

ESTIMATORS AND TESTS  
FOR VARIANCE COMPONENTS  
IN CROSS NESTED ORTHOGONAL DESIGNS

MIGUEL FONSECA, JOÃO TIAGO MEXIA

*Department of Mathematics, Faculty of Science and Technology*  
*New University of Lisbon*

*Monte da Caparica 2829-516 Caparica, Portugal*

**e-mail:** fonsecamig@yahoo.com

AND

ROMAN ZMYŚLONY

*Institute of Mathematics, University of Zielona Góra*

*Podgórna 50, 65-246 Zielona Góra, Poland*

**e-mail:** r.zmyslony@im.uz.zgora.pl

**Abstract**

Explicit expressions of UMVUE for variance components are obtained for a class of models that include balanced cross nested random models. These estimators are used to derive tests for the nullity of variance components. Besides the usual  $F$  tests, generalized  $F$  tests will be introduced. The separation between both types of tests will be based on a general theorem that holds even for mixed models. It is shown how to estimate the  $p$ -value of generalized  $F$  tests.

**Keywords:** hypothesis testing, generalized  $F$  distribution, adaptive test, nested orthogonal designs.

**2000 Mathematics Subject Classification:** 62H10, 62H15, 62J10.

## 1. INTRODUCTION

We start by presenting the models we consider. These include the usual balanced cross nested random models. Nextly we obtain UMVUE for the variance components. These estimators are then used to derive  $F$  tests for the nullity of single variance components. The separation between both types of tests will be based on a general theorem that holds also for mixed models.

Under certain conditions, exact distributions were obtained for the generalized  $F$  test statistics under the null hypothesis (see [7]). These expressions are useful in validating the estimation of  $p$ -values for these tests using Monte-Carlo methods. This estimation is convenient since the distribution of these statistics depends, as we shall see, on nuisance parameters.

When  $\mathbf{y}^n$  is normal, with mean vector  $\mathbf{X}\boldsymbol{\beta}^n$  and covariance matrix  $\boldsymbol{\Sigma}$ , we put  $\mathbf{y}^n \sim \mathcal{N}(\mathbf{X}\boldsymbol{\beta}^n, \boldsymbol{\Sigma})$ . Let

$$(1) \quad \boldsymbol{\Sigma} = \sum_{i=1}^k \sigma_i^2 \mathbf{V}_i$$

with  $\mathbf{X}$ ,  $\mathbf{V}_i$ ,  $i = 1, \dots, k-1$  known matrices, and  $\mathbf{V}_k = \mathbf{I}_n$ . With  $\mathbf{X}^+$  the Moore-Penrose generalized inverse of matrix  $\mathbf{X}$ , the orthogonal projection matrices on the range space  $R(\mathbf{X})$  of  $\mathbf{X}$  and its orthogonal complement will be  $\mathbf{P} = \mathbf{X}\mathbf{X}^+$  and  $\mathbf{M} = \mathbf{I}_n - \mathbf{P}$ , respectively. Moreover, if  $\mathbf{P}\mathbf{V}_i = \mathbf{V}_i\mathbf{P}$ ,  $i = 1, \dots, k$ , and the vector sub-space  $\mathcal{V} = \text{sp}\{\mathbf{M}\mathbf{V}_1, \dots, \mathbf{M}\mathbf{V}_k\}$ , spanned by the matrices  $\mathbf{M}\mathbf{V}_1, \dots, \mathbf{M}\mathbf{V}_k$ , is a quadratic sub-space (containing the square of its matrices), with dimension  $k$ , then, see [8] and [11], we have the minimal sufficient complete statistics

$$(2) \quad \mathbf{X}'\mathbf{y}^n \text{ and } \mathbf{y}^{n'}\mathbf{E}_i\mathbf{y}^n, i = 1, \dots, k$$

with  $\mathbf{E}_i = \mathbf{M}\mathbf{V}_i$ . Moreover, see [7], if  $\mathcal{V}$  is commutative, there exists a unique basis  $\{\mathbf{E}_1, \dots, \mathbf{E}_k\}$ , such that  $\mathbf{E}_i\mathbf{E}_j = \delta_{ij}\mathbf{E}_i$ ,  $i, j = 1, \dots, k$ , with  $\delta_{ij}$  the Kronecker delta. Thus the model will be called regular.

Now we consider a special case of regular models.

## 2. MODELS

Throughout the text, superscripts will indicate vector dimensions,  $\mathbf{1}^v$  and  $\mathbf{0}^v$  will have their  $v$  components equal to 1 and 0, respectively. Moreover,  $\mathbf{I}_v$  will be the  $v \times v$  identity matrix, while  $R(\mathbf{A})$  and  $N(\mathbf{A})$  will be the range space and the kernel of matrix  $\mathbf{A}$ . Besides the usual matrix product we will

also use the Kronecker matrix product ( $\otimes$ ). The properties of this product are studied in detail in detail in [10]. The transpose and inverse of matrix  $\mathbf{A}$  will be denoted by  $\mathbf{A}'$  and  $\mathbf{A}^{-1}$ , respectively. With  $\Gamma$  the set of vectors  $\mathbf{h}^L$  with integer components  $h_l = 0, \dots, u_l$ ,  $l = 1, \dots, L$ , we assume, for the vector  $\mathbf{y}^n$  of observations, the model

$$(3) \quad \mathbf{y}^n = \sum_{\mathbf{h}^L \in \Gamma} \mathbf{X}(\mathbf{h}^L) \boldsymbol{\beta}^{c(\mathbf{h}^L)}(\mathbf{h}^L) + \mathbf{e}^n$$

where  $\mathbf{X}(\mathbf{0}^L) = \mathbf{1}^n$ ,  $\boldsymbol{\beta}^{c(\mathbf{0}^L)}(\mathbf{0}^L) = \mu$ , and, with  $\mathbf{h}^L \neq \mathbf{0}^L$ ,  $\mathbf{X}(\mathbf{h}^L) = \bigotimes_{l=1}^L X_l(h_l) \otimes \mathbf{1}^r$ . The vectors  $\boldsymbol{\beta}^{c(\mathbf{h}^L)}(\mathbf{h}^L)$ , with  $\mathbf{h}^L \neq \mathbf{0}^L$ , and  $\mathbf{e}^n$  will be normal, independent, with null mean vectors and covariance matrices  $\sigma^2(\mathbf{h}^L) \mathbf{I}_{c(\mathbf{h}^L)}$  and  $\sigma^2 \mathbf{I}_n$ , respectively. The crossing was integrated in the model through the use of the Kronecker matrix product, while, to have nesting, we require that  $R(\mathbf{X}_l(h_l)) \subset R(\mathbf{X}_l(h_l + 1))$ ,  $h_l = 0, \dots, u_l - 1$ ,  $l = 1, \dots, L$ . This means that  $R(\mathbf{X}_l(h_l))$  is strictly included in  $R(\mathbf{X}_l(h_l + 1))$ . Lastly, model orthogonality is achieved by assuming that  $\mathbf{M}_l(h_l) = \mathbf{X}_l(h_l) \mathbf{X}_l'(h_l) = b(h_l) \mathbf{Q}_l(h_l)$ , with  $\mathbf{Q}_l(h_l)$  an orthogonal projection matrix,  $h_l = 0, \dots, u_l$ ,  $l = 1, \dots, L$ . In order to have  $\mathbf{X}(\mathbf{0}^L) = \mathbf{1}^n$  we must have  $\mathbf{X}_l(0) = \mathbf{1}^{n_l}$ ,  $l = 1, \dots, L$ , as well as  $n = r \prod_{l=1}^L n_l$ . Then  $\mathbf{Q}_l(0) = \frac{1}{n_l} \mathbf{J}_{n_l}$ ,  $l = 1, \dots, L$ , when  $\mathbf{J}_v = \mathbf{1}^v \mathbf{1}^{v'}$ . As we shall see, this last assumption will enable us to derive a Jordan algebra  $\mathcal{A}$  automatically associated to the model. Such algebra corresponds, as it is well known, to orthogonal partitions and so, in this way, the orthogonality of the model is ensured.

On account of nesting we will also have the orthogonal projection matrices  $\mathbf{B}_l(0) = \mathbf{Q}_l(0)$  and  $\mathbf{B}_l(t_l) = \mathbf{Q}_l(t_l) - \mathbf{Q}_l(t_l - 1)$ ,  $t_l = 1, \dots, u_l$ ,  $l = 1, \dots, L$ . The Kronecker matrix product enables us to the orthogonal projection matrices:

$$(4) \quad \begin{cases} \mathbf{Q}(\mathbf{h}^L) = \bigotimes_{l=1}^L \mathbf{Q}_l(h_l) \otimes \frac{1}{r} \mathbf{J}_r, & \mathbf{h}^L \in \Gamma \\ \mathbf{B}(\mathbf{t}^L) = \bigotimes_{l=1}^L \mathbf{B}_l(t_l) \otimes \frac{1}{r} \mathbf{J}_r, & \mathbf{t}^L \in \Gamma \end{cases}.$$

If matrices  $\mathbf{Q}_l(h_l)$ ,  $h_l = 0, \dots, u_l$ , have ranks  $r_l(0) = 1$  and  $r_l(h_l)$ ,  $h_l = 1, \dots, u_l$ , matrices  $\mathbf{B}_l(t_l)$ ,  $t_l = 0, \dots, u_l$ , will have ranks  $g_l(0) = 1$  and  $g_l(t_l) = r_l(t_l) - r_l(t_l - 1)$ ,  $t_l = 1, \dots, u_l$ ,  $l = 1, \dots, L$ . Moreover, matrices  $\mathbf{Q}(\mathbf{h}^L)$  [ $\mathbf{B}(\mathbf{t}^L)$ ] with  $\mathbf{u}^L \in \Gamma$  [ $\mathbf{t}^L \in \Gamma$ ] will have rank

$$r(\mathbf{h}^L) = t \prod_{l=1}^L r_l(h_l) \left[ g(\mathbf{t}^L) = \prod_{l=1}^L g_l(t_l) \right].$$

In order to illustrate these concepts, we now present an example: a three factor random model, in which the second factor nests the third. We then get the model

$$(5) \quad \mathbf{Y}^n = \sum_{\mathbf{h}^2 \in \Gamma} \mathbf{X}(\mathbf{h}^2) \beta^{c(\mathbf{h}^2)}(\mathbf{h}^2) + \mathbf{e}^n,$$

where  $\mathbf{h}^2 = (i_1, i_2)$ ,  $i = 0, 1$ ,  $j = 0, 1, 2$ . Taking  $a_1(1)$  as the number of levels of the first factor,  $a_2(1)$  the number of levels of the second and  $a_2(2)$  the number of levels of the third (nested on the second), with  $a_1(0) = a_0(2) = 1$ . We then get the matrices

$$(6) \quad \left\{ \begin{array}{l} \mathbf{X}(1, 0) = \mathbf{I}_{a_1(1)} \otimes \mathbf{1}^{a_2(1)a_2(2)} \\ \mathbf{X}(0, 1) = \mathbf{1}^{a_1(1)} \otimes \mathbf{I}_{a_2(1)} \otimes \mathbf{1}^{a_2(2)} \\ \mathbf{X}(0, 2) = \mathbf{1}^{a_1(1)} \otimes \mathbf{I}_{a_2(1)a_2(2)} \\ \mathbf{X}(1, 1) = \mathbf{I}_{a_1(1)a_2(1)} \otimes \mathbf{1}^{a_2(2)} \\ \mathbf{X}(1, 2) = \mathbf{I}_{a_1(1)a_2(1)a_2(2)} \end{array} \right. ,$$

having also  $\beta^1(0, 0) = \mu$  and  $\beta^{c(i,j)}(i, j) \sim \mathcal{N}(\mathbf{0}^{c(i,j)}, \sigma^2(i, j)\mathbf{I}_{c(i,j)})$ ,  $i = 0, 1$ ,  $j = 0, 1, 2$ ,  $(i, j) \neq (0, 0)$ . Consequently

$$(7) \quad \left\{ \begin{array}{l} \mathbf{Q}(1, 0) = \frac{1}{a_2(1)a_2(2)} \mathbf{I}_{a_1(1)} \otimes \mathbf{J}_{a_2(1)a_2(2)} \\ \mathbf{Q}(0, 1) = \frac{1}{a_1(1)a_2(2)} \mathbf{J}_{a_1(1)} \otimes \mathbf{I}_{a_2(1)} \otimes \mathbf{J}_{a_2(2)} \\ \mathbf{Q}(0, 2) = \frac{1}{a_1(1)} \mathbf{J}_{a_1(1)} \otimes \mathbf{I}_{a_2(1)a_2(2)} \\ \mathbf{Q}(1, 1) = \frac{1}{a_2(2)} \mathbf{I}_{a_1(1)a_2(1)} \otimes \mathbf{J}_{a_2(2)} \\ \mathbf{Q}(1, 2) = \mathbf{I}_{a_1(1)a_2(1)a_2(2)} \end{array} \right. .$$

Consequently:

$$(8) \quad \begin{cases} \mathbf{B}(1, 0) = \frac{1}{a_2(1)a_2(2)r} \mathbf{K}_{a_1(1)} \otimes \mathbf{J}_{a_2(1)a_2(2)r} \\ \mathbf{B}(0, 1) = \frac{1}{a_1(1)a_2(2)r} \mathbf{J}_{a_1(1)} \otimes \mathbf{K}_{a_2(1)} \otimes \mathbf{J}_{a_2(2)r} \\ \mathbf{B}(0, 2) = \frac{1}{a_1(1)r} \mathbf{J}_{a_1(1)} \otimes \mathbf{I}_{a_2(1)} \mathbf{K}_{a_2(2)} \otimes \mathbf{J}_r \\ \mathbf{B}(1, 1) = \frac{1}{a_2(2)r} \mathbf{K}_{a_1(1)} \otimes \mathbf{K}_{a_2(1)} \otimes \mathbf{J}_{a_2(2)r} \\ \mathbf{B}(1, 2) = \frac{1}{r} \mathbf{K}_{a_1(1)} \otimes \mathbf{I}_{a_2(1)} \otimes \mathbf{K}_{a_2(2)} \otimes \mathbf{J}_r \end{cases},$$

with  $r$  being the number of repetitions and  $\mathbf{K}_v = \mathbf{I}_v - \frac{1}{v} \mathbf{J}_v$ , and

$$(9) \quad \begin{cases} \text{rank}(\mathbf{B}(1, 0)) = a_1(1) - 1 \\ \text{rank}(\mathbf{B}(0, 1)) = a_2(1) - 1 \\ \text{rank}(\mathbf{B}(0, 2)) = a_2(1)(a_2(2) - 1) \\ \text{rank}(\mathbf{B}(1, 1)) = (a_1(1) - 1)(a_2(1) - 1) \\ \text{rank}(\mathbf{B}(1, 2)) = (a_1(1) - 1)a_2(1)(a_2(2) - 1) \end{cases}.$$

This is in fact the model used in [1].

Now, in order to relate matrices  $\mathbf{Q}(\mathbf{h}^L)$ ,  $\mathbf{u}^L \in \Gamma$ , and  $\mathbf{B}(\mathbf{t}^L)$ ,  $\mathbf{t}^L \in \Gamma$ , we are going to consider certain sub-sets of  $\Gamma$ . Let  $\mathbf{u}^L \wedge \mathbf{v}^L$  [ $\mathbf{u}^L \vee \mathbf{v}^L$ ] have components  $\min\{u_l, v_l\}$  [ $\max\{u_l, v_l\}$ ],  $l = 1, \dots, L$ , and put  $\mathbf{t}_-^L = (\mathbf{t}^L - \mathbf{1}^L) \vee \mathbf{0}^L$  and  $\mathbf{h}_+^L = (\mathbf{h}^L + \mathbf{1}^L) \wedge \mathbf{u}^L$ , as well as

$$(10) \quad \begin{cases} \sqcap(\mathbf{t}^L) = \{\mathbf{h}^L : \mathbf{t}_-^L \leq \mathbf{h}^L \leq \mathbf{t}^L\} \\ \sqcup(\mathbf{h}^L) = \{\mathbf{t}^L : \mathbf{h}^L \leq \mathbf{t}^L \leq \mathbf{h}_+^L\} \end{cases}$$

in order to get, first.

**Proposition 1.** *We have  $\mathbf{Q}(\mathbf{h}^L) = \sum_{\mathbf{t}^L \leq \mathbf{h}^L} \mathbf{B}(\mathbf{t}^L)$ , as well as*

$$\mathbf{B}(\mathbf{t}^L) = \sum_{\mathbf{h}^L \in \Pi(\mathbf{t}^L)} (-1)^{m(\mathbf{h}^L, \mathbf{t}^L)} \mathbf{Q}(\mathbf{h}^L)$$

with  $m(\mathbf{h}^L, \mathbf{t}^L)$  the number of components of  $\mathbf{h}^L$  that are smaller than the corresponding components of  $\mathbf{t}^L$ .

**Proof.** The first part of the thesis follows from  $\mathbf{Q}_l(h_l) = \sum_{t_l \leq h_l} \mathbf{B}_l(t_l)$ ,  $h_l = 0, \dots, u_l$ ,  $l = 1, \dots, L$ , and from the distributive properties of the Kronecker product. Next, the first part of the thesis enables us to write

$$\sum_{\mathbf{h}^L \in \Pi(\mathbf{t}^L)} (-1)^{m(\mathbf{h}^L, \mathbf{t}^L)} \mathbf{Q}(\mathbf{h}^L) = \sum_{\mathbf{h}^L \in \Pi(\mathbf{t}^L)} (-1)^{m(\mathbf{h}^L, \mathbf{t}^L)} \sum_{\mathbf{v}^L \leq \mathbf{h}^L} \mathbf{B}(\mathbf{v}^L).$$

If  $\mathbf{v}^L < \mathbf{t}^L$ ,  $\mathbf{B}(\mathbf{v}^L)$  will enter in the terms of the  $\mathbf{h}^L$  such that  $\mathbf{v}^L \leq \mathbf{h}^L \leq \mathbf{t}^L$ . Since we can choose in  $\binom{m(\mathbf{v}^L, \mathbf{t}^L)}{a}$  ways  $a$  of the components of  $\mathbf{h}^L$  that are smaller than the corresponding components of  $\mathbf{t}^L$ , this being the number of vectors  $\mathbf{h}^L$  such that  $m(\mathbf{h}^L, \mathbf{t}^L) = a$ . If  $m(\mathbf{v}^L, \mathbf{t}^L) > 0$  the coefficient of  $\mathbf{B}(\mathbf{v}^L)$  in the previous expression of  $\sum_{\mathbf{h}^L \in \Pi(\mathbf{t}^L)} (-1)^{m(\mathbf{h}^L, \mathbf{t}^L)} \mathbf{Q}(\mathbf{h}^L)$  will be

$$\sum_{a=0}^{m(\mathbf{v}^L, \mathbf{t}^L)} \binom{m(\mathbf{v}^L, \mathbf{t}^L)}{a} (-1)^a = 0$$

since  $a = m(\mathbf{h}^L, \mathbf{t}^L)$ . If  $m(\mathbf{h}^L, \mathbf{t}^L) = 0$  we must have  $\mathbf{v}^L = \mathbf{h}^L = \mathbf{t}^L$  and  $\mathbf{B}(\mathbf{v}^L)$  only enters, with coefficient one, in  $\sum_{\mathbf{h}^L \in \Pi(\mathbf{t}^L)} (-1)^{m(\mathbf{h}^L, \mathbf{t}^L)} \sum_{\mathbf{v}^L \leq \mathbf{h}^L} \mathbf{B}(\mathbf{v}^L)$ . Thus we get

$$\mathbf{B}(\mathbf{t}^L) = \sum_{\mathbf{v}^L \in \Pi(\mathbf{t}^L)} (-1)^{m(\mathbf{v}^L, \mathbf{t}^L)} \mathbf{Q}(\mathbf{v}^L).$$

■

**Corollary 1.** *With  $S(\mathbf{t}^L) = \|\mathbf{B}(\mathbf{t}^L)\mathbf{y}^n\|^2$  we have*

$$\begin{aligned} S(\mathbf{t}^L) &= \sum_{\mathbf{h}^L \in \Pi(\mathbf{t}^L)} (-1)^{m(\mathbf{h}^L, \mathbf{t}^L)} \|\mathbf{Q}(\mathbf{h}^L) \mathbf{y}^n\|^2 \\ &= \sum_{\mathbf{h}^L \in \Pi(\mathbf{t}^L)} \frac{(-1)^{m(\mathbf{h}^L, \mathbf{t}^L)}}{b(\mathbf{h}^L)} \|\mathbf{X}'(\mathbf{h}^L) \mathbf{y}^n\|^2 \end{aligned}$$

with  $b(\mathbf{h}^L) = r \prod_{l=1}^L b_l(h_l)$ .

**Proof.** The first part of the thesis follows directly from Proposition 1. As to the second part, since  $\mathbf{X}(\mathbf{h}^L)\mathbf{X}'(\mathbf{h}^L) = b(\mathbf{h}^L)\mathbf{Q}(\mathbf{h}^L)$  and orthogonal projection matrices are idempotent and symmetric,

$$\begin{aligned} \|\mathbf{Q}(\mathbf{h}^L)\mathbf{y}^n\|^2 &= \mathbf{y}^{n'} \mathbf{Q}'(\mathbf{h}^L) \mathbf{Q}(\mathbf{h}^L) \mathbf{y}^n = \mathbf{y}^{n'} \mathbf{Q}(\mathbf{h}^L) \mathbf{y}^n \\ &= \mathbf{y}^{n'} \left( \frac{1}{b(\mathbf{h}^L)} \mathbf{X}(\mathbf{h}^L) \mathbf{X}'(\mathbf{h}^L) \right) \mathbf{y}^n = \frac{1}{b(\mathbf{h}^L)} \|\mathbf{X}'(\mathbf{h}^L) \mathbf{y}^n\|^2, \end{aligned}$$

the rest of the proof being straightforward. ■

Besides matrices  $\mathbf{B}(\mathbf{t}^L)$ ,  $\mathbf{t}^L \in \Gamma$ , we take

$$(11) \quad \mathbf{B} = \mathbf{I}_n - \mathbf{B}(\mathbf{u}^L) = \mathbf{I}_n - \sum_{\mathbf{t}^L \in \Gamma} \mathbf{B}(\mathbf{t}^L)$$

to obtain a basis for a Jordan algebra  $\mathcal{A}$ . This basis, see [12], is the sole basis of  $\mathcal{A}$  constituted by mutually orthogonal orthogonal projection matrices. This will be the principal basis of  $\mathcal{A}$ . Moreover matrices  $\mathbf{Q}(\mathbf{h}^L)$ ,  $\mathbf{h}^L \in \Gamma$ , and  $\mathbf{I}_n$  belong to  $\mathcal{A}$ .  $\mathbf{R}(\mathbf{B}(\mathbf{t}^L))$ , with  $\mathbf{t}^L \in \Gamma$ , and  $\mathbf{R}(\mathbf{B})$  constitute an orthogonal partition of  $\mathbb{R}^n$  into mutually orthogonal sub-spaces.

The mean vector and covariance matrix of  $\mathbf{y}^n$  will be, with  $\sigma^2(\mathbf{0}^L) = 0$

$$(12) \quad \begin{cases} \mathbb{E}[\mathbf{y}^n] = \mathbf{1}^n \mu \\ \mathbb{V}[\mathbf{y}^n] = \sum_{\mathbf{h}^L \in \Gamma} \sigma^2(\mathbf{h}^L) \mathbf{M}(\mathbf{h}^L) + \sigma^2 \mathbf{I}_n = \sum_{\mathbf{t}^L \in \Gamma} \gamma(\mathbf{t}^L) \mathbf{B}(\mathbf{t}^L) + \sigma^2 \mathbf{B} \end{cases}$$

with

$$(13) \quad \gamma(\mathbf{t}^L) = \sigma^2 + \sum_{\mathbf{h}^L \geq \mathbf{t}^L} b(\mathbf{h}^L) \sigma^2(\mathbf{h}^L),$$

since  $\sigma^2(\mathbf{h}^L)\mathbf{M}(\mathbf{h}^L) = b(\mathbf{h}^L)\sigma^2(\mathbf{h}^L)\sum_{\mathbf{t}^L \leq \mathbf{h}^L} \mathbf{B}(\mathbf{t}^L)$ ,  $\mathbf{h}^L \in \Gamma$ . It is the fact that  $\mathbb{V}[\mathbf{y}^n]$  is a linear combination of the matrices in the basis of  $\mathcal{A}$ , which are mutually orthogonal, that gives the model its orthogonality.

Still using model (5), it is easy to check that

$$(14) \quad \left\{ \begin{array}{l} \gamma(1, 0) = a_2(1)a_2(2)r\sigma^2(1, 0) + a_2(2)r\sigma^2(1, 1) + r\sigma^2(1, 2) + \sigma^2 \\ \gamma(0, 1) = a_1(1)a_2(2)r\sigma^2(0, 1) + a_1(1)a_2(1)r\sigma^2(0, 2) + a_2(2)r\sigma^2(1, 1) \\ \quad + r\sigma^2(1, 2) + \sigma^2 \\ \gamma(0, 2) = a_1(1)a_2(1)r\sigma^2(0, 2) + a_2(2)r\sigma^2(1, 1) + r\sigma^2(1, 2) + \sigma^2 \\ \gamma(1, 1) = a_2(2)r\sigma^2(1, 1) + r\sigma^2(1, 2) + \sigma^2 \\ \gamma(1, 2) = r\sigma^2(1, 2) + \sigma^2 \end{array} \right. .$$

We now go over to the balanced case in which we will have  $L$  groups of  $u_1, \dots, u_L$  factors, *i.e.*, the generalization of model (5). When  $u_l > 1$  there is balanced nesting for the factors of the  $l$ -th group. This is, for each of the  $a_l(1)$  levels of the first factor in the group there will be  $a_l(2)$  levels of the second factor. Then, for each of the  $a_l(1) \times a_l(2)$  levels of the second factor there will be  $a_l(3)$  of the third factor, and so on.

$$(15) \quad \mathbf{X}_l(h_l) = \mathbf{I}_{c_l(h_l)} \otimes \mathbf{1}^{b_l(h_l)}, \quad h_l = 0, \dots, u_l, \quad l = 1, \dots, L$$

with  $c_l(h_l) = \prod_{t=1}^{h_l} a_l(t)$  and  $b_l = c_l(u_l)/c_l(h_l)$ , as well as

$$(16) \quad \mathbf{M}_l(h_l) = \mathbf{I}_{c_l(h_l)} \otimes \left( \frac{1}{b_l(h_l)} \mathbf{J}_{b_l(h_l)} \right) = \frac{1}{b_l(h_l)} \mathbf{Q}_l(h_l),$$

$$h_l = 0, \dots, u_l, \quad l = 1, \dots, L$$

where  $\mathbf{Q}_l(h_l) = \mathbf{I}_{c_l(h_l)} \otimes \left( \frac{1}{b_l(h_l)} \mathbf{J}_{b_l(h_l)} \right)$ ,  $h_l = 0, \dots, u_l$ ,  $l = 1, \dots, L$ . Defining  $\mathbf{K}_v = \mathbf{I}_v - \frac{1}{v} \mathbf{J}_v$  we also get



$$(17) \quad \mathbf{B}_l(t_l) = \mathbf{I}_{c_l(t_l-1)} \otimes \mathbf{K}_{a_l(t_l)} \otimes \left( \frac{1}{b_l(t_l)} \mathbf{J}_{b_l(t_l)} \right) = \mathbf{A}_l'(t_l) \mathbf{A}_l(t_l),$$

$$t_l = 1, \dots, u_l, l = 1, \dots, L$$

with

$$\mathbf{A}_l(t_l) = \frac{1}{\sqrt{b_l(t_l)}} \mathbf{I}_{c_l(t_l-1)} \otimes \mathbf{K}_{a_l(t_l)} \otimes \mathbf{1}^{b_l(t_l)'}, \quad t_l = 1, \dots, u_l, l = 1, \dots, L.$$

Putting  $\mathbf{A}_l(0) = \mathbf{1}^{n_l'}$ ,  $l = 1, \dots, L$ , and, if we have  $r$  replicates for each factor combination we define the following matrices

$$\mathbf{A}(\mathbf{t}^L) = \frac{1}{\sqrt{r}} \bigotimes_{l=1}^L \mathbf{A}_l(t_l) \otimes \mathbf{1}^{r'}, \quad \mathbf{t}^L \in \Gamma.$$

The sums of squares and corresponding degrees of freedom for a given factor or interaction will be

$$(18) \quad \begin{cases} S(\mathbf{t}^L) = \|\mathbf{A}(\mathbf{t}^L) \mathbf{y}^n\|^2, & \mathbf{t}^L \in \Gamma \\ g(\mathbf{t}^L) = \prod_{l=1}^L g_l(t_l) = \prod_{l=1}^L (c_l(t_l) - c_l(t_l - 1)), & \mathbf{t}^L \in \Gamma \end{cases}$$

with  $c_l(-1) = 0$  for  $l = 1, \dots, L$ .

Just for the record, we consider another special case. Let  $\mathbf{P}_l$  be a  $n_l \times n_l$  orthogonal projection matrix whose first column vector is  $\frac{1}{\sqrt{n_l}} \mathbf{1}^{n_l}$ ,  $l = 1, \dots, L$ . With  $c_l(0) = 1$  and  $c_l(h_l) < c_l(h_l + 1) \leq n_l$ ,  $h_l = 1, \dots, u_l - 1$ , we can take  $\mathbf{X}_l(h_l)$  to be constituted by the first  $c_l(h_l)$  columns of  $\mathbf{P}_l$ ,  $l = 1, \dots, L$ . It is easy to see that our assumptions hold in this case which is distinct from the balanced one. Putting  $\mathbf{A}(0) = \mathbf{X}_l'(0)$ ,  $l = 1, \dots, L$

$$(19) \quad \mathbf{X}_l(t_l) = \begin{bmatrix} \mathbf{X}_l(t_l - 1) & \mathbf{A}_l'(t_l) \end{bmatrix} \quad t_l = 1, \dots, u_l, l = 1, \dots, L$$

and  $\mathbf{A}(\mathbf{t}^L) = \bigotimes_{l=1}^L \mathbf{A}_l(t_l) \otimes \mathbf{1}^{r'}$  expressions in (18) continue to hold.

## 3. ESTIMATORS

We start by pointing out that for the models we are considering, the minimal complete statistics will be:  $y_{\bullet} = \frac{1}{n} \mathbf{1}^{n'} \mathbf{y}^n$ ,  $S(\mathbf{t}^L)$ , with  $\mathbf{t}^L \in \Gamma$ , and  $S = \|\mathbf{B} \mathbf{y}^n\|^2$ . The statistics  $S(\mathbf{t}^L)$ , with  $\mathbf{t}^L \in \Gamma$ , are given by the first expression in (18).

When  $\mathbf{0}^L < \mathbf{t}^L \leq \mathbf{u}^L$ ,  $\mathbf{B}(\mathbf{t}^L) \mathbf{y}^n$  will be normal with null mean vector and covariance matrix  $\gamma(\mathbf{t}^L) \mathbf{B}(\mathbf{t}^L)$ , so that  $S(\mathbf{t}^L)$  will be the product by  $\gamma(\mathbf{t}^L)$  of a central chi-square with  $g(\mathbf{t}^L) = \text{rank}(\mathbf{B}(\mathbf{t}^L))$  degrees of freedom, we put  $S(\mathbf{t}^L) \sim \gamma(\mathbf{t}^L) \chi_{(g(\mathbf{t}^L))}^2$ . Likewise we have  $S \sim \sigma^2 \chi_{(g)}^2$ , with  $g = n - r(\mathbf{u}^L)$ . Moreover,

$$(20) \quad \mathcal{F}(\mathbf{t}_1^L, \mathbf{t}_2^L) = \frac{g(\mathbf{t}_2^L)}{g(\mathbf{t}_1^L)} \cdot \frac{S(\mathbf{t}_1^L)}{S(\mathbf{t}_2^L)}$$

will be the product by  $\frac{\gamma(\mathbf{t}_1^L)}{\gamma(\mathbf{t}_2^L)}$  of a random variable with central  $F$  distribution with  $g(\mathbf{t}_1^L)$  and  $g(\mathbf{t}_2^L)$  degrees of freedom. Whenever  $g(\mathbf{t}_2^L) > 2$ ,

$$\mathbb{E}[\mathcal{F}(\mathbf{t}_1^L, \mathbf{t}_2^L)] = \frac{g(\mathbf{t}_2^L)}{g(\mathbf{t}_2^L) - 2} \cdot \frac{\gamma(\mathbf{t}_1^L)}{\gamma(\mathbf{t}_2^L)}$$

and so we have the UMVUE

$$(21) \quad \widehat{\left( \frac{\gamma(\mathbf{t}_1^L)}{\gamma(\mathbf{t}_2^L)} \right)} = \frac{g(\mathbf{t}_2^L) - 2}{g(\mathbf{t}_2^L)} \cdot \frac{S(\mathbf{t}_1^L)}{S(\mathbf{t}_2^L)}.$$

Later on we will use these estimators. If  $f_{q;r,s}$  is the quantile, for probability  $q$ , of the central  $F$  distribution with  $r$  and  $s$  degrees of freedom we have, for  $\frac{\gamma(\mathbf{t}_1^L)}{\gamma(\mathbf{t}_2^L)}$ , the  $(1 - q) \times 100\%$  confidence interval

$$\left[ \frac{\mathcal{F}(\mathbf{t}_1^L, \mathbf{t}_2^L)}{f_{1-\frac{q}{2};g_1,g_2}}, \frac{\mathcal{F}(\mathbf{t}_1^L, \mathbf{t}_2^L)}{f_{\frac{q}{2};g_1,g_2}} \right] \quad \text{with } g_i = g(\mathbf{t}_i^L), \quad i = 1, 2.$$

The  $1 - q$  level upper bound will likewise be  $\frac{\mathcal{F}(\mathbf{t}_1^L, \mathbf{t}_2^L)}{f_{1-q;g_1,g_2}}$

In the preceding section we defined the sets  $\sqcap(\mathbf{t}^L)$  and  $\sqcup(\mathbf{t}^L)$  and established Proposition 1. We now have.

**Proposition 2.** Whenever  $m(\mathbf{h}^L) = m(\mathbf{h}^L, \mathbf{u}^L) > 0$ ,

$$\sigma^2(\mathbf{h}^L) = \frac{1}{b(\mathbf{h}^L)} \sum_{\mathbf{t}^L \in \sqcup(\mathbf{h}^L)} (-1)^{m(\mathbf{h}^L, \mathbf{t}^L)} \gamma(\mathbf{t}^L).$$

**Proof.** From the definition of  $\gamma(\mathbf{t}^L)$ , we get

$$\sum_{\mathbf{t}^L \in \sqcup(\mathbf{h}^L)} (-1)^{m(\mathbf{h}^L, \mathbf{t}^L)} \gamma(\mathbf{t}^L) = \sum_{\mathbf{t}^L \in \sqcup(\mathbf{h}^L)} (-1)^{m(\mathbf{h}^L, \mathbf{t}^L)} \left( \sigma^2 + \sum_{\mathbf{v}^L \geq \mathbf{t}^L} b(\mathbf{v}^L) \sigma^2(\mathbf{v}^L) \right).$$

Now, in  $\sqcup(\mathbf{h}^L)$  we have  $\binom{m(\mathbf{h}^L)}{z}$  (with  $z$  ranging from 0 to  $m(\mathbf{h}^L)$ ) vectors  $\mathbf{t}^L$  such that  $m(\mathbf{h}^L, \mathbf{t}^L) = z$ , since this is the number of choices of  $z$  components of  $\mathbf{h}^L$ , from the  $m(\mathbf{h}^L)$  components that are smaller than those of  $\mathbf{u}^L$ . So, these components of  $\mathbf{h}^L$  can be increased by 1, the resulting vector still belonging to  $\sqcup(\mathbf{h}^L)$ . Thus, the coefficient of  $\sigma^2$  will be  $\sum_{z=0}^{m(\mathbf{h}^L)} \binom{m(\mathbf{h}^L)}{z} (-1)^z = 0$ . Likewise, if  $\mathbf{v}^L > \mathbf{h}^L$ ,  $\sigma^2(\mathbf{v}^L)$  appears in the terms associated with the  $\mathbf{t}^L \in \sqcup(\mathbf{h}^L)$  such that  $\mathbf{t}^L \leq \mathbf{v}^L$ . We can reason as before to show that there are  $\binom{m(\mathbf{h}^L, \mathbf{v}^L)}{z}$ ,  $z = 0, \dots, m(\mathbf{h}^L, \mathbf{v}^L)$ , such vectors  $\mathbf{t}^L \in \sqcup(\mathbf{h}^L)$  with  $m(\mathbf{h}^L, \mathbf{v}^L) = z$ , and so, the coefficient of  $b(\mathbf{h}^L) \sigma^2(\mathbf{h}^L)$  will be  $\sum_{z=0}^{m(\mathbf{h}^L, \mathbf{v}^L)} \binom{m(\mathbf{h}^L, \mathbf{v}^L)}{z} (-1)^z = 0$ . Lastly when  $\mathbf{v}^L = \mathbf{h}^L$ ,  $b(\mathbf{v}^L) \sigma^2(\mathbf{v}^L) = b(\mathbf{h}^L) \sigma^2(\mathbf{h}^L)$  only enters in the term associated with  $\mathbf{t}^L = \mathbf{h}^L$  with coefficient 1, so that  $\sum_{\mathbf{t}^L \in \sqcup(\mathbf{h}^L)} (-1)^{m(\mathbf{h}^L, \mathbf{t}^L)} \gamma(\mathbf{t}^L) = b(\mathbf{h}^L) \sigma^2(\mathbf{h}^L)$  and the thesis is established. ■

**Corollary 2.**  $\hat{\gamma}(\mathbf{t}^L) = \frac{S(\mathbf{t}^L)}{g(\mathbf{t}^L)}$ ,  $\mathbf{t}^L \in \Gamma$  and, if  $m(\mathbf{h}^L) > 0$ ,

$$\hat{\sigma}^2(\mathbf{h}^L) = \frac{1}{b(\mathbf{h}^L)} \sum_{\mathbf{t}^L \in \sqcup(\mathbf{h}^L)} (-1)^{m(\mathbf{h}^L, \mathbf{t}^L)} \hat{\gamma}(\mathbf{t}^L)$$

will be UMVUE.

Moreover, for  $\sigma^2$  and  $\sigma^2(\mathbf{u}^L)$  we have the UMVUE

$$(22) \quad \begin{cases} \hat{\sigma}^2 = \frac{S}{g} \\ \hat{\sigma}^2(\mathbf{u}^L) = \frac{1}{b(\mathbf{u}^L)} (\hat{\gamma}(\mathbf{u}^L) - \hat{\sigma}^2) \end{cases},$$

since  $S \sim \sigma \chi_{(g)}^2$  and  $\gamma(\mathbf{u}^L) = \sigma^2 + b(\mathbf{u}^L) \sigma^2(\mathbf{u}^L)$ . We point out that  $m(\mathbf{h}^L) = 0$  if and only if  $\mathbf{h}^L = \mathbf{u}^L$ .

Let  $\sqcup(\mathbf{h}^L)_+$  and  $\sqcup(\mathbf{h}^L)_-$  be the vectors of  $\sqcup(\mathbf{h}^L)$  for which  $m(\mathbf{h}^L, \mathbf{t}^L)$  is even or is odd. If  $m(\mathbf{h}^L) > 0$ , with

$$\sigma^2(\mathbf{h}^L)_+ = \frac{1}{b(\mathbf{h}^L)} \sum_{\mathbf{t}^L \in \sqcup(\mathbf{h}^L)_+} \gamma(\mathbf{t}^L)$$

and

$$\sigma^2(\mathbf{h}^L)_- = \frac{1}{b(\mathbf{h}^L)} \sum_{\mathbf{t}^L \in \sqcup(\mathbf{h}^L)_-} \gamma(\mathbf{t}^L)$$

we have

$$(23) \quad \sigma^2(\mathbf{h}^L) = \sigma^2(\mathbf{h}^L)_+ - \sigma^2(\mathbf{h}^L)_-.$$

We will also have the UMVUE estimators

$$(24) \quad \begin{cases} \hat{\sigma}^2(\mathbf{h}^L)_+ = \frac{1}{b(\mathbf{h}^L)} \sum_{\mathbf{t}^L \in \sqcup(\mathbf{h}^L)_+} \hat{\gamma}(\mathbf{t}^L) \\ \hat{\sigma}^2(\mathbf{h}^L)_- = \frac{1}{b(\mathbf{h}^L)} \sum_{\mathbf{t}^L \in \sqcup(\mathbf{h}^L)_-} \hat{\gamma}(\mathbf{t}^L). \end{cases}$$

We will now show that the estimators we obtained are efficient in the Fisher sense.

Using the formula for the determinants of matrices in a Jordan algebra given in the appendix we get

$$(25) \quad \det(\mathbb{V}[\mathbf{y}^n]) = \left( \prod_{\mathbf{t}^L \in \Gamma} \gamma(\mathbf{t}^L)^{g(\mathbf{t}^L)} \right) (\sigma^2)^g,$$

so that, if  $\boldsymbol{\eta}^v$  has components  $\mu, \gamma(\mathbf{t}^L), \mathbf{0}^L < \mathbf{t}^L \leq \mathbf{u}^L$ , and  $\sigma^2$  we have the log-likelihood

$$(26) \quad \begin{aligned} l(\boldsymbol{\eta}^v) = & -\frac{n}{2\gamma(\mathbf{0}^L)}(y_\bullet - \mu)^2 - \frac{1}{2} \sum_{\mathbf{0}^L \leq \mathbf{t}^L \leq \mathbf{u}^L} \frac{S(\mathbf{t}^L)}{\gamma(\mathbf{t}^L)} - \frac{S}{2\sigma^2} - \frac{n}{\log}(2\pi) \\ & - \frac{1}{2} \sum_{\mathbf{t}^L \in \Gamma} g(\mathbf{t}^L) \log \gamma(\mathbf{t}^L) - \frac{g}{2} \log \sigma^2. \end{aligned}$$

We omitted  $\gamma(\mathbf{0}^L)$  from the components of  $\boldsymbol{\eta}^v$  since this parameter is clearly a function of the  $\gamma(\mathbf{t}^L)$ ,  $\mathbf{0}^L < \mathbf{t}^L \leq \mathbf{u}^L$ . It is now straightforward to show that  $\tilde{\mu} = y_\bullet$ ,  $\tilde{\gamma}(\mathbf{t}^L)$ ,  $\mathbf{0}^L < \mathbf{t}^L \leq \mathbf{u}^L$ , and  $\tilde{\sigma}^2$  are maximum likelihood estimators and that their Fisher information matrix

$$(27) \quad \mathbf{Z}(\boldsymbol{\eta}^v) = \mathbb{E}[\text{grad}(l(\boldsymbol{\eta}^v)) (\text{grad}(l(\boldsymbol{\eta}^v)))']$$

is the diagonal matrix with principal elements  $\frac{n}{2\gamma(\mathbf{0}^L)}$ ,  $\frac{n}{2\gamma(\mathbf{0}^L)}$ ,  $\frac{g(\mathbf{t}^L)}{2\gamma(\mathbf{t}^L)}$ ,  $\mathbf{0}^L < \mathbf{t}^L \leq \mathbf{u}^L$ , and  $\frac{g}{2\sigma^2}$ . Since these are the inverses of the variances of these estimators we see that the Cramer-Rao lower bound for the variance of unbiased estimators is attained for these estimators,

$$(28) \quad \mathbb{V}[\hat{\boldsymbol{\eta}}^v]^{-1} = \mathbf{Z}(\boldsymbol{\eta}^v).$$

Let  $\boldsymbol{\theta}^v$  have components  $\mu$ ,  $\sigma^2(\mathbf{t}^L)$ ,  $\mathbf{0}^L < \mathbf{t}^L \leq \mathbf{u}^L$ , and  $\sigma^2$ . Then  $\boldsymbol{\theta}^v = \mathbf{K}\boldsymbol{\eta}^v$  with  $\mathbf{K}$  a regular matrix so that  $\boldsymbol{\eta}^v = \mathbf{K}^{-1}\boldsymbol{\theta}^v$  and, using the new parameters, we have for the gradient of the log-likelihood

$$(29) \quad \text{grad}(l(\boldsymbol{\theta}^v)) = (\mathbf{K}^{-1})' \text{grad}(l(\boldsymbol{\eta}^v)).$$

Thus the Fisher information matrix for the new parameters will be

$$(30) \quad \mathbf{Z}(\boldsymbol{\theta}^v) = (\mathbf{K}^{-1})' \mathbf{Z}(\boldsymbol{\eta}^v) \mathbf{K}^{-1}$$

and, with  $\boldsymbol{\theta}^v = \mathbf{K}\boldsymbol{\eta}^v$ , we have the covariance matrix

$$(31) \quad \mathbb{V}[\hat{\boldsymbol{\theta}}^v] = \mathbf{K} \mathbb{V}[\hat{\boldsymbol{\eta}}^v] \mathbf{K}' = \mathbf{K} \mathbf{Z}(\boldsymbol{\eta}^v) \mathbf{K}' = \mathbf{Z}(\boldsymbol{\theta}^v).$$

According to (28), to (31) and to the invariance principle of the maximum likelihood estimators we established.

**Proposition 3.**  *$\hat{\boldsymbol{\eta}}^v$  and  $\hat{\boldsymbol{\theta}}^v$  are maximum likelihood estimators for which the Cramer-Rao lower bounds for covariance matrices are attained.*

## 4. HYPOTHESIS TESTING

4.1.  $F$  Tests for regular models

We now consider the regular models in Section 2. The test for  $H_0 : \sigma_{i_0}^2 = 0$  versus  $H_1 : \sigma_{i_0}^2 \neq 0$  for  $i_0 < k$  is, see [12], based on an UMVUE for  $\sigma_{i_0}^2$ . Namely, in that paper it was shown that if the sub-model obtained under the null hypothesis was regular, the test statistic

$$(32) \quad \mathcal{F} = \frac{\mathbf{y}^{n'} \mathbf{A}_+ \mathbf{y}^n}{\mathbf{y}^{n'} \mathbf{A}_- \mathbf{y}^n},$$

where  $\mathbf{y}^{n'} \mathbf{A} \mathbf{y}^n = \mathbf{y}^{n'} \mathbf{A}_+ \mathbf{y}^n - \mathbf{y}^{n'} \mathbf{A}_- \mathbf{y}^n$  is the UMVUE of  $\sigma_{i_0}^2$ , has an  $F$  distribution when  $H_0$  holds,  $\mathbf{A}_+$  and  $\mathbf{A}_-$  being the positive and negative parts of the symmetric matrix  $\mathbf{A}$ .

Now we establish the inverse result

**Theorem 1.** *If  $\mathcal{F}$  has  $F$  distribution when  $H_0$  holds, the corresponding sub-model will be regular.*

**Proof.** From Lemma 6 in [7], it follows that for  $\mathcal{U} = \text{sp}\{\mathbf{M}\mathbf{V}_1, \dots, \mathbf{M}\}$  there exists a unique base  $\mathbf{E}_1, \dots, \mathbf{E}_k$  such that  $\mathbf{E}_i \mathbf{E}_j = \delta_{ij} \mathbf{E}_j$ , where  $\delta_{ij}$  is the Kronecker delta. Thus, from the assumption of  $F$  distribution the UMVUE for  $\sigma_{i_0}^2$  should have the structure

$$(33) \quad \mathbf{y}^{n'} \mathbf{A} \mathbf{y}^n = c \left( \frac{\mathbf{y}^{n'} \mathbf{E}_i \mathbf{y}^n}{\text{rank}(\mathbf{E}_i)} - \frac{\mathbf{y}^{n'} \mathbf{E}_j \mathbf{y}^n}{\text{rank}(\mathbf{E}_j)} \right), \quad i \neq j.$$

Clearly  $c \neq 0$ . Note that  $\mathbf{V}_l$  can be expressed as follows:

$$(34) \quad \mathbf{V}_l = \sum_{m=1}^k \lambda_{l,m} \mathbf{E}_m.$$

From (33) and the unbiasedness assumption it follows that for  $l \neq i_0$  and  $\sigma_l^2 \geq 0$ ,  $\sigma_l^2 c \lambda_{l,i} = \sigma_l^2 c \lambda_{l,j}$ . Thus, whenever  $\sigma_l^2 > 0$  we have  $c \sigma_l^2 \neq 0$  as well as  $\lambda_{l,i} = \lambda_{l,j}$  for  $l \neq i_0$ . This means that  $\mathcal{V}_{i_0} = \text{sp}\{\mathbf{E}_1, \dots, \mathbf{E}_i + \mathbf{E}_j, \dots, \mathbf{E}_k\}$  is a quadratic commutative sub-space, and so the corresponding sub-model is regular. ■

In the next two sections we present first  $F$  tests, and then generalized  $F$  tests for hypothesis on variance components.

#### 4.2. $F$ Tests

In this section we derive  $F$  tests for the hypothesis

$$(35) \quad \begin{cases} H_0(\mathbf{h}^L) : \sigma^2(\mathbf{h}^L) = 0, & m(\mathbf{h}^L) \leq 1 \\ \overline{H}_0(\mathbf{h}^L) : \gamma(\mathbf{h}^L) = \sigma^2, & \mathbf{h}^L \in \Gamma. \end{cases}$$

We start with  $H_0(\mathbf{u}^L)$ . Since  $S(\mathbf{u}^L) \sim (r\sigma^2(\mathbf{u}^L) + \sigma^2) \chi_{(g(\mathbf{u}^L))}^2$  independent from  $S \sim \sigma^2 \chi_{(g)}^2$ . The test statistic

$$(36) \quad \mathcal{F}(\mathbf{u}^L) = \frac{g}{g(\mathbf{u}^L)} \cdot \frac{S(\mathbf{u}^L)}{S}$$

will have central  $F$  distribution with  $g(\mathbf{u}^L)$  and  $g$  degrees of freedom when  $H_0(\mathbf{u}^L)$  holds. Likewise, if  $m(\mathbf{h}^L) = 1$ ,  $\sqcup(\mathbf{h}^L)_+ = \{\mathbf{h}^L\}$  and  $\sqcup(\mathbf{h}^L)_- = \{\mathbf{u}^L\}$ . Since

$$S(\mathbf{h}^L) \sim (rb(\mathbf{h}^L)\sigma^2(\mathbf{h}^L) + r\sigma^2(\mathbf{u}^L) + \sigma^2) \chi_{(g(\mathbf{h}^L))}^2$$

independent from  $S(\mathbf{u}^L)$ ,

$$(37) \quad \mathcal{F}(\mathbf{u}^L) = \frac{g(\mathbf{u}^L)}{g(\mathbf{h}^L)} \cdot \frac{S(\mathbf{h}^L)}{S(\mathbf{u}^L)}$$

will have a central  $F$  distribution with  $g(\mathbf{h}^L)$  and  $g(\mathbf{u}^L)$  degrees of freedom, when  $H_0(\mathbf{h}^L)$  holds.

When  $m(\mathbf{u}^L) > 1$  there is more than one vector in  $\sqcup(\mathbf{h}^L)_+$  and in  $\sqcup(\mathbf{h}^L)_-$  so that we cannot pursue this straightforward derivation of  $F$  tests. In the next section we will consider how to test  $H_0(\mathbf{h}^L)$  when  $m(\mathbf{h}^L) > 1$ . We refer the possibilities, see [4], of using either Bartlett-Scheffé  $F$  tests or the Satterthwaite approximation of the distribution of  $F$  statistics. We will not consider these approaches,

since the first one discards information, for instance see [4] pg. 43, while the second may lead to difficulties (see [4] pg. 40, in the control of test size).

No such problems arise while testing  $\overline{H}_0(\mathbf{h}^L)$ ,  $\mathbf{h}^L \in \Gamma$ . Since  $S(\mathbf{h}^L) \sim \gamma(\mathbf{h}^L)\chi^2_{(g(\mathbf{h}^L))}$  independent of  $S \sim \sigma^2\chi^2_{(g)}$  we get the  $F$  statistic

$$(38) \quad \overline{\mathcal{F}(\mathbf{h}^L)} = \frac{g}{g(\mathbf{h}^L)} \cdot \frac{S(\mathbf{h}^L)}{S}, \quad \mathbf{h}^L \in \Gamma$$

with  $g(\mathbf{h}^L)$  and  $g$  degrees of freedom, when  $\overline{H}_0(\mathbf{h}^L)$  holds.

As a parting remark we point out that  $\overline{H}_0(\mathbf{h}^L)$  holds if and only if the  $H_0(\mathbf{v}^L)$ , with  $\mathbf{h}^L \leq \mathbf{v}^L \leq \mathbf{u}^L$ , hold. From Theorem 1 it follows that this statistic can have  $F$  distribution if and only if  $m(\mathbf{h}^L) > 1$  we will have to consider generalized  $F$  tests. We now present such tests.

### 4.3. Generalized $F$ tests

#### 4.3.1. Test statistics

These statistics introduced by Michalski and Zmysłony, see [12] and [14], can be written as

$$(39) \quad \mathcal{F}(\mathbf{h}^L) = \frac{\hat{\sigma}^2(\mathbf{h}^L)_+}{\hat{\sigma}^2(\mathbf{h}^L)_-}$$

with  $\hat{\sigma}^2(\mathbf{h}^L)_+$  and  $\hat{\sigma}^2(\mathbf{h}^L)_-$  defined in (24).

It is interesting to observe that  $\sigma^2(\mathbf{h}^L) = \sigma^2(\mathbf{h}^L)_+ - \sigma^2(\mathbf{h}^L)_- \geq 0$ , thus, with  $\lambda(\mathbf{h}^L) = \frac{\sigma^2(\mathbf{h}^L)_+}{\sigma^2(\mathbf{h}^L)_-}$ , both  $H_0(\mathbf{h}^L)$  and the corresponding alternative may be written as

$$(40) \quad \begin{cases} H_0(\mathbf{h}^L) : \lambda(\mathbf{h}^L) = 1 \\ H_1(\mathbf{h}^L) : \lambda(\mathbf{h}^L) > 1. \end{cases}$$



We now consider a two-dimensional presentation of the behavior of the test statistic (Figure 1). Let  $f$  be the critical value.

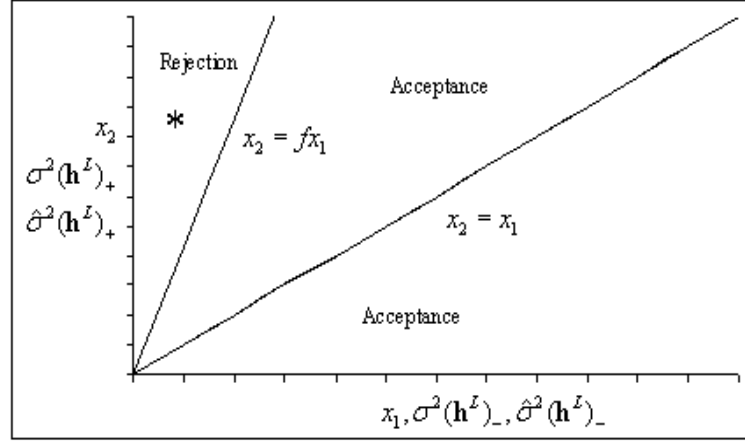


Figure 1. Acceptance and rejection regions.

In this presentation we have the points, with coordinates  $(x_1, x_2)$ ,  $(\sigma^2(\mathbf{h}^L)_-, \sigma^2(\mathbf{h}^L)_+)$  and  $(\hat{\sigma}^2(\mathbf{h}^L)_-, \hat{\sigma}^2(\mathbf{h}^L)_+)$ . The semi-straight line  $x_1 = f x_2$  separates the rejection from the acceptance region, while  $(*)$  indicates a clear alternative to  $H_0(\sigma^2(\mathbf{h}^L))$ . Since  $\hat{\sigma}^2(\mathbf{h}^L)_+$  and  $\hat{\sigma}^2(\mathbf{h}^L)_-$  are UMVUE of  $\sigma^2(\mathbf{h}^L)_+$  and  $\sigma^2(\mathbf{h}^L)_-$  it is expected that, when this alternative holds,  $(\hat{\sigma}^2(\mathbf{h}^L)_-, \hat{\sigma}^2(\mathbf{h}^L)_+)$  will be close to  $(*)$ , thus inside the rejection region. This heuristic argument clearly points towards the use of the test statistic  $\mathcal{F}(\mathbf{h}^L)$ .

After the figure, we now take a closer look at  $\mathcal{F}(\mathbf{h}^L)$  in order to be able to use it properly. To lighten the writing we put  $m = m(\mathbf{h}^L)$  taking  $w = 2^{m-1}$ ,  $\sqcup(\mathbf{h}^L)_+ = \{\mathbf{t}_1^L, \dots, \mathbf{t}_w^L\}$  and  $\sqcup(\mathbf{h}^L)_- = \{\mathbf{t}_{w+1}^L, \dots, \mathbf{t}_{2w}^L\}$ . Let  $\mathbf{g}_1^w$  and  $\mathbf{g}_2^w$  have components  $g_{1,j} = g(\mathbf{t}_j^L)$  and  $g_{2,j} = g(\mathbf{t}_{j+w}^L)$ ,  $j = 1, \dots, w$ , where we can assume that  $g_{i,1} = \min\{g_{i,1}, \dots, g_{i,w}\}$  and  $g_{i,w} = \max\{g_{i,1}, \dots, g_{i,w}\}$ ,  $i = 1, 2$ . Then

$$(41) \quad \mathcal{F}(\mathbf{h}^L) = \lambda(\mathbf{h}^L) \frac{\sum_{j=1}^w \frac{p_{1,j}}{g_{1,j}} \chi_{(g_{1,j})}^2}{\sum_{j=1}^w \frac{p_{2,j}}{g_{2,j}} \chi_{(g_{2,j})}^2},$$

with

$$p_{1,j} = \frac{\gamma(\mathbf{t}_j^L)}{\sigma^2(\mathbf{h}^L)_+} \text{ and } p_{2,j} = \frac{\gamma(\mathbf{t}_{j+w}^L)}{\sigma^2(\mathbf{h}^L)_-}, \quad j = 1, \dots, w,$$

and, as before,  $\lambda(\mathbf{h}^L) = \frac{\sigma^2(\mathbf{h}^L)_+}{\sigma^2(\mathbf{h}^L)_-}$ . Since

$$\sum_{j=1}^w p_{1,j} = \sum_{j=1}^w p_{2,j} = 1, \quad \mathcal{F}(\mathbf{h}^L)$$

will be the product by  $\lambda(\mathbf{h}^L)$  of the quotient of two convex combinations of independent central chi-squares divided by their degrees of freedom. Thus, it is natural to consider  $\mathcal{F}(\mathbf{h}^L)$  as a generalized  $F$  statistic. The  $p_{1,j}$  and  $p_{2,j}$ ,  $j = 1, \dots, w$  will be the components of vectors  $\mathbf{p}_1^w$  and  $\mathbf{p}_2^w$ . These vectors will be nuisance parameters. Since

$$(42) \quad \left\{ \begin{array}{l} p_{1,j} = \frac{\frac{\gamma(\mathbf{t}_j^L)}{\gamma(\mathbf{t}_w^L)}}{\frac{\sum_{j'=1}^w \gamma(\mathbf{t}_{j'}^L)}{\gamma(\mathbf{t}_w^L)}}, \quad j = 1, \dots, w \\ p_{2,j} = \frac{\frac{\gamma(\mathbf{t}_{j+w}^L)}{\gamma(\mathbf{t}_{2w}^L)}}{\frac{\sum_{j'=1}^w \gamma(\mathbf{t}_{j'+w}^L)}{\gamma(\mathbf{t}_{2w}^L)}}, \quad j = 1, \dots, w \end{array} \right.$$

we can use  $\widehat{\left( \frac{\gamma(\mathbf{t}_j^L)}{\gamma(\mathbf{t}_w^L)} \right)}$  and  $\widehat{\left( \frac{\gamma(\mathbf{t}_{j+w}^L)}{\gamma(\mathbf{t}_{2w}^L)} \right)}$  to estimate  $p_{1,j}$  and  $p_{2,j}$ ,  $j = 1, \dots, w$ .

Now, when  $H_0(\mathbf{h}^L)$  holds, we have

$$(43) \quad \mathcal{F}(\mathbf{h}^L) = \frac{Y_1}{Y_2}$$

with

$$Y_i = \sum_{j=1}^w p_{i,j} \frac{\chi_{(g_{i,j})}^2}{g_{i,j}}, \quad i = 1, 2.$$

Since the  $Y_1$  and  $Y_2$  are independent with mean value 1, and the partial derivatives of  $\frac{Y_1}{Y_2}$ , at point  $(1, 1)$ , are 1 and  $-1$ , we will have, when  $H_0(\mathbf{h}^L)$  holds

$$(44) \quad \mathcal{F}(\mathbf{h}^L) \approx 1 + (Y_1 - 1) - (Y_2 - 1)$$

so that, according to the independence of  $Y_1$  and  $Y_2$

$$(45) \quad \begin{cases} \mathbb{E}[\mathcal{F}(\mathbf{h}^L)] \approx 1 \\ \mathbb{V}[\mathcal{F}(\mathbf{h}^L)] \approx \mathbb{V}[Y_1] + \mathbb{V}[Y_2] \end{cases}.$$

We also have

$$(46) \quad \mathbb{V}[Y_i] = 2 \sum_{j=1}^w \frac{p_{i,j}^2}{g_{i,j}}, \quad i = 1, 2.$$

If we look at the Figure 2

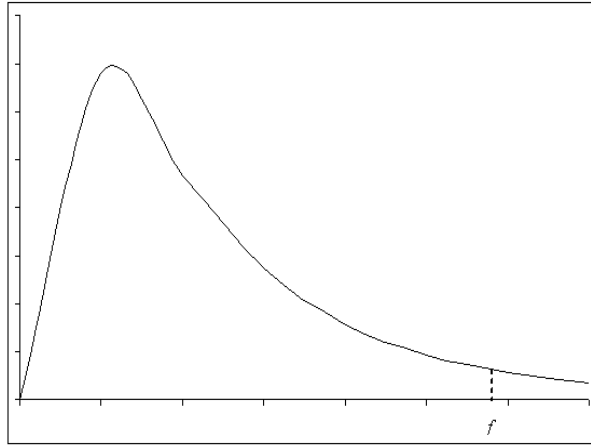


Figure 2. Critical value.

we are led to think that, when  $H_0(\mathbf{h}^L)$  holds, the probability of rejection increases with  $\sum_{i=1}^2 \mathbb{V}[Y_i]$ . Putting  $g_i = \sum_{j=1}^w g_{i,j}$  and  $u_{i,j} = p_{i,j} - \frac{1}{g_i} g_{i,j}$ ,  $j = 1, \dots, w$ ,  $i = 1, 2$ , we get

$$(47) \quad \mathbb{V}[Y_i] = \frac{2}{g_i} + 2 \sum_{j=1}^w \frac{u_{i,j}^2}{g_{i,j}}, \quad i = 1, 2$$

so that, according to (46) and (47),  $\frac{2}{g_i} < \mathbb{V}[Y_i] < \frac{2}{g_{i,1}}$ ,  $i = 1, 2$ .

It is interesting to observe that when  $\mathbb{V}[Y_i]$  attains its lower [upper] bound,  $Y_i$  is a central chi-square with  $g_i$  [ $g_{i,1}$ ] degrees of freedom divided by this number. Thus, when both lower or both upper bounds are attained,  $\mathcal{F}(\mathbf{h}^L)$  has, when  $H_0(\mathbf{h}^L)$  holds, the central  $F$  distribution with  $g_1$  and  $g_2$  or  $g_{1,1}$  and  $g_{2,1}$  degrees of freedom.

As we shall see, besides exact expressions for the distribution of  $\mathcal{F}(\mathbf{h}^L)$  that hold under certain conditions, Monte-Carlo methods may be used to “table” such a distribution. Thus it may be worthwhile to check if both  $F$  statistics represent the least and the most favorable possibilities for the control of the first type error.

In the balanced case we have some restrictions on the  $p_{i,j}$ ,  $j = 1, \dots, w$ ,  $i = 1, 2$ . When  $H_0(\mathbf{h}^L)$  holds  $\sigma^2(\mathbf{h}^L)_+ = \sigma^2(\mathbf{h}^L)_-$ , and we have  $\gamma(\mathbf{h}^L) = \gamma(\mathbf{t}_1^L) = \max\{\gamma(\mathbf{t}_1^L), \dots, \gamma(\mathbf{t}_w^L)\} \geq \max\{\gamma(\mathbf{t}_{w+1}^L), \dots, \gamma(\mathbf{t}_{2w}^L)\}$ , so that

$$(48) \quad \begin{cases} p_{1,1} = \max\{p_{1,1}, \dots, p_{1,w}\} \geq \frac{1}{w} > \frac{g(\mathbf{h}^L)}{g_1} \\ p_{1,1} \geq \max\{p_{2,1}, \dots, p_{2,w}\} \end{cases}.$$

When  $w = 2$  we can take  $p_{1,1} = p_1$ ,  $p_{1,2} = 1 - p_1$ ,  $p_{2,1} = p_2$  and  $p_{2,2} = 1 - p_2$  to get

$$(49) \quad \max\{p_2, 1 - p_2\} \leq p_1.$$

Getting back to the general case in order to control the first type error, we can

- place ourselves under the least favorable situation in which, as we saw, we have an  $F$  test with  $g_{1,1}$  and  $g_{2,1}$  degrees of freedom;
- use more completely the set of complete sufficient statistics. We then obtain, besides the test statistic, estimates to the nuisance parameters. In this case we will be carrying out an adaptative procedure similar to the one carried out in [1] and [15]. Likewise we may think we are using ancillary statistics, see [9].

#### 4.3.2. Exact distributions

In the preceding section we mentioned the possibility of estimating the  $p$ -value. This is feasible when we have an exact expression for the distribution of  $\mathcal{F}(\mathbf{h}^L)$  when  $H_0(\mathbf{h}^L)$  holds, since we can replace the argument in the

distribution by the value of the test statistic and the nuisance parameters by their estimators.

Let  $F(z|\mathbf{c}_1^r, \mathbf{c}_2^s, \mathbf{g}_1^r, \mathbf{g}_2^s)$  be the distribution of

$$(50) \quad Z = \frac{\sum_{j=1}^r c_{1,j} \frac{\chi_{(g_{1,j})}^2}{g_{1,j}}}{\sum_{j=1}^s c_{2,j} \frac{\chi_{(g_{2,j})}^2}{g_{2,j}}},$$

when all chi-squares are independent. Then, when  $H_0(\mathbf{h}^L)$  holds,  $\mathcal{F}(\mathbf{h}^L)$  will have distribution  $F(z|\mathbf{p}_1^w, \mathbf{p}_2^w, \mathbf{g}_1^w, \mathbf{g}_2^w)$ .

When  $\mathbf{g}_1^r = 2\mathbf{m}^r$  [ $\mathbf{g}_2^s = 2\mathbf{m}^s$ ] we say that the first [second] evenness condition holds. In [2] it is shown that

$$(51) \quad F(z|\mathbf{c}_1^r, \mathbf{c}_2^s, 2\mathbf{m}^r, \mathbf{g}_2^s) = \frac{1}{\prod_{v=2}^r (2a_v)^{b_v+1}} \sum_{k_1=0}^1 \sum_{j_1=0}^{k_1 b_1^+} \sum_{\left(\sum_{i=1}^r t_{1,i}=j_1\right)} \cdots \sum_{k_r=0}^1 \sum_{j_r=0}^{k_r b_r^+} \\ \sum_{\left(\sum_{i=1}^1 t_{r,i}=j_r\right)} \frac{(-1)^{\sum_{i=1}^r (k_i+j_i-t_{i,1})}}{(2a_1)^{j_1}} \prod_{i=2}^r \left( \frac{b_i^+!}{b_i! \prod_{v=1}^{i-1} t_{v,i+1-v}!} \left( \frac{1}{2a_i} - d(\mathbf{k}^{i-1}) \right)^{j_i-b_i^+-1} \right) \\ \frac{\left( \sum_{i=1}^r t_{i,1} \right)! z^{\sum_{v=1}^r t_{v,1}}}{\prod_{i=1}^r t_{i,1}!} \sum_{\left( \sum_{i=r+1}^{r+s} l_i = \sum_{v=1}^r t_{v,1} \right)} \prod_{i=r+1}^{r+s} \left( \frac{\Gamma(l_i + \frac{g_i}{2})}{l_i! \Gamma(\frac{g_i}{2})} \cdot \frac{1}{(2a_i)^{\frac{g_i}{2}}} \right. \\ \left. \left( \frac{1}{2a_i} + d(\mathbf{k}^r) z \right)^{-(l_i + \frac{g_i}{2})} \right)$$

with  $a_i = \frac{c_{1,i}}{g_{1,i}}$ ,  $i = 1, \dots, r$ ,  $a_{i+r} = \frac{c_{2,i}}{g_{2,i}}$ ,  $i = 1, \dots, s$ ,  $b_i = m_i - 1$ ,  $i = 1, \dots, r$ ,  $g_{i+r} = g_{2,i}$ ,  $i = 1, \dots, s$ ,

$$(52) \quad b_i^+ = b_i + \sum_{u=1}^{i-1} t_{u,i+1-u}, \quad i = 1, \dots, r$$

and, given  $w_u$  the largest index for non null components of  $\mathbf{k}^u = (k_1, \dots, k_u)'$ ,

$$(53) \quad d(\mathbf{k}^u) = \frac{1}{a_{w_u}}, \quad u = 1, \dots, r.$$

Moreover we also have

$$(54) \quad F(z|\mathbf{c}_1^r, \mathbf{c}_2^s, \mathbf{g}_1^r, 2\mathbf{m}^s) = 1 - F\left(\frac{1}{z} \middle| \mathbf{c}_2^s, \mathbf{c}_1^r, 2\mathbf{m}^s, \mathbf{g}_1^r\right).$$

When  $H_0(\mathbf{h}^L)$  holds,  $\mathcal{F}(\mathbf{h}^L)$  has distribution  $F(z|\mathbf{p}_1^w, \mathbf{p}_2^w, \mathbf{g}_1^w, \mathbf{g}_2^w)$  and so we can apply the results given above when either of the evenness conditions hold. Moreover these exact formulas can be used to show that there is an excellent agreement, see [1], between the results obtained through their use and those obtained through Monte-Carlo methods. Also corroborating this statement is the *Glivenko-Cantelli Theorem* (see [5], page 20), which states that the empirical sample distribution tends to the distribution of the population, providing thus a solid basis for the use of simulation techniques.

We now restrict ourselves to the balanced case. Our goal is to see when one or both of the evenness conditions hold. Since

$$(55) \quad g(\mathbf{h}^L) = \prod_{l=1}^L g_l(h_l)$$

it suffices that one of the  $g_l(h_l)$ ,  $l = 1, \dots, L$ , is even for  $g(\mathbf{h}^L)$  to be even. Besides this, with  $0 \leq h_l \leq u_l$ , we have as we saw

$$(56) \quad \begin{cases} g_l(0) = 1, & l = 1, \dots, L \\ g_l(h_l) = \left( \prod_{t_l=1}^{h_l-1} a_l(t_l) \right) (a_l(h_l) - 1), & h_l = 1, \dots, u_l, l = 1, \dots, L \end{cases}$$

where  $a_l(h_l)$  is the number of factors of the  $h_l$ -th factor in the  $l$ -th group of nested factors, and  $\prod_{t=1}^0 a_l(t) = 1$ ,  $l = 1, \dots, L$ . Thus if  $g_l(h_l)$ , with  $l \geq 1$ , is odd,  $a_l(h_l)$  must be even as well as  $g_l(h_*)$  with  $h_l \leq h_* \leq u_l$ . Thus in the sequence  $g_l(1), \dots, g_l(u_l)$  there can be only one odd number. This observation clearly points towards evenness conditions holding quite often. We will consider the sets of indexes  $\mathcal{C}(\mathbf{h}^L) = \{j : h_j < u_j\}$  and  $\mathcal{D}(\mathbf{h}^L) = \{l : h_l > 0\}$ , putting  $u(\mathbf{h}^L) = \#(\mathcal{C}(\mathbf{h}^L))$  and  $o(\mathbf{h}^L) = \#(\mathcal{D}(\mathbf{h}^L)) - 1$ .

With  $q_l(h_l)$  the number of even numbers in the pair  $(g_l(h_l), g_l(h_l + 1))$ , we get  $q_l(0) \leq 1$  since  $g_l(0) = 1$  so that only  $g_l(1)$  may be even. We also put

$$(57) \quad \begin{cases} k(\mathbf{h}^L) = \max_{l \in \mathcal{C}(\mathbf{h}^L)} q_l(h_l) \\ t(\mathbf{h}^L) = \min_{l \in \mathcal{C}(\mathbf{h}^L)} q_l(h_l) \end{cases}.$$

Let us establish

**Proposition 4.** *If  $k(\mathbf{h}^L) = 2$  or if there exists  $l' \notin \mathcal{C}(\mathbf{h}^L)$  such that  $g_{l'}(u_{l'})$  is even both evenness conditions are satisfied.*

**Proof.** If  $k(\mathbf{h}^L) = 2$  there will be  $l \in \mathcal{C}(\mathbf{h}^L)$  such that  $g_l(h_l)$  and  $g_l(h_l + 1)$  are even. Since either of these will be a factor in  $g(\mathbf{h}_*^L)$  whatever  $\mathbf{h}_*^L \in \sqcup(\mathbf{h}^L)$ , the first part of the thesis is established. The second part follows from  $g_{l'}(u_{l'})$  being, if it exists, a factor from whatever  $g(\mathbf{h}_*^L)$  with  $\mathbf{h}_*^L \in \sqcup(\mathbf{h}^L)$ . ■

Nextly we get

**Proposition 5.** *If  $k(\mathbf{h}^L) = t(\mathbf{h}^L) = 1$  one of the two evenness conditions holds.*

**Proof.** If  $k(\mathbf{h}^L) = t(\mathbf{h}^L) = 1$ , whenever  $l \in \mathcal{C}(\mathbf{h}^L)$ ,  $g_l(h_l)$  or  $g_l(h_l + 1)$  is even but not both since  $k(\mathbf{h}^L) = 1$ .

For  $g(\mathbf{h}_*^L)$ , with  $\mathbf{h}_*^L \in \sqcup(\mathbf{h}^L)$ , to be odd we must have, for  $l \in \mathcal{C}(\mathbf{h}^L)$ ,  $h_{*l} = h_l [h_l + 1]$  when  $g_l(h_l)$  is odd [even]. Thus there is at most one vector  $\mathbf{h}_0^L \in \sqcup(\mathbf{h}^L)$  such that  $g(\mathbf{h}_0^L)$  is odd.

Then, if  $\mathbf{h}_0^L \in \sqcup(\mathbf{h}^L)_+ [\sqcup(\mathbf{h}^L)_-]$  the first [second] evenness condition holds. ■

**Corollary 3.** *If  $o(\mathbf{h}^L) = L - 1$  at least one of the evenness conditions holds.*

**Proof.** If  $o(\mathbf{h}^L) = L - 1$  we have  $h_l \geq 1$ ,  $l = 1, \dots, L$  and so  $q_l(h_l) \geq 1$ , for all  $l \in \mathcal{C}(\mathbf{h}^L)$ , since, when  $g_l(h_l)$  is odd  $g_l(h_l + 1)$  is even, thus  $t(\mathbf{h}^L) \geq 1$  and the thesis follows from Propositions 3 and 4. ■

**Corollary 4.** *If  $a_{l,1}$  is odd,  $l = 1, \dots, L$ , at least one of the evenness conditions holds for all  $\mathbf{h}^L \in \Gamma$ .*

**Proof.** We will have  $q_l(0) = 1$ ,  $l = 1, \dots, L$ , as well as  $q_l(h_l) \geq 1$ ,  $h_l = 1, \dots, u_{l-1}$ . Thus  $t(\mathbf{h}^L) \geq 1$  for all  $\mathbf{h}^L \in \Gamma$ . To complete the proof we have only to apply Propositions 4 and 5. ■

These results show that quite often we can get exact expressions for  $F(z|\mathbf{c}_1^r, \mathbf{c}_2^s, \mathbf{g}_1^r, \mathbf{g}_2^s)$ . This is interesting in itself and enables us to check the accuracy obtained applying Monte-Carlo methods to evaluate  $F(z|\mathbf{c}_1^r, \mathbf{c}_2^s, \mathbf{g}_1^r, \mathbf{g}_2^s)$  (see [2]).

Monte-Carlo methods can be used to table distributions  $F(z|\mathbf{c}_1^r, \mathbf{c}_2^s, \mathbf{g}_1^r, \mathbf{g}_2^s)$  in a straightforward way since there is no problem in generating independent central chi-squares.

#### 4.3.3. Estimated $p$ -values

The distribution of the generalized  $F$  statistics depends on nuisance parameters, for which we obtained estimates. If either one or both of the evenness conditions hold the exact expression of the distribution  $F(z|\mathbf{c}_1^r, \mathbf{c}_2^s, \mathbf{g}_1^r, \mathbf{g}_2^s)$  can be used to obtain an estimated  $p$ -value. To do this we have only to replace  $z$  by the generalized  $F$  statistics and the nuisance parameters by their estimators.

If neither of the evenness conditions hold, we can use Monte-Carlo methods generating the required chi-squares and replacing, as before, the nuisance parameters by their estimators.

Moreover, when  $m(\mathbf{h}^L) = 1$  we can even obtain upper bounds for  $p$ -values. Then, as seen above, the probability of first type error will increase with  $p_1$  and for fixed  $p_1$ , have a minimum for  $p_2$  at  $\frac{g_3}{g_3+g_4}$ . Since

$$p_1 = \frac{\frac{\gamma_1}{\gamma_2}}{1 + \frac{\gamma_1}{\gamma_2}} \quad \text{and} \quad p_2 = \frac{\frac{\gamma_3}{\gamma_4}}{1 + \frac{\gamma_3}{\gamma_4}}, \quad p_1 \quad \text{and} \quad p_2 \quad \text{increase with} \quad \frac{\gamma_1}{\gamma_2} \quad \text{and} \quad \frac{\gamma_3}{\gamma_4}.$$

Given the independence between

$$\widehat{\left(\frac{\gamma_1}{\gamma_2}\right)} \quad \text{and} \quad \widehat{\left(\frac{\gamma_3}{\gamma_4}\right)}, \quad \text{if} \quad \mathbb{P}\left[\frac{\gamma_1}{\gamma_2} \in ]0, a[\right] = q_1 \quad \text{and} \quad \mathbb{P}\left[\frac{\gamma_3}{\gamma_4} \in ]b, c[\right] = q_2,$$

we will have

$$(58) \quad \mathbb{P}\left[\left(p_1 \in \left]0, \frac{a}{1+a}\right[\right) \wedge \left(p_2 \in \left]\frac{b}{1+b}, \frac{c}{1+c}\right[\right)\right] = q_1 q_2.$$

When  $0 < p_1 \leq \frac{a}{1+a}$  and  $\frac{b}{1+b} \leq p_2 \leq \frac{c}{1+c}$ , the maximum for the probability



of the first type error will be attained at

$$\left( \frac{a}{1+a}, \frac{b}{1+b} \right) \text{ if } \frac{c}{1+c} < \frac{g_3}{g_3+g_4}, \text{ at } \left( \frac{a}{1+a}, \frac{c}{1+c} \right) \text{ if } \frac{g_3}{g_3+g_4} < \frac{b}{1+b}$$

or at any of the two former points if  $\frac{b}{1+b} < \frac{g_3}{g_3+g_4} < \frac{c}{1+c}$ , since, see Figure 3,

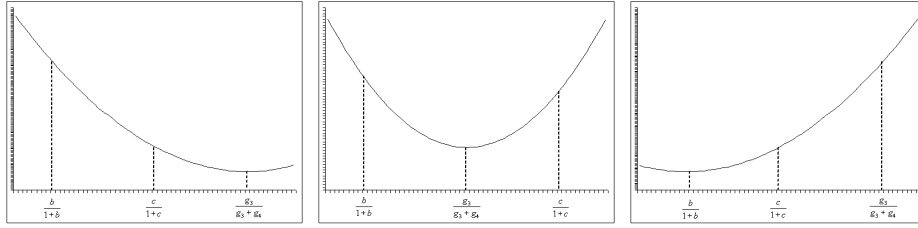


Figure 3. Maximum for first type error.

we are certainly in one of these three cases.

As a parting remark we point out that, in this section, we are following an approach inspired in adaptative tests, see [9].

## Appendix

### A. Determinants and Jordan Algebras

Let  $\nabla_j$ ,  $j = 1, \dots, v$ , be the range spaces of the mutually orthogonal matrices  $\mathbf{Q}_j$ ,  $j = 1, \dots, v$ , in the principal basis of the Jordan algebra  $\mathcal{A}$ . If the orthogonal direct sum of the  $\nabla_j$ ,  $j = 1, \dots, v$ ,  $\Omega$ , is a proper sub-space of  $\mathbb{R}^n$  we can complete this basis adding to it the orthogonal projection matrix  $\overline{\mathbf{Q}}$  on the orthogonal complement  $\Omega^\perp$  of  $\Omega$ . We thus have a new Jordan algebra  $\mathcal{A}$  containing the identity matrix  $\mathbf{I}_n$  as well as all the matrices in  $\mathcal{A}$ . We will say that a Jordan algebra containing the identity matrix is complete. We now establish

**Proposition 6.** *If  $\{\mathbf{Q}_1, \dots, \mathbf{Q}_v\}$  is the principal basis of a complete Jordan algebra we have*

$$\det \left( \sum_{j=1}^v c_j \mathbf{Q}_j \right) = \prod_{j=1}^v c_j^{g_j}$$

with  $g_j = \text{rank}(\nabla_j)$ ,  $j = 1, \dots, v$ .

**Proof.** With  $\bar{g}_0 = 0$  and

$$\bar{g}_j = \sum_{l=1}^j g_l, \quad j = 1, \dots, v, \quad \text{let } \left\{ \alpha_{\bar{g}_{j-1}+1}^n, \dots, \alpha_{\bar{g}_j}^n \right\}$$

be an orthonormal basis for  $\nabla_j$ ,  $j = 1, \dots, v$ . Since these spaces are mutually orthogonal and  $\bar{g}_v = n$ ,  $\{\alpha_1^n, \dots, \alpha_n^n\}$  will be an orthonormal basis for  $\mathbb{R}^n$ . The vectors in this basis will be the row vectors of an orthogonal matrix  $\mathbf{P}$  such that  $\mathbf{P}\mathbf{Q}_j\mathbf{P}' = \mathbf{D}_j$ , with  $\mathbf{D}_j$  the diagonal matrix whose non null principal elements have indexes between  $\bar{g}_{j-1} + 1$  and  $\bar{g}_j$  and are equal to 1. Thus

$$\det \left( \sum_{j=1}^v c_j \mathbf{Q}_j \right) = \det \left( \mathbf{P} \left( \sum_{j=1}^v c_j \mathbf{Q}_j \right) \mathbf{P}' \right)$$

and the thesis follows from  $\sum_{j=1}^v c_j \mathbf{D}_j$  being the matrix whose principal elements with indexes between  $\bar{g}_{j-1} + 1$  and  $\bar{g}_j$  are equal to  $c_j$ . ■

#### REFERENCES

- [1] M. Fonseca, J.T. Mexia and R. Zmyślony, *Estimating and testing of variance components: an application to a grapevine experiment*, Biometrical Letters **40** (1) (2003), 1–7.
- [2] M. Fonseca, J.T. Mexia and R. Zmyślony, *Exact distribution for the generalized F tests*, Discuss. Math.–Probability and Statistics **1,2** (2002), 37–51.
- [3] A.I. Khuri, *Direct Products: A powerful tool for the analysis of balanced data*, Commun. Stat. Theory, **11** (1977), 2903–2929.
- [4] A.I. Khuri, T. Matthew and B.K. Sinha, *Statistical Tests for Mixed Linear Models*, Wiley 1997.
- [5] M. Loeve, *Probability Theory*, D. van Nostrand Company 1963.
- [6] J. Seely and G. Zyskind, *Linear spaces and minimum variance unbiased estimation*, Ann. Math. Stat. **42** (1971), 691–703.
- [7] J. Seely and G. Zyskind, *Quadratic subspaces and completeness*, Ann. Math. Stat. **42** (1971), 710–721.
- [8] J. Seely, *Minimal sufficient statistics and completeness for multivariate normal families*, Sankhyā, Ser. (A) **39** (1977), 170–185.
- [9] T.A. Severini, *Likelihood Methods in Statistics*, Oxford University Press 2000.

- [10] W.H. Steeb, *Kronecker Product and Applications*, Manheim 1991.
- [11] R. Zmyślony, *Completeness for a family of normal distributions*, Mathematical Statistics, Banach Cent. Publ. **6** (1980), 355–357.
- [12] A. Michalski and R. Zmyślony, *Testing hypotheses for variance components in mixed linear models*, Statistics (3-4) **27** (1996), 297–310.
- [13] R. Zmyślony and S. Zontek, *On robust estimation of variance components via von Mises functionals*, Discuss. Math.–Algebra and Stoch. Methods, **15** (2) (1995), 349–362.
- [14] A. Michalski and R. Zmyślony, *Testing hypotheses for linear functions of parameters in mixed linear models*, Tatra Mt. Math. Publ. **17** (1999), 103–110.
- [15] S. Zontek, *Robust estimation in linear models to spatially located sensors and random input*, Tatra Mountain Publications **17** (1999), 301–310.

Received 17 October 2003  
Revised 17 December 2003