{v}_{ij0}) = \sigma^2 \boldsymbol{I}_{n-1}.$$

Here we have used the facts that $E(\boldsymbol{y}_{ij}) = \mu \mathbf{1}_n$ and Q_n is an $n \times n$ orthogonal matrix with the first column given by $\frac{1}{\sqrt{n}} \mathbf{1}_n$. Denote by $\boldsymbol{v}_0$ the $ab(n-1) \times 1$ vector consisting of the $\boldsymbol{v}_{ij0}' s$, $\forall i = 1, \cdots, a$ and $j = 1, \cdots, b$. Then

$$E(\boldsymbol{v}_0) = \mathbf{0}, \; Var(\boldsymbol{v}_0) = \sigma^2 \boldsymbol{I}_{ab(n-1)}. \tag{23}$$

It should be clear that the components of $\boldsymbol{v}_0$ are the $ab(n-1)$ *error contrasts*; we recall that under the two-way model with interaction and balanced data, the error sum of squares has df $= ab(n-1)$.

Now let's consider v_{ij1}, the first element of the vector $\boldsymbol{v}_{ij}' s$. The model for the y_{ijm}'s imply the following model for v_{ij1}:

$$v_{ij1} = \sqrt{n}\mu + \sqrt{n}(\alpha_i + \beta_j + \gamma_{ij}) + \sqrt{n}\; \bar{\epsilon}_{ij},$$

where $\bar{\epsilon}_{ij} = \frac{1}{n} \sum_{m=1}^{n} \epsilon_{ijm}$. Denote by $\boldsymbol{v}_1$ the vector consisting of the v_{ij1}'s, $j = 1, \cdots, b$ and $i = 1, \cdots, a$. Note that $\boldsymbol{v}_1$ is an $ab \times 1$ vector. We then we have the model

$$\boldsymbol{v}_1 = \sqrt{n}[\mu \mathbf{1}_{ab} + (\boldsymbol{I}_a \otimes \mathbf{1}_b)\boldsymbol{\alpha} + (\mathbf{1}_a \otimes \boldsymbol{I}_b)\boldsymbol{\beta} + \boldsymbol{\gamma} + \bar{\boldsymbol{\epsilon}}],$$

where $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ and $\bar{\boldsymbol{\epsilon}}$ are vectors consisting of the α_i, β_j, γ_{ij} and $\bar{\epsilon}_{ij}$, respectively. Thus

$$E(\boldsymbol{v}_1) = \sqrt{n}\mu \mathbf{1}_{ab}$$

$$V(\boldsymbol{v}_1) = n\sigma_\alpha^2(\boldsymbol{I}_a \otimes \mathbf{1}_b\mathbf{1}_b') + n\sigma_\beta^2(\mathbf{1}_a\mathbf{1}_a' \otimes \boldsymbol{I}_b) + n\sigma_\gamma^2\boldsymbol{I}_{ab} + \sigma^2\boldsymbol{I}_{ab}.$$

We now make further transformations of the vector $\boldsymbol{v}_1$. Let $Q_a = \left(\frac{1}{\sqrt{a}}1_a, Q_a^*\right)$ and $Q_b = \left(\frac{1}{\sqrt{b}}1_b, Q_b^*\right)$ be defined similar to Q_n, but with dimensions $a \times a$ and $b \times b$, respectively. Define

$$w_0 = \left(\frac{1}{\sqrt{a}}1_a' \otimes \frac{1}{\sqrt{b}}1_b'\right)\boldsymbol{v}_1, \ \boldsymbol{w}_1 = \left(Q_a^{*'} \otimes \frac{1}{\sqrt{b}}1_b'\right)\boldsymbol{v}_1,$$
$$\boldsymbol{w}_2 = \left(\frac{1}{\sqrt{a}}1_a' \otimes Q_b^{*'}\right)\boldsymbol{v}_1, \quad \text{and} \quad \boldsymbol{w}_3 = \left(Q_a^{*'} \otimes Q_b^{*'}\right)\boldsymbol{v}_1. \tag{24}$$

We note that w_0 is a scalar, and the vectors $\boldsymbol{w}_1$, $\boldsymbol{w}_2$ and $\boldsymbol{w}_3$ are of dimensions $(a-1)\times 1$, $(b-1)\times 1$, and $(a-1)(b-1)\times 1$, respectively. It is readily verified that w_0, $\boldsymbol{w}_1$, $\boldsymbol{w}_2$ and $\boldsymbol{w}_3$ are all uncorrelated. We shall denote the elements of the vectors $\boldsymbol{w}_1$, $\boldsymbol{w}_2$ and $\boldsymbol{w}_3$ as $w_{1i'}$, $w_{2j'}$ and $w_{3i'j'}$, respectively, for $i' = 1, 2, \cdots, a-1$, and $j' = 1, 2, \cdots, b-1$, and we have the following means and variances:

$$E(w_0) = \sqrt{abn}\mu, \ Var(w_0) = \sigma_0^2 = bn\sigma_\alpha^2 + an\sigma_\beta^2 + n\sigma_\gamma^2 + \sigma^2,$$
$$E(w_{1i'}) = 0, \ Var(w_{1i'}) = \sigma_1^2 = bn\sigma_\alpha^2 + n\sigma_\gamma^2 + \sigma^2, \ i' = 1, 2, ..., a-1,$$
$$E(w_{2j'}) = 0, \ Var(w_{2j'}) = \sigma_2^2 = an\sigma_\beta^2 + n\sigma_\gamma^2 + \sigma^2, \ j' = 1, 2, ..., b-1,$$
$$E(w_{3i'j'}) = 0, \ Var(w_{3i'j'}) = \sigma_3^2 = n\sigma_\gamma^2 + \sigma^2, \ i' = 1, 2, ..., a-1, \quad j' = 1, 2, \cdots, b-1.$$

Let $v_{0m'}$ denote the m'th element of the vector $\boldsymbol{v}_0$ defined in (23), so that $E(v_{0m'}) = 0$ and $Var(v_{0m'}) = \sigma^2$, $m' = 1, 2, \cdots, ab(n-1)$. Furthermore, the $v_{0m'}$s are uncorrelated. If the normality assumption holds for all the random effects and the error terms in the two-way random model, then the following four sets of random variables follow independent normal distributions with means all equal to zero, and variances as specified above: (i) $v_{0m'}$, $m' = 1, 2, \cdots, ab(n-1)$, (ii) $w_{1i'}$, $i' = 1, 2, \cdots, a-1$, (iii) $w_{2j'}$, $j' = 1, 2, \cdots, b-1$, and (iv) $w_{3i'j'}$, $i' = 1, 2, \cdots, a-1$, $j' = 1, 2, \cdots, b-1$. In other words, under the assumption of normality, the random variables given in (i), (ii), (iii) and (iv) can be treated as samples of sizes $ab(n-1)$, $a-1$, $b-1$ and $(a-1)(b-1)$ from four independent normal distributions with zero means. Independent smooth alternatives to normality can now be specified for each of the four sets (i), (ii), (iii) and (iv), and score tests can be derived for simultaneously testing normality of the error term and the normality of the random effects.

5.2. Testing normality in a two-way mixed model with balanced data

Now consider the two-way mixed effect model with interaction and balanced data:

$$y_{ijm} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijm},$$

where the α_i, $i = 1, 2, \cdots, a$, are assumed to be the fixed effects, and the rest of the effects are random effects. Here we make the usual assumption: $\sum_{i=1}^a \alpha_i = 0$. The standard normality assumptions that are imposed on the random effects and the error terms are the same as those given in the previous section: $\beta_j \sim N(0, \sigma_\beta^2)$, $\gamma_{ij} \sim N(0, \sigma_\gamma^2)$, and $\epsilon_{ijm} \sim N(0, \sigma^2)$, $i = 1, \cdots, a$, $j = 1, \cdots, b$ and $m = 1, \cdots, n$.

Starting with the transformation based on the Helmert matrix Q_n, we can arrive at the $ab(n-1)$ error contrast vector $\boldsymbol{v}_0$ given in the previous section, having zero mean and the variance-covariance matrix $\sigma^2 I_{ab(n-1)}$; see (23). Let's now consider the quantities w_0, $\boldsymbol{w}_1$, $\boldsymbol{w}_2$ and $\boldsymbol{w}_3$, defind in (24), and denote the elements of $\boldsymbol{w}_2$ and $\boldsymbol{w}_3$ as $w_{2j'}$ and $w_{3i'j'}$, respectively, for $i' = 1, 2, \cdots, a-1$, and $j' = 1, 2, \cdots, b-1$. It can once again be verified that w_0 , $\boldsymbol{w}_1$, $\boldsymbol{w}_2$, and $\boldsymbol{w}_3$ are all uncorrelated, and we have the following means and variances:

$$
\begin{aligned}
&E(w_0) = \sqrt{abn}\mu,\ Var(w_0) = an\sigma_\beta^2 + n\sigma_\gamma^2 + \sigma^2,\\
&E(\boldsymbol{w}_1) = \sqrt{b}Q_a^{*'}\boldsymbol{\alpha},\ Var(\boldsymbol{w}_1) = (n\sigma_\gamma^2 + \sigma^2)\boldsymbol{I}_{a-1},\\
&E(w_{2j'}) = 0,\ Var(w_{2j'}) = an\sigma_\beta^2 + n\sigma_\gamma^2 + \sigma^2,\ j' = 1, 2, \cdots, b-1,\\
&E(w_{3i'j'}) = 0,\ Var(w_{3i'j'}) = n\sigma_\gamma^2 + \sigma^2,\ i' = 1, 2, \cdots, a-1, \quad j' = 1, 2, \cdots, b-1.
\end{aligned}
$$

The vectors $\boldsymbol{v}_0$, $\boldsymbol{w}_2$ and $\boldsymbol{w}_3$ are also uncorrelated, have mean zeros, and covariance matrices $\sigma^2 I_{ab(n-1)}$, $(an\sigma_\beta^2 + n\sigma_\gamma^2 + \sigma^2)\boldsymbol{I}_{b-1}$ and $(n\sigma_\gamma^2 + \sigma^2)\boldsymbol{I}_{(a-1)(b-1)}$, respectively. Three independent smooth alternatives to normality can now be defined, as noted in the previous section, and smooth tests can be derived. Note that the scalar quantity w_0, and the vector $\boldsymbol{w}_1$ have means $\sqrt{abn}\mu$ and $\sqrt{b}Q_a^{*'}\boldsymbol{\alpha}$, respctively, which are unknown nuisance parameters to be estimated. Thus these components will not contribute to the test for normality. In this section and in the previous section, we have not brought up the issue of testing homoscedasticity.

We believe that the approach outlined in this section and the previous section can be adopted to any mixed or random effects model with balanced data. However, the same approach will not go through when we have unbalanced data. Let's briefly indicate why this is so. Consider the case of the random effects model. A key step in the derivations is the transformation in (24) leading to the uncorrelated quantities w_0, $\boldsymbol{w}_1$, $\boldsymbol{w}_2$ and $\boldsymbol{w}_3$. It is not difficult to note that such a transformation leading to uncorrelated quantities is not possible when we have unbalanced data. In short, when we have a linear model with a structured covariance matrix, which is the case for any mixed or random effects model, it is not clear how we can define a smooth alternative by taking the structure into account.

References

Bogdan, M. (1996). Data driven smooth tests for bivariate normality, preprint IM PWr 42/96, Technical University of Wroclaw.

Bogdan, M. (1999). Data driven smooth tests for bivariate normality. *Journal of Multivariate Analysis*, **68**, 26-53.

Chang, C.-H., Pal, N. and Lin, J.-J. (2017). A revisit to test the equality of variances of several populations. *Communications in Statistics - Simulation and Computation*, **46**, 6360-6384.

Inglot, T., Kallenberg, W. and Ledwina, T. (1994). On selection rules, with an application to goodness-of-fit for composite hypotheses. (Memorandum, Faculty of Mathematical Sciences; No. 1242). University of Twente, Faculty of Mathematical Sciences.

Inglot, T., Kallenberg, W. and Ledwina, T. (1997). Data driven smooth tests for composite hypotheses. *The Annals of Statistics*, **25**, 1222-1250.

Inglot, T. and Ledwina, T. (1996). Asymptotic optimality of data-driven Neymans tests for uniformity. *The Annals of Statistics*, **24**, 1982-2019.

Janic, A. and Ledwina, T. (2009). Data-driven smooth tests for a location-scale family revisited, *Journal of Statistical Theory and Practice*, **3**, 645-664.

Kallenberg, W. and Ledwina, T. (1995). Consistency and Monte Carlo simulation of a data driven version of smooth goodness-of-fit tests. *The Annals of Statistics*, **23**, 1594-1608.

Kallenberg, W. and Ledwina, T. (1997a). Data-driven smooth tests for composite hypotheses: Comparison of powers. *Journal of Statistical Computation and Simulation*, **59**, 101-121.

Kallenberg, W. and Ledwina, T. (1997b). Data-driven smooth tests when the hypothesis is composite. *Journal of the American Statistical Association*, **92**, 1094-1104.

Kallenberg, W., Ledwina, T. and Rafajlowicz, E. (1997). Testing bivariate independence and normality. *Sankhy*ã*, Series A*, **59**, 42-59.

Krishnamoorthy, K. and Mathew, T. (2009). *Statistical Tolerance Regions: Theory, Applications and Computation*. John Wiley, New York.

Ledwina, T. (1994). Data-driven version of Neyman's smooth test of fit. *Journal of the American Statistical Association*, **89**, 1000-1005.

Montgomery, D. C. (2020). *Design and Analysis of Experiments*, 10th Edition, John Wiley, New York.

Neyman, J. (1937) Smooth test for goodness of fit. *Skandinaviska Aktuarietidskrift*, **20**, 149-199.

Peña, E. A. and Slate E. H. (2006). Global validation of linear model assumptions. *Journal of the American Statistical Association*, **101**, 341-354.

Rayner, J. C. W., Thas O. and Best, D. J. (2009). *Smooth Tests of Goodness of Fit : Using R*. John Wiley, New York.

Thomas, J. D. and Hultquist, R. A. (1978). Interval estimation for the unbalanced case of the one-way random effects model. *The Annals of Statistics*, **6**, 582-587.

Xu, W., Li, Y. and Song, D. (2013). Testing normality in mixed models using a transformation method. *Statistical Papers*, **54**, 71-84.

Yang, Y. (2016). *The Simultaneous Assessment of Normality and Homoscedasticity in Some Linear Models*. Doctoral dissertation submitted to the University of Maryland Baltimore County.

Yang, Y. and Mathew, T. (2018). The simultaneous assessment of normality and homoscedasticity in linear fixed effects models. *Journal of Statistical Theory and Practice*, **12**, 66-81.

APPENDIX

The Score Vector and its Variance-Covariance Matrix

Here we shall give only the expressions for the score vector and its variance-covariance matrix (without providing their derivations) under both the Legendre polynomial based and Hermite polynomial based alternatives. In order to derive these, it is necessary to use expressions for the derivatives of $\ln\{C(\boldsymbol{\theta}, \mu, \boldsymbol{\sigma}, \sigma_\alpha)\}$. General results on such derivatives are given in Rayner, Thas and Best (2009, Section 6.1).

A. The score vector, variances and covariances under the Legendre polynomial based alternative

Recall that the elements of the score vector and those of the variance-covariance matrix are to be evaluated under the null hypothesis. For the log-likelihood function given in (10), the components of the score vector (under the null hypothesis) are as follows:

$$\frac{\partial \ln L}{\partial \theta_{1r}} = \sum_{l=1}^{a} b_r(u_{l1}), \ \frac{\partial \ln L}{\partial \theta_{2s}} = \sum_{l=1}^{a}\sum_{j=2}^{n_l} b_s(u_{lj}), \ r = 1, 2, \cdots, k_1, \ s = 1, 2, \cdots, k_2$$

$$\frac{\partial \ln L}{\partial \eta_c} = \sum_{l=1}^{a}(z_{l1}^2 - 1)\frac{\sigma q_{lc}}{\sigma^2 + n_l\sigma_\alpha^2} + \sum_{l=1}^{a}\sum_{j=2}^{n_l}(z_{lj}^2 - 1)\frac{q_{lc}}{\sigma}, \ c = 1, 2, \cdots, a,$$

$$\frac{\partial \ln L}{\partial \sigma_\alpha} = \sum_{l=1}^{a}(z_{l1}^2 - 1)\frac{n_l\sigma_\alpha}{\sigma^2 + n_l\sigma_\alpha^2}, \ \frac{\partial \ln L}{\partial \mu} = \sum_{l=1}^{a}\frac{\sqrt{n_l}z_{l1}}{\sqrt{\sigma^2 + n_l\sigma_\alpha^2}}$$

where σ^2 is the common variance under the null hypothesis, and q_{lc}'s are the elements of the $a \times a$ Helmert matrix Q defined in Section 2, and $\boldsymbol{\eta} = Q'\boldsymbol{\sigma}$; see (9).

The expressions for the variances and covariances among the components of the score vector involve certain constants c_i and e_i, $i = 1, 2, \cdots$. We shall first give these before giving the variance and covariance terms. Let $f(z)$ denote the density of a standard normal random variable Z. The required constants c_i and e_i are given by

$$\begin{aligned} c_i &= \text{Cov}[b_i(\Phi(Z)), Z] = \int_{-\infty}^{\infty} b_i(\Phi(z))zf(z)dz, \\ e_i &= \text{Cov}[b_i(\Phi(Z)), Z^2] = \int_{-\infty}^{\infty} b_i(\Phi(z))z^2f(z)dz, \end{aligned}$$

$i = 1, 2, 3, \cdots$. By using the expressions for the Legendre polynomials, it can be verified that $c_i = 0$ for i even, and $e_i = 0$ for i odd. When c_i's and e_i's are non-zero, they can be computed numerically. A few such values are given below; see Bogdan (1996, 1999).

$c_1 = 0.977205023801135, \ c_3 = 0.1830082402700861, \ c_5 = 0.0816989764273946,$
$c_7 = 0.04772936798473241, \ c_9 = 0.031880431223894;$
$e_2 = 1.232808888123174, \ e_4 = 0.5211245854593028, \ e_6 = 0.3045144697203598,$
$e_8 = 0.2055889833015625, \ e_{10} = 0.150770690085310.$

The variances and covariances among the components of the score vector are as follows (under the null hypothesis):

$$\text{Var}\left(\frac{\partial \ln L}{\partial \theta_{1r}}\right) = a, \quad \text{and Cov}\left(\frac{\partial \ln L}{\partial \theta_{1r}}, \frac{\partial \ln L}{\partial \theta_{1r'}}\right) = 0,\ r \neq r';\ r, r' = 1, 2, \cdots, k_1,$$

$$\text{Var}\left(\frac{\partial \ln L}{\partial \theta_{2s}}\right) = N - a,\ \text{where } N = \sum_{l=1}^{a} n_l, \quad \text{and Cov}\left(\frac{\partial \ln L}{\partial \theta_{2s}}, \frac{\partial \ln L}{\partial \theta_{2s'}}\right) = 0,$$

$s \neq s';\ s, s' = 1, 2, \cdots, k_2,$

$$\text{Cov}\left(\frac{\partial \ln L}{\partial \theta_{1r}}, \frac{\partial \ln L}{\partial \theta_{2s}}\right) = 0,\ r = 1, 2, \cdots, k_1,\ s = 1, 2, \cdots, k_2,$$

$$\text{Var}\left(\frac{\partial \ln L}{\partial \eta_c}\right) = 2\sum_{l=1}^{a}\left(\frac{\sigma^2}{(\sigma^2 + n_l\sigma_\alpha^2)^2} + \frac{(n_l - 1)}{\sigma^2}\right) q_{lc}^2,\ c = 1, 2, \cdots, a,$$

$$\text{Cov}\left(\frac{\partial \ln L}{\partial \eta_c}, \frac{\partial \ln L}{\partial \eta_{c'}}\right) = 2\sum_{l=1}^{a}\left(\frac{\sigma^2}{(\sigma^2 + n_l\sigma_\alpha^2)^2} + \frac{(n_l - 1)}{\sigma^2}\right) q_{lc}q_{lc'},\ c \neq c';\ c, c' = 1, 2, \cdots, a,$$

$$\text{Cov}\left(\frac{\partial \ln L}{\partial \theta_{1r}}, \frac{\partial \ln L}{\partial \eta_c}\right) = e_r \sum_{l=1}^{a} \frac{\sigma}{\sigma^2 + n_l\sigma_\alpha^2} q_{lc} \text{ for } r \text{ even, and 0 for } r \text{ odd},\ c = 1, \cdots, a,$$

$$\text{Cov}\left(\frac{\partial \ln L}{\partial \theta_{2s}}, \frac{\partial \ln L}{\partial \eta_c}\right) = (e_r/\sigma)\sum_{l=1}^{a}(n_l - 1)q_{lc} \text{ for } s \text{ even, and 0 for } s \text{ odd},\ c = 1, \cdots, a,$$

$$\text{Var}\left(\frac{\partial \ln L}{\partial \sigma_\alpha}\right) = 2\sum_{l=1}^{a} \frac{n_l^2\sigma_\alpha^2}{(\sigma^2 + n_l\sigma_\alpha^2)^2},$$

$$\text{Cov}\left(\frac{\partial \ln L}{\partial \theta_{1r}}, \frac{\partial \ln L}{\partial \sigma_\alpha}\right) = e_r \sum_{l=1}^{a} \frac{n_l\sigma_\alpha}{\sigma^2 + n_l\sigma_\alpha^2} \text{ for } r \text{ even, and 0 for } r \text{ odd},$$

$$\text{Cov}\left(\frac{\partial \ln L}{\partial \theta_{2s}}, \frac{\partial \ln L}{\partial \sigma_\alpha}\right) = 0,\ s = 1, \cdots, k_2,$$

$$\text{Cov}\left(\frac{\partial \ln L}{\partial \eta_c}, \frac{\partial \ln L}{\partial \sigma_\alpha}\right) = 2\sum_{l=1}^{a} \frac{\sigma\sigma_\alpha}{(\sigma^2 + n_l\sigma_\alpha^2)^2} n_l q_{lc},\ c = 1, \cdots, a,$$

$$\text{Var}\left(\frac{\partial \ln L}{\partial \mu}\right) = \sum_{l=1}^{a} \frac{n_l}{\sigma^2 + n_l\sigma_\alpha^2},$$

$$\text{Cov}\left(\frac{\partial \ln L}{\partial \theta_{1r}}, \frac{\partial \ln L}{\partial \mu}\right) = c_r \sum_{l=1}^{a} \frac{\sqrt{n_l}}{\sqrt{\sigma^2 + n_l\sigma_\alpha^2}} \text{ for } r \text{ odd, and 0 for } r \text{ even},$$

$$\text{Cov}\left(\frac{\partial \ln L}{\partial \theta_{2s}}, \frac{\partial \ln L}{\partial \mu}\right) = 0,\ s = 1, \cdots, k_2,$$

$$\text{Cov}\left(\frac{\partial \ln L}{\partial \eta_c}, \frac{\partial logL}{\partial \mu}\right) = 0,\ c = 1, \cdots, a, \quad \text{Cov}\left(\frac{\partial \ln L}{\partial \sigma_\alpha}, \frac{\partial \ln L}{\partial \mu}\right) = 0.$$

A.1. The case of balanced data

For balanced data, the n_l's have to be replaced with their common value n in all the expressions. As a result, some of the covariances will become 0, these are:

$$\text{Cov}\left(\frac{\partial \ln L}{\partial \eta_c}, \frac{\partial \ln L}{\partial \eta_{c'}}\right) = 0, \quad c, c' = 1, 2, \cdots, a$$
$$\text{Cov}\left(\frac{\partial \ln L}{\partial \theta_{1r}}, \frac{\partial \ln L}{\partial \eta_c}\right) = 0, \quad r = 1, 2, \cdots, k_1, \ c = 2, ..., a$$
$$\text{Cov}\left(\frac{\partial \ln L}{\partial \theta_{2s}}, \frac{\partial \ln L}{\partial \eta_c}\right) = 0, \quad s = 1, 2, \cdots, k_2, \ c = 2, ..., a$$
$$\text{Cov}\left(\frac{\partial \ln L}{\partial \eta_c}, \frac{\partial logL}{\partial \sigma_\alpha}\right) = 0, \quad c = 2, \cdots, a$$

B. The score vector, variances and covariances under the Hermite polynomial based alternative

The components of the score vector (under the null hypothesis) can be derived similar to those obtained under the Legendre polynomial case. The scores with respect to θ_{1r} and θ_{2s} are given below, and those with respect η_c, σ_α, and μ are not given since they are the same as in the Legendre case. The scores corresponding to θ_{1r} and θ_{2s} can be shown to be equal to zero for $r = 1$, 2 and for $s = 1$, 2.

$$\frac{\partial \ln L}{\partial \theta_{1r}} = \sum_{l=1}^{a} h_r(z_{l1}), \ \frac{\partial \ln L}{\partial \theta_{2s}} = \sum_{l=1}^{a}\sum_{j=2}^{n_l} h_s(z_{lj}), \ r = 3, 4, \cdots, k_1, \ s = 3, 4, \cdots, k_2.$$

Several of the variance and covariance terms are the same as those for the Legendre polynomial case. The terms that are different from the Legendre case are given below, and are in fact zeros.

$$\text{Cov}\left(\frac{\partial \ln L}{\partial \theta_{1r}}, \frac{\partial \ln L}{\partial \eta_c}\right) = 0, \quad c = 2, \cdots ., a$$
$$\text{Cov}\left(\frac{\partial \ln L}{\partial \theta_{2r}}, \frac{\partial \ln L}{\partial \eta_c}\right) = 0$$
$$\text{Cov}\left(\frac{\partial \ln L}{\partial \theta_{2s}}, \frac{\partial \ln L}{\partial \eta_c}\right) = 0, \quad c = 2, \cdots, a.$$