Estimation in additive models and ANOVA-like applications

Patrícia Antunes, Sandra S. Ferreira, Dário Ferreira, Célia Nunes & João Tiago Mexia

Published online: 04 Feb 2020.

Submit your article to this journal ⤤

View related articles ⤤

View Crossmark data ⤤

Taylor & Francis
Taylor & Francis Group

Check for updates

Estimation in additive models and ANOVA-like applications

Patrícia Antunes[a], Sandra S. Ferreira [a,b], Dário Ferreira [a,b], Célia Nunes [a,b] and João Tiago Mexia[c]

[a]Center of Mathematics, University of Beira Interior, Covilhã, Portugal; [b]Department of Mathematics and Center of Mathematics and Applications, University of Beira Interior, Covilhã, Portugal; [c]Center of Mathematics and Its Applications, Faculty of Sciences and Technology, New University of Lisbon, Lisbon, Portugal

ABSTRACT
A well-known property of cumulant generating function is used to estimate the first four order cumulants, using least-squares estimators. In the case of additive models, empirical best linear unbiased predictors are also obtained. Pairs of independent and identically distributed models associated with the treatments of a base design are used to obtain unbiased estimators for the fourth-order cumulants. An application to real data is presented, showing the good behaviour of the least-squares estimators and the great flexibility of our approach.

1. Introduction

According to Withers *et al.* [21], there has been much interest in deriving expressions for moments and cumulants using available computer technology. For example, Zheng [22] provides an implementation of two methods for expressing moments in terms of cumulants and vice versa, using the Mathematica software. In [17], the same computer package is used to derive relations between moments, cumulants, and quasi-moments. In [1], it is shown how procedures for computing moments and cumulants may themselves be derived from a few elementary identities. An efficient method for symbolic computation of moments and cumulants of sampling distributions is presented in [8]. For most excellent accounts of the literature, we refer the readers to [2,14].

While formulas involving the first, second and third central moments are easy to apply, mathematical complications arise with the fourth central moment. In this paper, we use a well-known property of cumulant generating function to obtain unbiased estimators for the first four order moments in additive models. These models are given by

$$Y = X_0 \beta_0 + XL, \tag{1}$$

where X_0 and X are design matrices, β_0 is fixed, $X = [X_1, \ldots, X_m]$, $L = [L_1, \ldots, L_m]$, where the $L_1, \ldots, L_m$ are independent, with $c_1, \ldots, c_m$ components, with variances

---

CONTACT Sandra S. Ferreira ✉ sandraf@ubi.pt

$\sigma_1^2, \ldots, \sigma_m^2$ (and third and fourth central moments, $\upsilon_1, \ldots, \upsilon_m$ and $Z_1, \ldots, Z_m$, respectively). These models are easy to implement, not requiring structural conditions to be fulfilled, see [10]. Such conditions, such as blocks with the same size, and orthogonal block structure, have played an important part in the study of models (see for instance [4,5]). The fact that these conditions are no longer required makes additive models much more comprehensive.

As stated above, we will show how to use cumulant generating function (CGF) to obtain least-squares estimators for these moments. In order to obtain unbiased estimators for the $Z_1, \ldots, Z_m$, we will consider a pair of models with independent and identically distributed, i.i.d., observation vectors $\boldsymbol{Y}(1)$ and $\boldsymbol{Y}(2)$. Moreover, we will also show how to estimate $\boldsymbol{\beta}_0$ and estimable vectors $\boldsymbol{\eta} = \boldsymbol{G}\boldsymbol{\beta}_0$ and how to obtain empirical best linear unbiased predictors, EBLUP, for $\boldsymbol{L}$. So, we will show how to carry out ANOVA-like inference for the moments and estimable vectors associated with the treatments in the base design.

The rest of the paper is organized as follows. In Section 2, we will recall certain results on CGF and cumulants, which will be useful to show how to obtain estimators for the variances, third and fourth central moments. In Section 3, we will consider additive models and show how to use cumulants to carry out the estimation of the parameters in such models. Next, in Section 4, we present further results on estimation. In Section 5, we shall see that when we have a pair of i.i.d. models, for each treatment of a base design, we will be able to assess the influence of the factors, in that base design, on the various second-, third- and fourth-order central moments, as well as on estimable functions. In Section 6, we present an application to real data illustrating our approach, with quite good results. Moreover, the considered models have only six observations each, being therefore very small. This plays in favour of the presented approach because, as it may be seen, the obtained estimates are very close to the true values. The paper ends with some conclusions in Section 7.

2. Cumulants and generating functions

In statistics, the most commonly used generating functions are the moment generating function and the cumulant generating function. An excellent source about relations between moments and cumulants is the book [9]. We begin with a definition of moment generating function:

Definition 2.1 (Moment generating function): Given the random vector $\boldsymbol{W}$, with components $W_1, \ldots, W_m$, its moment generating function (MGF) will be

$$\varphi(\boldsymbol{d} \mid \boldsymbol{W}) = E(e^{\boldsymbol{d}^\top \boldsymbol{W}}),$$

if expected value, E, exists and is finite and defined for all real vector $\boldsymbol{d}$. It must be mentioned that not all random vectors possess an MGF (see [15]). Follows

Definition 2.2 (Cumulant generating function): The cumulant generating function (CGF) is

$$\psi(\boldsymbol{d} \mid \boldsymbol{W}) = \log(\varphi(\boldsymbol{d} \mid \boldsymbol{W})).$$

There are some advantages to using cumulants (see, e.g. [16]). The rth cumulant is a polynomial function of the first r moments. Moreover, it is shift-invariant if $r > 1$, which

means that for a random variable X and constants a, the rth cumulant of the probability distribution of $X + a$ is the same as the rth cumulant of the probability distribution of X, i.e. $C_r(X + a) = C_r(X)$, where $C_r(X)$ refers to rth cumulant. Because of shift invariance, one may take the rth cumulant to be a polynomial in the first r central moments.

If the components $W_1, \ldots, W_m$ of $\boldsymbol{W}$ are independent, we have (see [7]) the following property of CGF

$$\psi(\boldsymbol{d} \mid \boldsymbol{a}^\top \boldsymbol{W}) = \sum_{l=1}^{m} \psi(a_l d_l \mid W_l), \tag{2}$$

with $\boldsymbol{a} = (a_1, \ldots, a_m)^\top$ and

$$\psi^{\langle r \rangle}(\boldsymbol{d} \mid \boldsymbol{a}^\top \boldsymbol{W}) = \sum_{l=1}^{m} a_l^r \psi^{\langle r \rangle}(d_l \mid W_l), \tag{3}$$

with $\langle r \rangle$ indicating the rth derivative in order to $d_1, \ldots, d_m$. Since the rth cumulant equals the rth derivative of the CGF at the origin we represent them by $\mathcal{O}_r(\boldsymbol{a}^\top \boldsymbol{W})$ and $\mathcal{O}_r(W_l)$, $l = 1, \ldots, m, r = 1, 2, \ldots$, respectively. So we have the relation (see [13,19])

$$\mathcal{O}_r(\boldsymbol{a}^\top \boldsymbol{W}) = \sum_{l=1}^{m} a_l^r \mathcal{O}_r(W_l). \tag{4}$$

It is useful to recall that for any random variable U with mean value $\mu(u)$, [variance, third and fourth central moments], $[\sigma^2(u), \mu_3(u)$ and $\mu_4(u)]$, according to [3,9], we have

$$\begin{aligned}
\mathcal{O}_1(U) &= \mu(u), \\
\mathcal{O}_2(U) &= \sigma^2(u), \\
\mathcal{O}_3(U) &= \mu_3(u), \\
\mathcal{O}_4(U) &= \mu_4(u) - 3(\sigma^2(u))^2.
\end{aligned} \tag{5}$$

While estimating the $\mathcal{O}_r(U)$, $r = 1, 2, 3$, presents no problem, the difficulties arise when we want to estimate $\mathcal{O}_4(U)$ due to the fourth and higher order cumulants not being equal to the central moments. We overcome this difficulties considering a pair of i.i.d. models, each providing an unbiased estimator for the $\sigma^2(u)$.

3. Adjustments

Let us consider a matrix $\boldsymbol{A}^\top = [\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_{\dot{n}}]$, where $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_{\dot{n}}$ constitute an orthonormal basis for the orthogonal complement $\Omega^\perp$, of the space $\Omega = R(\boldsymbol{X}_0)$, the range space of $\boldsymbol{X}_0$ spanned by the model's mean vector $\boldsymbol{\mu} = \boldsymbol{X}_0 \boldsymbol{\beta}_0$. If the dimension of Ω is $k = rank(\boldsymbol{X}_0)$, the dimension of $\Omega^\perp$ is $\dot{n} = n - k$. Then, with

$$\boldsymbol{\alpha}_h^\top \boldsymbol{X}_l = (a_{h,l,1}, \ldots, a_{h,l,c_l}), \quad l = 1, \ldots, m, \quad h = 1, \ldots, \dot{n}, \tag{6}$$

we have

$$\dot{Y}_h = \boldsymbol{\alpha}_h^\top \boldsymbol{Y} = \sum_{l=1}^{m} (\boldsymbol{\alpha}_h^\top \boldsymbol{X}_l) \boldsymbol{L}_l = \sum_{l=1}^{m} \sum_{v=1}^{c_l} a_{h,l,v} L_{l,v}, \tag{7}$$

with

$$E(\dot{Y}_h) = 0, \quad h = 1, \ldots, \dot{n}. \tag{8}$$

Thus,

$$\mathcal{O}_r(\dot{Y}_h) = \sum_{l=1}^{m} \sum_{v=1}^{c_l} a_{h,l,v}^r \mathcal{O}_r(L_{l,v}) = \sum_{l=1}^{m} b_{h,l}(r) \chi_l(r), \tag{9}$$

where

$$b_{h,l}(r) = \sum_{v=1}^{c_l} a_{h,l,v}^r, \tag{10}$$

and the rth order cumulants of the $L_{l,v}$ are

$$\chi_l(r) = \mathcal{O}_r(L_{l,v}), \quad v = 1, \ldots, c_l, \quad l = 1, \ldots, m, \ r = 2, 3, 4. \tag{11}$$

Expression (9) may be used to obtain the least square estimator (LSE) for the $\chi_l(r)$, $l = 1, \ldots, m, r = 2, 3, 4$, since we have the unbiased estimators

$$\tilde{\mathcal{O}}_2(\dot{Y}_h) = \dot{Y}_h^2,$$
$$\tilde{\mathcal{O}}_3(\dot{Y}_h) = \dot{Y}_h^3, \tag{12}$$
$$\tilde{\mathcal{O}}_4(\dot{Y}_h) = \frac{\dot{Y}_h^4(1) + \dot{Y}_h^4(2)}{2} - 3\dot{Y}_h(1)^2 \dot{Y}_h(2)^2,$$

$h = 1, \ldots, \dot{n}$ (see [3,19,22]). The estimation of the $\tilde{\mathcal{O}}_2(\dot{Y}_h)$ and $\tilde{\mathcal{O}}_3(\dot{Y}_h)$ requires only one observation vector, unlike the estimation of $\tilde{\mathcal{O}}_4(\dot{Y}_h)$ that requires a pair of i.i.d. observation vectors, in order to have an unbiased estimator for $\mathcal{O}_4(\dot{Y}_h)$.

Taking matrix

$$\boldsymbol{B}(r) = \left[b_{h,l}(r) \right], \quad r = 2, 3, 4, \tag{13}$$

we get the LSE for $\chi(r)$,

$$\tilde{\chi}(r) = (\boldsymbol{B}(r)^\top \boldsymbol{B}(r))^+ \boldsymbol{B}(r)^\top \tilde{\mathcal{O}}_r(\dot{\boldsymbol{Y}}), \quad r = 2, 3, 4, \tag{14}$$

where $+$ indicates the Moore–Penrose inverse of a matrix and

$$\dot{\boldsymbol{Y}}^\top = [\dot{\boldsymbol{Y}}_1, \ldots, \dot{\boldsymbol{Y}}_{\dot{n}}]. \tag{15}$$

In what follows, we consider estimable vectors for which we will use generalized least-squares estimators (GLSE) according to [20]. The estimator of variance components is given by $\tilde{\chi}^\top(2) = [\tilde{\chi}_1(2), \ldots, \tilde{\chi}_m(2)]$. Through LSE, the estimator for $\boldsymbol{\beta}_0$ is given by

$$\tilde{\boldsymbol{\beta}}_0 = (\boldsymbol{X}_0^\top \not\!\!\boldsymbol{\Sigma}(\tilde{\chi}(2))\boldsymbol{X}_0)^+ \boldsymbol{X}_0^\top \not\!\!\boldsymbol{\Sigma}(\tilde{\chi}(2))^+ \boldsymbol{Y}, \tag{16}$$

with

$$\not\!\!\boldsymbol{\Sigma}(\tilde{\chi}(2)) = \sum_{l=1}^{m} \tilde{\chi}_l(2)\boldsymbol{X}_l\boldsymbol{X}_l^\top. \tag{17}$$

With $\boldsymbol{G} = \boldsymbol{A}\boldsymbol{X}_0$ and $\tilde{\boldsymbol{\beta}}_0$ an unbiased estimator of $\boldsymbol{\beta}_0$, such as the LSE, one may obtain $\tilde{\boldsymbol{\eta}} = \boldsymbol{G}\tilde{\boldsymbol{\beta}}_0$ which will be an unbiased estimator of $\boldsymbol{\eta}$. We point out that $\boldsymbol{\beta}_0$ is itself an estimable vector.

4. Further results on estimation and EBLUP

Up to now we have assumed the model

$$Y = X_0\boldsymbol{\beta}_0 + \sum_{l=1}^{m} X_l L_l. \tag{18}$$

Let us now add a term $e = (e_1, \ldots, e_n)$ of measurement errors. These will be generated as the sum of a great number of small 'impacts' in a way that resembles the probabilistic model behind the central limit theorem. We can then assume that e is normal, with null mean vector and covariance matrix $\sigma^2 I_n$. Thus the $(e_1, \ldots, e_n)$ will be i.i.d. with

$$\mathcal{O}_1(e_i) = \mathcal{O}_3(e_i) = \mathcal{O}_4(e_i) = 0; \quad \mathcal{O}_2(e_i) = \sigma^2, \quad i = 1, \ldots, n. \tag{19}$$

Throughout this section, matrix $\overline{T}$ will denote the orthogonal projection matrix on $R[(X_0, \ldots, X_m)]$, so that, with $\overline{k} = rank(\overline{T})$, we have the unbiased estimator

$$\tilde{\sigma}^2 = \frac{Y^\top (I_n - \overline{T}) Y}{n - \overline{k}}, \tag{20}$$

for σ^2 (see [12]). To estimate the $\sigma_l^2 = \chi_l(2), l = 1, \ldots, m$, we have to replace the former estimators $\tilde{\sigma_l^2}$ by $\tilde{\sigma_l^2} - \tilde{\sigma}^2$, while the estimators of third and fourth orders are left unchanged, since $\mathcal{O}_3(e_i) = \mathcal{O}_4(e_i) = 0, i = 1, \ldots, n$.

Given a random vector H, with mean vector $\boldsymbol{\eta}$ and covariance matrix V, it may be seen that we have

$$E(\|H - \boldsymbol{\eta}\|^2) = \text{tr}(V), \tag{21}$$

where $\text{tr}(V)$ denotes the trace of matrix V. So, we have the model

$$Y = X_0\boldsymbol{\beta}_0 + \sum_{l=1}^{m} X_l L_l + e, \tag{22}$$

with the covariance matrix

$$\varSigma(Y) = \sum_{l=1}^{m} \chi_l(2) M_l + \sigma^2 I_n. \tag{23}$$

Therefore, for $\text{tr}(V)$, we have the unbiased estimator

$$\text{tr}(\tilde{V}) = \sum_{l=1}^{m} \tilde{\chi}_l(2) t_l + n\sigma^2, \tag{24}$$

where $t_l = \text{tr}(M_l), l = 1, \ldots, m$. Rewriting the model as

$$Y = X_0\boldsymbol{\beta}_0 + XL + e \tag{25}$$

and considering $\boldsymbol{\lambda}$ any known vector, we have for $\boldsymbol{\lambda}^\top L$ the EBLUP (see [6])

$$\boldsymbol{\lambda}^\top L = (Y - X_0\tilde{\boldsymbol{\beta}}_0)^\top \tilde{f}, \tag{26}$$

where

$$\tilde{f} = \left(\sum_{l=1}^{m} \tilde{\chi}_l(2)M_l + \tilde{\sigma}^2 I_n\right)^{-1} \not\!\Sigma(Y, \lambda^\top L) \tag{27}$$

and the cross covariance matrix

$$\not\!\Sigma(Y, \lambda^\top L) = \not\!\Sigma(X_0\beta_0 + XL + e; \lambda^\top L) = X \not\!\Sigma(L)\lambda, \tag{28}$$

since e and L are independent. Now, L has the blockwise diagonal covariance matrix $D(\chi_1(2)I_{c_1}, \ldots, \chi_m(2)I_{c_m})$ so that, with $\lambda = [\lambda_1^\top, \ldots, \lambda_m^\top]^\top$, we have the extended cross-covariance matrix

$$\not\!\Sigma(Y, \lambda^\top L) = XD(\chi_1(2)I_{c_1}, \ldots, \chi_m(2)I_{c_m})[\lambda_1^\top, \ldots, \lambda_m^\top]^\top = \sum_{l=1}^{m} \tilde{\chi}_l(2)X_l\lambda_l. \tag{29}$$

5. Pairs of models

Here we will consider pairs of i.i.d. models

$$Y(u) = X_0\beta_0(u) + \sum_{l=1}^{m} X_l L_l(u), \quad u = 1, 2, \tag{30}$$

where the vectors $L_l(u)$, $l = 1, \ldots, m$, $u = 1, 2$, have variance–covariance matrices $\chi_l(2)(u)I_{c_l}, l = 1, \ldots, m$, with c_l components, $l = 1, \ldots, m$ and $u = 1, 2$.

Using the previous results, one may obtain the estimator

$$\tilde{E}(u) = \begin{bmatrix} \tilde{\beta}_0(u) \\ \tilde{\chi}(2)(u) \\ \tilde{\mu}_3(u) \end{bmatrix}, \quad u = 1, 2. \tag{31}$$

Hence, we will use the pair $(Y(1), Y(2))$ to obtain

$$\tilde{\mathcal{O}}_4 = \tilde{\mu}_4 - 3\tilde{\chi}(2)(1)\tilde{\chi}(2)(2), \tag{32}$$

with

$$\tilde{\mu}_4 = \frac{\tilde{\mu}_4(1) + \tilde{\mu}_4(2)}{2}. \tag{33}$$

Furthermore, one may combine the estimators into

$$\tilde{E} = \begin{bmatrix} \frac{1}{2}(\tilde{\beta}_0(1) + \tilde{\beta}_0(2)) \\ \frac{1}{2}(\tilde{\chi}(2)(1) + \tilde{\chi}(2)(2)) \\ \frac{1}{2}(\tilde{\mu}_3(1) + \tilde{\mu}_3(2)) \\ \tilde{\mu}_4 - 3\tilde{\chi}(2)(1)\tilde{\chi}(2)(2) \end{bmatrix}$$

and use the pair

$$d_0 = \|\tilde{E}(1) - \tilde{E}(2)\|,$$

$$r_0 = \frac{\tilde{E}(1)\tilde{E}(2)}{\prod_{u=1}^{2} \|\tilde{E}(u)\|} \tag{34}$$

to measure the precision of the estimation, where r_0 acts like a correlation measure, to see what is the precision obtained with the pair of models. r_0 should be as small as possible, because the smaller the better.

6. An application

In order to illustrate the theory, we applied the obtained results to real data in a grapevine study. The data were taken from [11], considering two models with two crossed factors (Clone, which is a random factor with two levels and Location, which is a fixed-effects factor with three levels), of length 6. All computations were performed using R software, see [18].

The considered models may be written as

$$Y(u) = X_0\boldsymbol{\beta}_0(u) + X_1L_1(u) + X_2L_2(u), \quad u = 1, 2, \tag{35}$$

where

$$X_0 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}^\top,$$

$$X_1 = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}^\top,$$

$$X_2 = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}^\top,$$

$\boldsymbol{\beta}_0$ is fixed and $L_1(u)$ and $L_2(u)$, $u = 1, 2$, are random and uncorrelated vectors, with null mean vectors and variance–covariance matrices $\chi_1(2)(u)I_2$ and $\chi_2(2)(u)I_3$, $u = 1, 2$, with I_c the $c \times c$ identity matrix.

The considered observation vectors are

$$Y(1) = [1.85, 0.70, 1.50, 1.50, 0.65, 0.90]^\top,$$

$$Y(2) = [1.45, 0.40, 1.15, 1.25, 0.90, 0.35]^\top.$$

To make the adjustments, we needed to have $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_5$ constituting an orthonormal basis for $\Omega^\perp$, the orthogonal complement of the space Ω. Thus, we considered

$$A = \begin{bmatrix} -\frac{1}{2} & 0 & \frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{\sqrt{12}} & -\frac{2}{\sqrt{12}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} & -\frac{2}{\sqrt{12}} & \frac{1}{\sqrt{12}} \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{1}{2} & 0 & -\frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{2} \\ -\frac{1}{\sqrt{12}} & -\frac{2}{\sqrt{12}} & -\frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} & -\frac{2}{\sqrt{12}} & \frac{1}{\sqrt{12}} \end{bmatrix},$$

being $\boldsymbol{\alpha}_h^\top$ the hth row of A, $h = 1, \ldots, 5$. So, from (7), we obtained $\dot{Y}(1) = AY(1)$ and $\dot{Y}(2) = AY(2)$, obtaining

$$\tilde{\mathcal{O}}_2(\dot{Y}) = \frac{\dot{Y}^2(1) + \dot{Y}^2(2)}{2},$$

$$\tilde{\mathcal{O}}_3(\dot{Y}) = \frac{\dot{Y}^3(1) + \dot{Y}^3(2)}{2},$$

$$\tilde{\mathcal{O}}_4(\dot{Y}) = \frac{\dot{Y}^4(1) + \dot{Y}^4(2)}{2} - 3\dot{Y}^2(1)\dot{Y}^2(2).$$

The obtained values for $\tilde{\mathcal{O}}_2(\dot{Y})$, $\tilde{\mathcal{O}}_3(\dot{Y})$ and $\tilde{\mathcal{O}}_4(\dot{Y})$ are given by

$$\tilde{\mathcal{O}}_2(\dot{Y}) = [0.29, 0.49, 0.10, 0.05, 0.20]^\top,$$

$$\tilde{\mathcal{O}}_3(\dot{Y}) = [-0.16, 0.39, -0.03, -0, 01, -0.10]^\top,$$

$$\tilde{\mathcal{O}}_4(\dot{Y}) = [-0.153, -0.173, -0.006, -0.00005, -0.003]^\top.$$

In order to compare our results, we also calculated $\mathcal{O}_2(\dot{Y})$, $\mathcal{O}_3(\dot{Y})$ and $\mathcal{O}_4(\dot{Y})$, using the 'moments' package in software R. The results obtained are the same as those obtained using our approach.

From (13), with $r = 2$, we obtain

$$B(2) = \begin{bmatrix} 0 & 2 \\ 0 & 2 \\ 3 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Note that matrix $B(2)$ has many null elements. This happens because the orthogonal projection matrix on the range space of X_0 commutes with matrices $X_l X_l^\top$, $l = 1, 2$. This shows the great flexibility of this approach, which allows these matrices to commute, or not.

From (14) we obtained the LSE for the variance components, $\chi(2)$,

$$\tilde{\chi}(2) = [0.03 \quad 0.20].$$

We also estimated the variance components using the 'VCA' package in R software. The obtained estimates are $[0.04 \quad 0.23]$ and $[0.00 \quad 0.07]$, using Y_1 and Y_2, as observation vectors, respectively. Note that when we estimated the variance components using our approach, we obtained $\tilde{\mathcal{O}}_2(\dot{Y})$ by combining both observation vectors, while with the 'VCA' package we estimated the variance components, using first Y_1 and Y_2. Therefore, as expected, the same values were not obtained. However, the difference between the estimates obtained with our approach and the mean of those obtained using the 'VCA' package is small. Finally, from (16), we obtained the estimator of $\boldsymbol{\beta}_0$, given by

$$\tilde{\boldsymbol{\beta}}_0 = [0.79 \quad 0.16].$$

7. Conclusion

The paper has demonstrated two main ideas. The first one is how to avoid complications in the calculus of the fourth-order central moment in additive models. One may consider cumulants and use a pair of independent and identically distributed models, obtaining thus unbiased estimators. The second idea is that one may use cumulants for additive models

to carry out the estimation of the parameters of such models, for example, the variance components. It is possible to estimate not only the second-order moments but their third- and fourth-order central moments, besides the remaining estimable vectors.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Sandra S. Ferreira 🔵 http://orcid.org/0000-0002-9209-7772
Dário Ferreira 🔵 http://orcid.org/0000-0001-9095-0947
Célia Nunes 🔵 http://orcid.org/0000-0003-0167-4851

References

[1] D.F. Andrews, *Asymptotic expansions of moments and cumulants*, Stat. Comput. 11 (2001), pp. 7–16.
[2] D.F. Andrews and J.E. Stafford, *Symbolic Computation for Statistical Inference*, Oxford University Press, Oxford, 2000.
[3] N. Balakrishnan, N.L. Johnson, and S. Kotz, *A note on relationships between moments, central moments and cumulants from multivariate distributions*, Stat. Probab. Lett. 1 (1998), pp. 49–54.
[4] T. Caliński and S. Kageyama, *Block Designs: A Randomization Approach. Vol. I: Analysis*, Lecture Notes in Statistics, 150, Springer-Verlag, New York, 2000.
[5] T. Caliński and S. Kageyama, *Block Designs: A Randomization Approach. Vol. II: Design*, Lecture Notes in Statistics, 170, Springer-Verlag, New York, 2003.
[6] R. Christensen, *Plane Answers to Complex Questions: The Theory of Linear Models*, 4th ed., Springer-Verlag, New York, 2002.
[7] C.C. Craig, *On a property of the semi-invariants of Thiele*, Ann. Math. Statist. 2 (1931), pp. 154–164.
[8] E. Di Nardo, G. Guarino, and D. Senato, *Symbolic computation of moments of sampling distributions*, Comput. Statist. Data. Anal. 52 (2008), pp. 4909–4922.
[9] J.B. Douglas, *Analysis with Standard Contagious Distributions*, International Co-operative Publishing House, Fairland, MD, 1980.
[10] D. Ferreira, S.S. Ferreira, C. Nunes, M. Fonseca, A. Silva, and J.T. Mexia, *Estimation and incommutativity in mixed models*, J. Multivariate Anal. 161 (2017), pp. 58–67.
[11] M. Fonseca, J.T. Mexia, and R. Zmyślony, *Estimating and testing of variance components: An application to a grapevine experiment*, Biom. Lett. 40 (2003), pp. 1–7.
[12] J. Johnston and J. DiNardo, *Econometric Methods*, 4th ed., University of California, Irvine, CA, 2001.
[13] M.G. Kendall, A. Stuart, and J.K. Ord, *Kendall's Advanced Theory of Statistics*, Griffin, London, 1987.

[14] P. McCullagh, *Tensor Methods in Statistics*, Chapman and Hall, London, 1987.

[15] A. Mukherjea, M. Rao, and S. Suen, *A note on moment generating functions*, Stat. Probab. Lett. 76 (2006), pp. 1185–1189.

[16] G. Pistone and H.P. Wynn, *Cumulant varieties*, J. Symbolic Comput. 2 (2006), pp. 210–221.

[17] I.E. Poloskov, *CAS Mathematica in random studies*, Lect. Notes Comput. Sci. 2657 (2003), pp. 781–790.

[18] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, 2013. Available at http://www.R-project.org/.

[19] P.J. Smith, *A recursive formulation of the old problem of obtaining moments from cumulants and vice versa*, Amer. Statist. 49 (2012), pp. 217–218.

[20] K.T. Takeaki and H. Kurata, *Generalized Least Squares*, John Wiley & Sons Ltd, Chichester, 2004.

[21] C.S. Withers, S. Nadarajah, and S.H. Shih, *Moments and cumulants of a mixture*, Methodol. Comput. Appl. Probab. 17 (2015), pp. 541–564.

[22] Q. Zheng, *Computing relations between moments and cumulants*, Comput. Stat. 17 (2002), pp. 507–515.