



MICROCOPY RESOLUTION TEST CHART NATIONAL BUREAU UN STANDARDS-1963-A

OTIC FILE COPY

٤.

Exact Tests for Variance Component Models With Unequal Cell Frequencies in the Last Stage

2

by

A. I. Khuri Department of Statistics University of Florida Gainesville, FL 32611



**88 2** 05 v10

Technical Report Number 296 January 1988

PREPARED UNDER GRANT NO. N00014-86-K-0059 FROM THE OFFICE OF NAVAL RESEARCH. ANDRÉ I. KHURI AND RAMON C. LITTELL, PRINCIPAL INVESTIGATORS

DESTRIBUTION STATEMENT A

Approved for public releases Distribution Unitation

REPORT DOCUMENTATION P	AGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
REPORT NUMBER	GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
007		
		S TYPE OF REPORT & REPORT COVERED
Exact Tests for Variance Component	Models with	
nequal Cell Frequencies in the Las	it Stage	6. PERFORMING ORG. REPORT NUMBER
AUTHOR(+)		S. CONTRACT OR GRANT NUMBER(s)
		N00014-86-K-0059
A. I. Khuri		K & T 411455201
		ACCE. NO. 49101623439
PERFORMING ORGANIZATION NAME AND ADDRESS		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
96 Little Hall. University of Flor	ida	
ainesville. FL 32611		
CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE
Office of News] Becamet		
Mathematical Sciences Division		IS. NUMBER OF PAGES
Arlington, VA 22217-5000		21
MONITORING AGENCY NAME & ADDRESS(I different f	rom Controlling Office)	18. SECURITY CLASS. (of this report)
		Unclassified
		134. DECLASSIFICATION/ DOWNGRADING SCHEDULE
DISTRIBUTION STATEMENT (of this Report)	ibution Unlimit	ed
DISTRIBUTION STATEMENT (of this Report) Approved for Public Release: Distr DISTRIBUTION STATEMENT (of the abouract entitied in	Block 29, 11 different fro	ed
DISTRIBUTION STATEMENT (of this Report) Approved for Public Release: Distr DISTRIBUTION STATEMENT (of the ebetract entitied in	Block 20, 11 different fro	ed m Report)
DISTRIBUTION STATEMENT (of this Report) Approved for Public Release: Distr DISTRIBUTION STATEMENT (of the abouract entered in	Block 20, 11 different fro	ed Report)
DISTRIBUTION STATEMENT (of this Report) Approved for Public Release: Distr DISTRIBUTION STATEMENT (of the abouract enroyed in	Dieck 20, 11 different fro	red m Report)
DISTRIBUTION STATEMENT (of this Report) Approved for Public Release: Distr DISTRIBUTION STATEMENT (of the ebetract entered in	Block 20, 11 different fro	ed (m Report)
DISTRIBUTION STATEMENT (of this Report) Approved for Public Release: Distr DISTRIBUTION STATEMENT (of the aboutact entered in SUPPLEMENTARY NOTES	Diock 20, 11 dillorent fro	ed Report)
DISTRIBUTION STATEMENT (of this Report) Approved for Public Release: Distr DISTRIBUTION STATEMENT (of the abotract antified in SUPPLEMENTARY NOTES	Diock 20, 11 different fro	red Report)
DISTRIBUTION STATEMENT (of this Report) Approved for Public Release: Distr DISTRIBUTION STATEMENT (of the abouract antered in SUPPLEMENTARY NOTES	Dieck 20, 11 different fro	red m Report)
DISTRIBUTION STATEMENT (of this Report) Approved for Public Release: Distr DISTRIBUTION STATEMENT (of the abouract entered in SUPPLEMENTARY NOTES	Diock 20, 11 dillorent fro	red m Report)
DISTRIBUTION STATEMENT (of this Report) Approved for Public Release: Distr DISTRIBUTION STATEMENT (of the abotract entered in SUPPLEMENTARY NOTES KEY WORDS (Continue on reverse side if necessary and i	ibution Unlimit	red Report)
DISTRIBUTION STATEMENT (of this Report) Approved for Public Release: Distr DISTRIBUTION STATEMENT (of the obstract entered in SUPPLEMENTARY HOTES KEY WORDS (Continue on reverse side if necessary and i	ibution Unlimit	Report)
DISTRIBUTION STATEMENT (of this Report) Approved for Public Release: Distr DISTRIBUTION STATEMENT (of the ebetract entered in SUPPLEMENTARY NOTES KEY WORKS (Continue on reverse eide if necessary and f Jubalanced random models; Hypothesi	Dieck 20, 11 different fro	Report)
DISTRIBUTION STATEMENT (of this Report) Approved for Public Release: Distr DISTRIBUTION STATEMENT (of the abotract entered in SUPPLEMENTARY NOTES KEY WORKS (Continue on reverse eide if necessary and i Unbalanced random models; Hypothesi and crossed classification models.	ibution Unlimit Block 20, 11 different fro Mentify by block number) Is testing; Powe	red Report) er-of exact tests; Nested
DISTRIBUTION STATEMENT (of this Report) Approved for Public Release: Distr DISTRIBUTION STATEMENT (of the aborract entered in SUPPLEMENTARY NOTES KEY WORKS (Continue on reverse elde II necessary and I Inbalanced random models; Hypothesi and crossed classification models.	Diock 20, 11 different fro	Report)
DISTRIBUTION STATEMENT (of this Report) Approved for Public Release: Distr DISTRIBUTION STATEMENT (of the observer envored in SUPPLEMENTARY NOTES REY WORKS (Continue on reverse side if necessary and i Unbalanced random models; Hypothesi and crossed classification models.	Diock 20, 11 different fro Monthly by block number) Is testing; Power Conting by block number)	r Report)
DISTRIBUTION STATEMENT (of this Report) Approved for Public Release: Distr DISTRIBUTION STATEMENT (of the obsirect entered in SUPPLEMENTARY NOTES KEY WORKS Continue on reverse side if necessary and i Inbalanced random models; Hypothesi and crossed classification models.	dentify by block number)	Report)
DISTRIBUTION STATEMENT (of this Report) Approved for Public Release: Distr DISTRIBUTION STATEMENT (of the abstract entered in SUPPLEMENTARY NOTES KEY WORKS Continue on reverse side if necessary and i Inbalanced random models; Hypothesi and crossed classification models.	dentify by block number) struct exact te	r Report) er of exact tests; Nested sts for the variance com-
DISTRIBUTION STATEMENT (of this Report) Approved for Public Release: Distr DISTRIBUTION STATEMENT (of the abetract entered in SUPPLEMENTARY NOTES KEY WORKS (Continue on reverse elde II necessary and I Inbalanced random models; Hypothesi and crossed classification models. ABSTRACT (Continue on reverse elde II necessary and IC Chis paper presents a method to cont bonents in a general unbalanced ran	dentity by block number) struct exact tendom model. It	red Report) ar of exact tests; Nested sts for the variance com- is assumed that the im-
DISTRIBUTION STATEMENT (of this Report) Approved for Public Release: Distr DISTRIBUTION STATEMENT (of the observer entered in SUPPLEMENTARY NOTES KEY WORKS Continue on reverse side if necessary and i Unbalanced random models; Hypothesi and crossed classification models. ABSTRACT (Continue on reverse side if necessary and is conents in a general unbalanced ran balance affects only the last stage	dentify by block number) struct exact te ndom model. It e of the design	sts for the variance com- is assumed that the im- with no missing cells.
DISTRIBUTION STATEMENT (of this Report) Approved for Public Release: Distr DISTRIBUTION STATEMENT (of the observer entered in SUPPLEMENTARY NOTES KEY WORKS (Continue on reverse side if necessary and i Inbalanced random models; Hypothesi and crossed classification models. ABSTRACT (Continue on reverse side if necessary and is chis paper presents a method to com- bonents in a general unbalanced ran colance affects only the last stage	Diock 20, if different fro Block 20, if different fro identify by block number) Is testing; Power dentify by block number) struct exact te ndom model. It a of the design	sts for the variance com- is assumed that the im- with no missing cells.
DISTRIBUTION STATEMENT (of this Report) Approved for Public Release: Distr DISTRIBUTION STATEMENT (of the ebetrect entired in SUPPLEMENTARY NOTES KEY WORKS Continue on reverse eide if necessary and i Inbalanced random models; Hypothesi and crossed classification models. ABSTRACT (Continue on reverse eide if necessary and ic Distriction on the last stage	ibution Unlimit Block 20, 11 different fro dentify by block number) ts testing; Power dentify by block number) struct exact te ndom model. It e of the design	red Report) ex-of exact tests; Nested, sts for the variance com- is assumed that the im- with no missing cells. Keyever
DISTRIBUTION STATEMENT (of this Report) Approved for Public Release: Distr DISTRIBUTION STATEMENT (of the observed entered in SUPPLEMENTARY NOTES KEY WORKS (Continue on reverse elde if necessary and i Inbalanced random models; Hypothesi and crossed classification models. ABSTRACT (Continue on reverse elde if necessary and is chis paper presents a method to com- conents in a general unbalanced ran calance affects only the last stage	dentify by block number) struct exact tendom model. It e of the design	er of exact tests; Nested, sts for the variance com- is assumed that the im- with no missing cells. Weyever

.

80000000

# EXACT TESTS FOR VARIANCE COMPONENT MODELS WITH UNEQUAL CELL FREQUENCIES IN THE LAST STAGE

André I. KHURI

Department of Statistics, University of Florida, Gainesville, FL 32611, USA

Abstract: Khuri and Littell (1987) derived exact tests for testing hypotheses concerning the variance components in an unbalanced random two-way model. The method used in the development of these tests can be extended to more general unbalanced random models. In this article, such an extension is established for models that are unbalanced only with respect to the last stage of their associated designs. A numerical example is given to illustrate the implementation of the proposed methodology.

AMS Subject Classification: Primary 62J10; Secondary 62F03.

Key words and phrases: Unbalanced random models; Hypothesis testing; Power of exact tests; Nested and crossed classification models.

#### 1. Introduction

The traditional analysis of data from an unbalanced random model uses approximate F tests that are based on sums of squares, which, in general, are neither independent nor distributed as scaled chi-square variates. The true critical values and power functions of these tests are unknown and depend on variance components other than those under consideration. Furthermore, the tests, particularly those that depend on Satterthwaite's procedure, can be quite unreliable. This was demonstrated by Tietjen (1974), in the case of the unbalanced random two-fold nested model, and by Cummings and Gaylor (1974).

-1-

Avail and/or Special

Ulst

The method developed by Khuri and Littell (1987) produced exact tests for the variance components in an unbalanced random two-way classification with interaction model. This method is based on a particular transformation that reduces the analysis of the unbalanced model to that of a balanced one. In this article, a demonstration is given of the applicability of the same kind of transformation to any unbalanced random model provided that the imbalance occurs only in the last stage of the associated design.

#### 2. Notation and preliminaries

The model for a general unbalanced design whose imbalance is caused by unequal cell frequencies in the last stage can be written in the form

$$\mathbf{y}_{\theta} = \sum_{i=0}^{\nu} \gamma_{\theta_{i}(\bar{\theta}_{i})} + \epsilon_{\theta}, \qquad (2.1)$$

where  $\theta = (k_1, k_2, ..., k_s)$  is a complete set of subscripts that identify a typical response y. The term,  $\gamma_{\theta_i(\bar{\theta}_i)}$ , denotes the i<sup>th</sup> effect in the model, where  $\bar{\theta}_i$  and  $\theta_i$  are, respectively, the corresponding sets of rightmost and nonrightmost bracket subscripts (see Section 2 in Khuri (1982)). By definition,  $\theta_i = \bar{\theta}_i = \phi$ , the empty set, for i=0 and the corresponding  $\gamma$  is the grand mean, usually denoted by  $\mu$ . It is assumed that  $\gamma_{\theta_i(\bar{\theta}_i)}$ , for i=1,2,..., $\nu$ , and  $\epsilon_{\theta}$  are independent and normally distributed random variables with zero means and variances  $\sigma_1^2, \sigma_2^2, ..., \sigma_{\nu}^2$ , and  $\sigma_{\epsilon}^2$ , respectively.

Since the design is balanced except for its last stage, the ranges of subscripts  $k_1, k_2, ..., k_s$  can be expressed in the form

$$k_{j} = \frac{1, 2, ..., a_{j}}{1, 2, ..., n_{\tau}}, \quad \text{for } j=1,2,...,s-1$$
(2.2)

where  $\tau$  is a subset of  $\theta$  consisting of the first s-1 subscripts, that is,

$$\tau = (\mathbf{k}_1, \, \mathbf{k}_2, \, \dots, \, \mathbf{k}_{s-1}). \tag{2.3}$$

Except for  $\epsilon_{\theta}$ , all the effects on the right side of (2.1) are indexed by the subscripts in  $\tau$ . Let T be the set of all (s-1)-tuples as in (2.3), and let c denote the number of elements in T, that is,

$$\mathbf{T} = \left\{ \tau = (\mathbf{k}_1, \mathbf{k}_2, ..., \mathbf{k}_{s-1}) : \mathbf{k}_j = 1, 2, ..., \mathbf{a}_j; \quad j = 1, 2, ..., s-1 \right\}$$
(2.4)

$$\mathbf{c} = \prod_{j=1}^{s-1} \mathbf{a}_j. \tag{2.5}$$

It is assumed that  $n_{\tau} \geq 1$  for all  $\tau \epsilon T$  and, for reasons to be seen later, that

$$N > 2c-1,$$
 (2.6)

where  $N = \sum_{\tau \in T} n_{\tau}$  is the total number of observations. Let  $\psi_i = (\theta_i : \overline{\theta}_i)$  be the set of subscripts associated with the i<sup>th</sup> effect, which results from combining the elements of  $\theta_i$  and  $\overline{\theta}_i$  (i=0,1,..., $\nu$ ). The complement of  $\psi_i$  with respect to  $\tau$  is denoted by  $\psi_i^c$  (i=0,1,..., $\nu$ ).

We note that if it were not for  $\epsilon_{\theta}$ , the model in (2.1) would be of the same form as that of a balanced model (see Khuri (1982) for a general representation of a balanced model). This fact will be quite useful in the development of the exact tests in Section 3.

The model in (2.1) can be written in matrix form as

$$\underline{y} = \sum_{i=0}^{\nu} \underline{X}_i \underline{\beta}_i + \underline{\epsilon}, \qquad (2.7)$$

where y and  $\underline{\epsilon}$  are the vectors of observations and random errors, respectively,  $\underline{X}_{i}$  is a matrix of zeros

and ones of order  $N \times c_i$   $(i=0,1,...,\nu)$  with  $X_0$  being equal to  $1_N$ , the vector of ones of order  $N \times 1$ , and  $\beta_i$  is a vector consisting of the  $c_i$  elements of  $\gamma_{\theta_i(\bar{\theta}_i)}$ . The integer  $c_i$  is given by

$$c_{i} = \prod_{k_{j} \in \psi_{i}} a_{j}, \quad i = 0, 1, ..., \nu,$$
 (2.8)

where, if we recall,  $\psi_i$  is the set of subscripts associated with the i<sup>th</sup> effect. Note that  $c_i = 1$  for i = 0.

Let  $\bar{y}_{\tau}$  denote the mean of  $y_{\theta}$  averaged over the range of subscript  $k_s$  for a given  $\tau = (k_1, k_2, ..., k_{s-1})$  in T, that is,

$$\bar{\mathbf{y}}_{\tau} = \frac{1}{\bar{\mathbf{n}}_{\tau}} \sum_{\mathbf{k_s}=1}^{\bar{\mathbf{n}}_{\tau}} \mathbf{y}_{\theta}, \quad \tau \in \mathbf{T}.$$
(2.9)

From (2.1) we have

$$\bar{\mathbf{y}}_{\tau} = \sum_{i=0}^{\nu} \gamma_{\theta_i}(\bar{\theta}_i) + \bar{\epsilon}_{\tau}, \qquad (2.10)$$

where  $\bar{\epsilon}_{\tau} = \left(\sum_{k_s=1}^{n_{\tau}} \epsilon_{\theta}\right)/n_{\tau}$ . Formula (2.10) may be written in matrix form as

$$\bar{\mathbf{y}} = \sum_{i=0}^{\nu} \mathbf{H}_{i} \hat{\boldsymbol{\beta}}_{i} + \bar{\boldsymbol{\xi}}, \qquad (2.11)$$

where  $\underline{H}_{i}$  is a matrix of order  $c \times c_{i}$ . Since  $\sum_{i=0}^{\nu} \gamma_{\theta_{i}}(\overline{\theta}_{i})$  in (2.10) is of the same form as in a balanced model, the matrix  $\underline{H}_{i}$  can be expressed as a direct product of the form (see Khuri (1982), p. 2908)

$$\mathbf{\tilde{H}}_{i} = \bigotimes_{\ell=1}^{s-1} \tilde{\mathbf{L}}_{i\ell}, \quad i = 0, 1, ..., \nu,$$
(2.12)

where  $L_{i\ell}$  is given by

$$\mathbf{L}_{i\ell} = \frac{\mathbf{l}_{a\ell}}{\mathbf{l}_{a\ell}}, \quad \begin{array}{l} \mathbf{k}_{\ell} \in \psi_{i} \\ \mathbf{k}_{\ell} \in \psi_{i}^{c} \\ \mathbf{k}_{\ell} \in \psi_{i}^{c} \end{array} \quad i = 0, 1, \dots, \nu; \\ \ell = 1, 2, \dots, s-1. \end{array}$$
(2.13)

In (2.13),  $I_{a_{\ell}}$  and  $I_{a_{\ell}}$  denote, respectively, the identity matrix of order  $a_{\ell} \times a_{\ell}$  and the vector of ones of order  $a_{\ell} \times 1$ . We recall that  $\psi_i^c$  in (2.13) is the complement of  $\psi_i$  with respect to  $\tau$ . Let  $A_i = H_i H_i'$ , then

$$\underline{A}_{i} = \bigotimes_{\ell=1}^{s-1} \underline{M}_{i\ell}, \quad i = 0, 1, ..., \nu,$$
(2.14)

where

$$\underline{M}_{i\ell} = \frac{\mathbf{I}_{a\ell}}{\mathbf{J}_{a\ell}}, \quad \begin{array}{l} \mathbf{k}_{\ell} \in \psi_{i} \\ \mathbf{J}_{a\ell}, \quad \mathbf{k}_{\ell} \in \psi_{i}^{c} \\ \ell = 1, 2, \dots, s-1. \end{array}$$
(2.15)

In (2.15),  $J_{a_{\ell}}$  denotes the matrix of ones of order  $a_{\ell} \times a_{\ell}$ .

The variance-covariance matrix of  $\bar{y}$  in (2.11) can now be written as

$$\operatorname{Var}(\bar{\mathbf{y}}) = \sum_{i=1}^{\nu} \sigma_i^2 \mathbf{A}_i + \sigma_{\epsilon}^2 \mathbf{K}, \qquad (2.16)$$

where

$$\tilde{\mathbf{K}} = \operatorname{diag}\left(\frac{1}{n_{\tau}}\right)_{\tau \in \mathbf{T}}.$$
(2.17)

The right side of (2.17) denotes a diagonal matrix of order c×c. Its diagonal elements are the reciprocals of the  $n_{\tau}$ 's for  $\tau \epsilon T$ , the set of all values of  $\tau$  in (2.3). It can be verified that  $A_i A_{i'} = A_{i'} A_i$  for  $i \neq i'$ . It follows that there exists an orthogonal matrix, Q, of order c×c such that

$$QA_iQ' = A_i, \quad i = 0, 1, ..., \nu,$$
 (2.18)

where  $\Lambda_i$  is a diagonal matrix. The construction of the matrix Q will be described in Section 3.

#### 3. The development of the exact tests

Let us again consider model (2.10). As was noted earlier in Section 2, the first part to the right of this model, namely,  $\sum_{i=0}^{\nu} \gamma_{\theta_i}(\bar{\theta}_i)$ , is not affected by the imbalance in the last stage of the design, that is, in the  $\tau$ -cells, where  $\tau$  is described in (2.3). Let us therefore consider the derived model

$$\mathbf{z}_{\tau} = \sum_{i=0}^{\nu} \gamma_{\theta_i}(\bar{\theta}_i). \tag{3.1}$$

Model (3.1) is balanced with one observation in each  $\tau$ -cell. Let  $\underline{P}_i$  be a c×c matrix associated with the sum of squares for the i<sup>th</sup> effect (i=0,1,..., $\nu$ ). From Khuri (1982) we have the following lemmas: Lemma 3.1.

- (i)  $P_i$  is idempotent  $(i=0,1,\ldots,\nu)$ .
- (ii)  $P_i P_{i'} = 0$  for  $i \neq i'$ . (iii)  $\sum_{i=0}^{\nu} P_i = I_c$ .

Lemma 3.2. The matrix  $\underline{P}_i$  can be expressed in terms of the  $\underline{A}_i$ 's in (2.14) as

$$P_{i} = \sum_{j=0}^{\nu} \left( \lambda_{ij} / b_{j} \right) A_{j}, \quad i = 0, 1, ..., \nu,$$
(3.2)

where  $\lambda_{ij}$  is a known constant equal to the coefficient of the j<sup>th</sup> admissible mean in the i<sup>th</sup> component for the balanced model in (3.1) (possible values of  $\lambda_{ij}$  are -1, 0, and 1), and b<sub>j</sub> is given by

$$b_{j} = \prod_{\substack{k_{\ell} \in \psi_{j}^{c}}} a_{\ell}, \quad j = 0, 1, ..., \nu,$$
(3.3)

where  $v_j^c$  is the complement of  $\psi_j$  with respect to  $\tau$ .

Lemma 3.3. The  $A_i$  and  $P_i$  matrices in (3.2) are related by the following identity:

$$A_{j}P_{i} = \kappa_{ij}P_{i}, \quad i = 0, 1, ..., \nu;$$
  
 $j = 0, 1, ..., \nu,$  (3.4)

where  $\kappa_{ij}$  is given by

$$\kappa_{ij} = \begin{matrix} 0, & \psi_i \not\subset \psi_j \\ b_j, & \psi_i \subset \psi_j, \end{matrix}$$
(3.5)

and  $\psi_i$  is the set of subscripts associated with the i<sup>th</sup> effect in (3.1), i = 0, 1, ...,  $\nu$ .

Let  $m_i$  be the rank of  $P_i$   $(i=0,1,...,\nu)$ , then by Lemma 3.1,  $\sum_{i=0}^{\nu} m_i = c$ . Let  $Q_i$  be a matrix of order  $m_i \times c$  and rank  $m_i$  whose rows are orthonormal and span the row space of  $P_i$   $(i=0,1,...,\nu)$ . The matrix  $Q_i$  can be easily obtained as the result of a Gram-Schmidt orthonormal factorization of the rows (or columns) of  $P_i$ . The proof of the following lemma is given in Appendix A: Lemma 3.4. The matrices  $Q_0, Q_1, ..., Q_{\nu}$  have the following properties:

(i)  $Q_0 = \frac{1}{c} \sqrt{c}$ .

(ii) 
$$\begin{array}{ll} Q_{i} \ Q_{i}' = I_{m_{i}}, & i = 0, 1, ..., \nu\\ Q_{i} \ Q_{i}' = 0, & i \neq i'. \end{array}$$

(iii) 
$$A_j Q'_i = \begin{array}{ccc} 0, & \psi_i \not\subset \psi_j & i = 0, 1, \dots, \nu; \\ b_j Q'_i, & \psi_i \subset \psi_j & j = 0, 1, \dots, \nu. \end{array}$$

Let us now define the matrix

$$\mathbf{Q} = \left[\mathbf{Q}_0': \mathbf{Q}_1': \cdots: \mathbf{Q}_\nu'\right]', \tag{3.6}$$

which is of order cxc. By Lemma 3.4, Q is an orthogonal matrix and diagonalizes  $A_0$ ,  $\Lambda_1$ , ...,  $A_{\nu}$  simultaneously as in (2.18). From (2.16) we have

$$\operatorname{Var}\left(\underline{Q}_{i} \ \overline{y}\right) = \underline{Q}_{i}\left(\sum_{j=1}^{\nu} \sigma_{j}^{2} \underline{A}_{j} + \sigma_{\epsilon}^{2} \underline{K}\right) \underline{Q}_{i}^{\prime}, \quad i = 1, 2, ..., \nu,$$

$$= \sum_{j=1}^{\nu} \sigma_{j}^{2} (Q_{i} A_{j} Q_{i}') + \sigma_{\epsilon}^{2} (Q_{i} K Q_{i}'), \quad i = 1, 2, ..., \nu,$$
(3.7)

where  $\bar{y}$  is the vector of all  $\tau$ -cell means defined in (2.11). But, by Lemma 3.4 ((ii), (iii))

 $\sum_{j=1}^{\nu} \sigma_{j}^{2} \left( Q_{i} A_{j} Q_{i}^{\prime} \right) = \delta_{i} I_{m_{i}}, \quad i = 1, 2, ..., \nu,$ (3.8)

where

$$\delta_{i} = \sum_{j \in W_{i}} b_{j} \sigma_{j}^{2}.$$
(3.9)

In (3.9),  $W_i$  is the set

$$W_{i} = \left\{ j : \psi_{i} \subset \psi_{j}, \quad 1 \leq j \leq \nu \right\}.$$

From (3.7) and (3.8) we obtain

$$\operatorname{Var}\left(\underline{Q}_{i}\overline{\underline{y}}\right) = \delta_{i}\underline{I}_{\mathbf{m}_{i}} + \sigma_{\epsilon}^{2}\left(\underline{Q}_{i}\underline{K}\underline{Q}_{i}'\right), \quad i = 1, 2, ..., \nu.$$
(3.10)

Furthermore, for  $i \neq i'$ ,

$$\operatorname{Cov}\left(\underline{Q}_{i} \underline{\bar{y}}, \underline{\bar{y}}' \underline{Q}_{i'}'\right) = \underline{Q}_{i} \left(\sum_{j=1}^{\nu} \sigma_{j}^{2} \underline{A}_{j} + \sigma_{\epsilon}^{2} \underline{K}\right) \underline{Q}_{i'}'$$
$$= \sigma_{\epsilon}^{2} \left(\underline{Q}_{i} \underline{K} \underline{Q}_{i'}'\right). \tag{3.11}$$

Let u be a vector of order (c-1)×1 defined as

$$\underline{\mathbf{u}} = \left[\underline{\mathbf{Q}}_1' : \underline{\mathbf{Q}}_2' : \cdots : \underline{\mathbf{Q}}_{\nu}'\right]' \overline{\underline{\mathbf{y}}}.$$
(3.12)

By (3.10) and (3.11), the variance-covariance matrix of u is of the form

$$\operatorname{Var}(\underline{u}) = \operatorname{diag}\left(\delta_{1}\underline{I}_{m_{1}}, \delta_{2}\underline{I}_{m_{2}}, \dots, \delta_{\nu}\underline{I}_{m_{\nu}}\right) + \sigma_{\epsilon}^{2}\underline{G}, \qquad (3.13)$$

where

$$\tilde{\mathbf{G}} = \tilde{\mathbf{Q}} \tilde{\mathbf{K}} \tilde{\mathbf{Q}}', \qquad (3.14)$$

and  $\tilde{\mathbf{Q}} = \begin{bmatrix} \mathbf{Q}_1' : \mathbf{Q}_2' : \cdots : \mathbf{Q}_{\nu}' \end{bmatrix}'$ .

Now, the vector  $\bar{y}$  can be expressed as  $\bar{y} = \bar{D}y$ , where y is the vector of observations in (2.7) and  $\bar{D}$  is the direct sum

$$\bar{\mathbf{D}} = \mathop{\oplus}_{\tau \in \mathbf{T}} \left( \frac{\mathbf{l}'_n}{\tau} / \mathbf{n}_{\tau} \right), \tag{3.15}$$

where  $\tau$  is defined in (2.3) and T is the set of all values of  $\tau$ . The matrix D is of order  $c \times N$ , where N is the total number of observations. Also, the residual sum of squares for the original unbalanced model in (2.1) can be written as

$$SS_{E} = \sum_{\theta} (y_{\theta} - \bar{y}_{\tau})^{2}$$
$$= \sum_{\tau \in T} \left[ \sum_{k_{s}=1}^{n_{\tau}} (y_{\theta} - \bar{y}_{\tau})^{2} \right], \qquad (3.16)$$

where, if we recall,  $\theta = (\tau, k_s)$ . Formula (3.16) can be rewritten as

$$SS_{E} = \underline{y}' \underline{R} \underline{y}, \qquad (3.17)$$

where

 $\tilde{\mathbf{R}} = \bigoplus_{\tau \in T} \left( \underline{\mathbf{I}}_{n_{\tau}} - \underline{\mathbf{J}}_{n_{\tau}} / n_{\tau} \right).$ (3.18)

Lemma 3.5.

- (i) R is idempotent of rank N-c.
- (ii)  $\mathbf{D}\mathbf{R} = \mathbf{0}$ .
- (iii)  $RX_i = 0$ ,  $i = 1, 2, ..., \nu$ ,

where  $X_1, X_2, ..., X_{\nu}$  are the matrices given in (2.7).

Proof. See Appendix B.

From Lemma 3.5 we conclude that

$$\mathbf{D}\Sigma\mathbf{R} = \mathbf{0},\tag{3.19}$$

where  $\Sigma$  is the variance-covariance matrix of y in (2.7), which is equal to

$$\Sigma = \sum_{i=1}^{\nu} \sigma_i^2 X_i X_i' + \sigma_{\epsilon}^2 I_N.$$
(3.20)

Furthermore,

$$\mathbf{R}\Sigma/\sigma_{\epsilon}^2 = \mathbf{R}.\tag{3.21}$$

It follows from (3.19) and (3.21) that  $\bar{y} = Dy$  is independent of SS<sub>E</sub> and that SS<sub>E</sub>/ $\sigma_{\epsilon}^2$  has the chisquare distribution with N-c degrees of freedom, the rank of R.

Since R is idempotent of rank N-c, then it can be written as

$$\mathbf{R} = \mathbf{C} \mathbf{A} \mathbf{C}', \tag{3.22}$$

where C is an orthogonal matrix and  $\Lambda$  is a diagonal matrix with N-c ones and c zeros. By the assumption in (2.6),  $\Lambda$  can be partitioned as

$$\underline{\Lambda} = \operatorname{diag}(\underline{I}_{\xi_1}, \underline{I}_{\xi_2}, \underline{0}), \tag{3.23}$$

where

$$\begin{aligned} \xi_1 &= c - 1 \\ \xi_2 &= N - 2c + 1 > 0, \end{aligned} \tag{3.24}$$

-10-

and 0 in (3.23) is a zero matrix of order c×c. Accordingly, the matrix C in (3.22) can be partitioned as  $C = [C_1 : C_2 : C_3]$ , where  $C_1, C_2$ , and  $C_3$  are of orders N× $\xi_1$ , N× $\xi_2$ , and N×c, respectively. We then have

$$\mathbf{R} = \mathbf{C}_1 \mathbf{C}_1' + \mathbf{C}_2 \mathbf{C}_2'. \tag{3.25}$$

Note that

Let us now define the random vector  $\boldsymbol{\omega}$  as

$$\omega = \underline{u} + \left(\lambda_{\max}\underline{I}_{\xi_1} - \underline{G}\right)^{\frac{1}{2}}\underline{C}_1'\underline{y}, \qquad (3.27)$$

where G is the matrix in (3.14) and  $\lambda_{\max}$  is its largest eigenvalue,  $\left(\lambda_{\max}I_{\xi_1} - G\right)^{\frac{1}{2}}$  is a symmetric matrix with eigenvalues equal to the square roots of the eigenvalues of  $\lambda_{\max}I_{\xi_1} - G$ , which are nonnegative. Let  $\omega$  be partitioned just like u in (3.12) as

$$\boldsymbol{\omega} = \left(\boldsymbol{\omega}_1', \, \boldsymbol{\omega}_2', \, \dots, \, \boldsymbol{\omega}_\nu'\right)',\tag{3.28}$$

where  $\omega_i$  is of order  $m_i \times 1$  and  $m_i$  is the number of rows  $Q_i$  (i=1,2,..., $\nu$ ) in (3.12). The distributional properties of the  $\omega_i$ 's are given in the following lemma:

Lemma 3.6.

- (i)  $E(\omega_i) = 0$ ,  $i = 1, 2, ..., \nu$ .
- (ii)  $\omega_1, \omega_2, ..., \omega_{\nu}$  are independently distributed as normal random vectors with  $\omega_i$  having the variance-covariance matrix

-11-

$$\operatorname{Var}(\omega_{i}) = \left(\delta_{i} + \lambda_{\max}\sigma_{\epsilon}^{2}\right) I_{m_{i}}, \quad i = 1, 2, \dots, \nu, \qquad (3.29)$$

where  $\delta_i$  is defined in (3.9).

(iii)  $\psi_1, \psi_2, ..., \psi_{\nu}$  are independent of  $SS_E^{(2)}$ , where  $SS_E^{(2)} = y' C_2 C_2' y$  is the portion of the residual sum of squares,  $SS_E$ , in (3.17), which corresponds to the matrix  $C_2$  in (3.25).

Proof.

**•**1

(i) From (2.11) and Lemma 3.4,  $E(Q_i \overline{y}) = Q_i H_0 \beta_0 = Q_i l_c \beta_0 = 0$  for  $i = 1, 2, ..., \nu$ . Hence,  $E(\underline{u}) = \underline{0}$  by (3.12). From (3.18) we also have that  $\underline{R} \underline{l}_N = \underline{0}$ , which by (3.25) can be rewritten as

$$\left(\bar{\mathcal{C}}_{1}\bar{\mathcal{C}}_{1}'+\bar{\mathcal{C}}_{2}\bar{\mathcal{C}}_{2}'\right)\underline{\mathbf{i}}_{N}=\underline{\mathbf{0}}.$$
(3.30)

Using (3.26) in (3.30) we get  $\zeta'_{11N} = 0$ . It follows from (2.7) that  $E(\zeta'_{1y}) = \zeta'_{11N}\beta_0 = 0$ . The mean of  $\omega$  in (3.27) is therefore equal to zero.

(ii) It is clear that  $\omega$  is normally distributed. Now, the vector  $\mathbf{u}$  in (3.27) is independent of  $C_1' \mathbf{y}$ . To show this, we note from (3.19), (3.25), and (3.26) that

$$\mathbf{D}\boldsymbol{\Sigma}\mathbf{C}_1 = \boldsymbol{0}. \tag{3.31}$$

Hence.

$$\operatorname{Cov}\left(\bar{y}, \, y'C_1\right) = D\Sigma C_1 = 0,$$

since  $\bar{y} = \bar{D}y$ . It follows that  $\bar{y}$ , and hence u in (3.12), is independent of  $C'_{1}y$ .

The variance-covariance matrix of  $\omega$  can therefore be expressed as

$$\operatorname{Var}(\boldsymbol{\omega}) = \operatorname{Var}(\underline{\boldsymbol{u}}) + \left(\lambda_{\max}\underline{\boldsymbol{I}}_{\xi_{1}} - \underline{\boldsymbol{G}}\right)^{\frac{1}{2}} \underline{\boldsymbol{C}}_{1}' \underline{\boldsymbol{\Sigma}} \underline{\boldsymbol{C}}_{1} \left(\lambda_{\max}\underline{\boldsymbol{I}}_{\xi_{1}} - \underline{\boldsymbol{G}}\right)^{\frac{1}{2}}.$$
(3.32)

But, from (3.21) we have

$$\mathbf{R} \boldsymbol{\Sigma} \mathbf{R} = \sigma_{\epsilon}^2 \mathbf{R}. \tag{3.33}$$

Also, from (3.25), (3.26), and (3.33) it can be verified that

$$\underline{C}_1' \underline{\Sigma} \underline{C}_1 = \sigma_{\epsilon}^2 \underline{C}_1' \underline{C}_1 = \sigma_{\epsilon}^2 \underline{I}_{\xi_1}.$$
 (3.34)

From (3.32) and (3.34) we then have

$$\operatorname{Var}(\omega) = \operatorname{Var}(\underline{u}) + \sigma_{\epsilon}^{2} \left( \lambda_{\max} \underline{I}_{\xi_{1}} - \underline{G} \right).$$
(3.35)

By using (3.13) in (3.35) we get

$$\operatorname{Var}(\omega) = \operatorname{diag}\left(\delta_{1}\mathfrak{l}_{\mathfrak{m}_{1}}, \, \delta_{2}\mathfrak{l}_{\mathfrak{m}_{2}}, \, \dots, \, \delta_{\nu}\mathfrak{l}_{\mathfrak{m}_{\nu}}\right) + \lambda_{\max}\sigma_{\epsilon}^{2}\mathfrak{l}_{\xi_{1}}.$$
(3.36)

From (3.36) we conclude that  $\omega_1, \omega_2, ..., \omega_{\nu}$  are independent and that  $Var(\omega_i)$  has the form described in (3.29).

(iii)  $SS_E^{(2)}$  is independent of  $\underline{u}$  (since  $SS_E$  is independent of  $\underline{\overline{y}}$  by (3.19)) and is also independent of  $\underline{C}'_1 \underline{y}$ . This is true because

$$\mathcal{L}_1' \mathcal{\Sigma} \mathcal{L}_2 = 0, \tag{3.37}$$

which follows from (3.25), (3.26), and (3.33). Consequently,  $SS_E^{(2)}$  and  $\omega$  in (3.27) are independent.

From Lemma 3.6 we conclude that if  $SS_i = \omega_i' \omega_i$  (i=1,2,..., $\nu$ ), then the sums of squares,  $SS_1$ ,  $SS_2, ..., SS_{\nu}$  are independent and  $SS_i/(\delta_i + \lambda_{\max}\sigma_{\epsilon}^2)$  is distributed as a central chi-square variate with  $m_i$  degrees of freedom (i=1,2,..., $\nu$ ). Furthermore, the  $SS_i$ 's are independent of  $SS_E^{(2)}/\sigma_{\epsilon}^2$ , which -13has the chi-square distribution with  $\xi_2$  degrees of freedom, where  $\xi_2$  is defined in (3.24).

It follows that  $SS_1, SS_2, ..., SS_{\nu}$  and  $SS_E^{(2)}$  act like sums of squares in an ANOVA table for a balanced random model. In other words, the analysis concerning the variance components,  $\sigma_1^2, \sigma_2^2, ..., \sigma_{\nu}^2$ , can proceed using these sums of squares just like in a balanced data situation.

In particular, if the data set is balanced, that is, if  $n_{\tau} = n$ , then  $K = I_c/n$  (see (2.17)) and  $G = I_{\xi_1}/n$ , where G and  $\xi_1$  are defined in (3.14) and (3.24), respectively. Consequently,  $\lambda_{\max}$ , the largest eigenvalue of G, is equal to 1/n. The vectors  $\omega$  and  $\underline{u}$  in (3.27) are therefore identical. In this case,

$$= \bar{y}' (\underline{I}_{c} - \underline{J}_{c}/c) \bar{y} \quad (by \text{ Lemma 3.4})$$

$$= \bar{y}' \Big( \sum_{i=1}^{\nu} \bar{P}_i \Big) \bar{y} \qquad (by \text{ Lemma 3.1(iii) and the fact that } \bar{P}_0 = \bar{J}_c/c \text{ by formula (3.2)} \Big).$$

Thus,

187.57.23 197.197.23

$$\sum_{i=1}^{\nu} SS_i = \sum_{i=1}^{\nu} \overline{y}' \underline{P}_i \overline{y}.$$

Now, from (3.2) we can write

$$\bar{\mathbf{y}}' \mathbf{P}_{\mathbf{i}} \bar{\mathbf{y}} = \bar{\mathbf{y}}' \left[ \sum_{\mathbf{j}=0}^{\nu} \left( \lambda_{\mathbf{i}\mathbf{j}} / \mathbf{b}_{\mathbf{j}} \right) \bar{\mathbf{A}}_{\mathbf{j}} \right] \bar{\mathbf{y}}$$

$$= \bar{\mathbf{y}}' \left[ \sum_{\mathbf{j}=0}^{\nu} \left( \lambda_{\mathbf{i}\mathbf{j}} / \mathbf{b}_{\mathbf{j}} \right) \left( \sum_{\ell=1}^{s-1} \mathbf{M}_{\mathbf{j}\ell} \right) \right] \bar{\mathbf{y}}. \qquad (by (2.14))$$
(3.38)

Since  $\bar{y} = (\underline{I}_c \otimes \underline{I}'_n) \underline{y}/n$ , formula (3.38) can be expressed as



$$\overline{y}' \mathbf{P}_{i} \overline{y} = \frac{1}{n} \underline{y}' \left[ \sum_{j=0}^{\nu} \frac{\lambda_{ij}}{nb_{j}} \left\{ \begin{pmatrix} s-1 \\ \ell = 1 \end{pmatrix} M_{j\ell} \right\} \right] \underline{y}.$$

$$(3.39)$$

Formula (3.39) shows that  $n\overline{y}'\underline{P}_{i}\overline{y}$  is the usual sum of squares for the i<sup>th</sup> effect in a balanced model of the form given in (2.1). In other words,  $nSS_1$ ,  $nSS_2$ , ..., and  $nSS_{\nu}$  reduce to the usual balanced ANOVA sums of squares associated with the corresponding  $\nu$  effects whenever the data set is balanced.

#### 4. Power of the exact tests

Power values for the exact tests in Section 3 can be easily obtained just like in a balanced model situation. As in Khuri and Littell (1987), it is easy to show that such power values are monotone decreasing with respect to  $\lambda_{max}$ , the largest eigenvalue of the matrix G in (3.14). Upper and lower bounds on  $\lambda_{max}$  are given by the double inequality

$$\frac{1}{c} \sum_{\tau \in \mathbf{T}} \frac{1}{\mathbf{n}_{\tau}} \le \lambda_{\max} \le \frac{1}{\min(\mathbf{n}_{\tau})}, \tag{4.1}$$

where  $n_{\tau}$  is the frequency of the  $\tau$ -cell (see (2.3)), and T is the set of all values of  $\tau$ . The proof of (4.1) is similar to the one given in Lemma 2 in Khuri and Littell (1987) and will, therefore, be omitted. We note that the lower bound in (4.1) is the reciprocal of the harmonic mean of the  $\tau$ -cell frequencies.

#### 5. A numerical example

An example is given in Milliken and Johnson (1984, p. 264) of a study concerning the efficiency of workers in assembly lines at several plants. Three plants were randomly selected. Four assembly sites and three workers were randomly selected in each plant. For convenience, the efficiency scores are reproduced in Table 1.

The model for this experiment is

$$y_{ijk\ell} = \mu + \alpha_i + \beta_{i(j)} + \gamma_{i(k)} + (\beta\gamma)_{i(jk)} + \epsilon_{ijk\ell},$$

where  $\alpha_i$  is the effect of the i<sup>th</sup> plant (i=1,2,3),  $\beta_{i(j)}$  is the effect of the j<sup>th</sup> site within the i<sup>th</sup> plant (j=1,2,3,4),  $\gamma_{i(k)}$  is the effect of the k<sup>th</sup> worker within the i<sup>th</sup> plant (k=1,2,3),  $(\beta\gamma)_{i(jk)}$  is the interaction effect of sites and workers within plant i, and  $\epsilon_{ijk\ell}$  is the error term. The variance components are  $\sigma_{\alpha}^2$ ,  $\sigma_{\beta}^2$ ,  $\sigma_{\gamma}^2$ ,  $\sigma_{\beta\gamma}^2$  and  $\sigma_{\epsilon}^2$ , respectively.

To facilitate the understanding of the application of the exact testing procedure to this example, the reader is referred to Table 2 which lists the values of some key quantities used in the development of the exact tests. The expected mean square values of  $MS_i = SS_i/m_i$  (i=1,2,3,4) and  $MS_E^{(2)} = SS_E^{(2)}/\xi_2$  are given in Table 3.

From Tables 2 and 3 it can be seen that the value of the exact F-statistic for testing the hypothesis  $H_0$ :  $\sigma_{\beta\gamma}^2 = 0$  versus  $H_a$ :  $\sigma_{\beta\gamma}^2 > 0$  is  $F = MS_4/MS_E^{(2)} = 7.090$  with 18 and 47 degrees of freedom. The significance level is  $3.4 \times 10^{-8}$ . The second hypothesis to be tested is  $H_0$ :  $\sigma_{\beta}^2 = 0$  versus  $H_a$ :  $\sigma_{\beta}^2 > 0$ . The corresponding value of the F-statistic is  $F = MS_2/MS_4 = .994$  with 9 and 18 degrees of freedom. This is a nonsignificant test. Next, the value of the F-statistic for the hypothesis  $H_0$ :  $\sigma_{\gamma}^2 = 0$  is  $F = MS_3/MS_4 = 3.293$  with 6 and 18 degrees of freedom. The level of significance in this case is .023.

Finally, the testing of the hypothesis  $H_0$ :  $\sigma_{\alpha}^2 = 0$  versus  $H_a$ :  $\sigma_{\alpha}^2 > 0$  requires the use of Satterthwaite's procedure since no mean square exists in Table 3 with an expected value equal to that of MS<sub>1</sub> under  $H_0$ . The test statistic in this case is given by

$$F = \frac{MS_1}{MS_2 + MS_3 - MS_4} = 5.184,$$

which is approximately distributed as an F random variable with 2 and  $\eta$  degrees of freedom, where

$$\eta = \frac{\left(MS_2 + MS_3 - MS_4\right)^2}{\left(MS_2\right)^2/9 + \left(MS_3\right)^2/6 + \left(MS_4\right)^2/18} = 5.477.$$

The level of significance is .055.

#### Appendix A

This appendix gives the proof of Lemma 3.4.

#### Proof.

- (i) From (3.2),  $\underline{P}_0 = \underline{A}_0/b_0$  and  $\underline{b}_0 = c$  by (3.3) since  $\psi_0$  is the empty set, hence  $\psi_0^c = \tau$ . But, by (2.14),  $\underline{A}_0 = \underline{J}_c$ . It follows that  $\underline{P}_0 = \underline{J}_c/c$ . Consequently,  $\underline{Q}_0 = \underline{J}_c'/\sqrt{c}$ .
- (ii)  $Q_i Q'_i = I_{m_i}$ . This follows by the definition of  $Q_i$ . Furthermore,  $Q_i = V_i P_i$  for some matrix  $V_i$  of order  $m_i \times c$  (i=0,1,..., $\nu$ ). Since  $P_i P_{i'} = 0$  for  $i \neq i'$ , then  $Q_i Q'_{i'} = V_i P_i P_i V'_{i'} = 0$ .

(iii) From (3.4) and (3.5) it can be seen that  $A_j Q'_i = A_j P_i V'_i = 0$ , if  $\psi_i \not\subset \psi_j$ , and  $A_j Q'_i = A_j P_i V'_i = b_j P_i V'_i = b_j Q'_i$ , if  $\psi_i \subset \psi_j$ .

#### Appendix B

This appendix gives the proof of Lemma 3.5.

Proof.

(i) This is straightforward.

(ii) 
$$\mathbb{D}\mathbb{R} = \begin{bmatrix} \bigoplus_{\tau \in \mathbf{T}} (\underline{1}'_{n_{\tau}}/n_{\tau}) \prod \mathbb{I}_{N} - \bigoplus_{\tau \in \mathbf{T}} (\underline{J}_{n_{\tau}}/n_{\tau}) \end{bmatrix}$$
 (by (3.15) and (3.18))  
$$= \bigoplus_{\tau \in \mathbf{T}} (\underline{1}'_{n_{\tau}}/n_{\tau}) - \bigoplus_{\tau \in \mathbf{T}} (\underline{1}'_{n_{\tau}}/n_{\tau}) = 0.$$

(iii) The matrix  $X_i$  can be partitioned into c submatrices that correspond to the values of

 $\tau$  in (2.3). The submatrix corresponding to a particular  $\tau$  is of order  $n_{\tau} \times c_i$ , where  $c_i$  is the number of columns of  $X_i$  (i=1,2,..., $\nu$ ). Let us denote such a submatrix by  $U_{\tau}$ . Each column of  $U_{\tau}$  consists of either  $n_{\tau}$  zeros or  $n_{\tau}$  ones. Therefore,  $RX_i$  can be partitioned into c submatrices of orders  $n_{\tau} \times c_i$  for the different values of  $\tau$ . For a particular  $\tau$ , the corresponding submatrix is of the form

$$\left(\underline{\mathbf{I}}_{\mathbf{n}_{\tau}} - \underline{\mathbf{J}}_{\mathbf{n}_{\tau}}/\mathbf{n}_{\tau}\right)\underline{\mathbf{U}}_{\tau} = \underline{\mathbf{U}}_{\tau} - \underline{\mathbf{U}}_{\tau} = \underline{\mathbf{0}},$$

by the property of  $U_{\tau}$  described earlier. It follows that  $\mathbf{R}\mathbf{X}_{i} = \mathbf{0}$  for  $i=1,2,\ldots,\nu$ .

#### Acknowledgements

This research was partially supported by the Office of Naval Research under Grant N00014-86-K-0059.

#### References

- Cummings, W.B. and D.W. Gaylor (1974). Variance component testing in unbalanced nested designs. Journal of the American Statistical Association 69,765-771.
- Khuri, A.I. (1982). Direct products: A powerful tool for the analysis of balanced data. Communications in Statistics A11, 2903-2920.
- Khuri, A.I. and R.C. Littell (1987). Exact tests for the main effects variance components in an unbalanced random two-way model. *Biometrics* 43, 545-560.
- Milliken, G.A. and D.E. Johnson (1984). Analysis of Messy Data. Lifetime Learning Publications, Belmont, CA.
- Tietjen, G.L. (1974). Exact and approximate tests for unbalanced random effects designs. Biometrics 30, 573-581.

## Table 1

۶

Data for the numerical example

Plant 1				
Vorker	Site			
	1	2	3	4
	100.6	110.0	100.0	98.2
	106.8	105.8	102.5	<b>99.</b> 5
	100.6		97.6	
			98.7	
			98.7	100.0
	92.3	103.2	96.4	108.0
	92.0	100.5		108.9
	97.2	100.2		107.9
	93.9	97.7		
	93.0			
	96.9	92.5	86.8	94.4
	96.1	85.9		93.0
	100.8	85.2		91.0
		89.4		
		88.7		
Plant 2	82.6	96.5	87.9	83.6
		100.1	93.5	82.7
		101.9	88.9	87.7
		97.9	92.8	88.0
		95.9		82.5
	72.7	71.7	78.4	82.1
		72.1	80.4	79.9
		72.4	83.8	81.9
		71.4	77.7	82.6
			81.2	<u>78.6</u>
	82.5	80.9	96.3	77.7
	82.1	84.0	92.4	78. <b>6</b>
	82.0	82.2	92.0	77.2
		83.4	95.8	78.8
		81.5	<u>`</u>	80.5
Plant 3	107.6	96.1	101.1	109.1
	108.8	98.5		
	107.2	97.3		
	104.2	93.5		
	105.4			
	97.1	91.9	88.0	<b>89.6</b>
	94.2		91.4	<b>86.0</b>
	91.5		90.3	91.2
	99.2		91.5	87.4
			85.7	
	87.1	97.8	95.9	101.4
		95.9	89.7	100.1
				102.1
				98.4

-19-

1,11,22

**1 1 1 1** 

**1** 

1° . .

1

3

1. 1. 1

0.1

### Table 2

Quantity	Formula cited	Corresponding value
ν	(2.1)	4
τ	(2.3)	(i,j,k)
c	(2.5)	36
Ν	(2.6)	118
ь <sub>0</sub>	(3.3)	36
<sup>b</sup> 1	(3.3)	12
<sup>b</sup> 2	(3.3)	3
հ <sub>3</sub>	(3.3)	4
b <sub>4</sub>	(3.3)	1
P <sub>1</sub>	(3.2)	$(I_3 \otimes J_{12})/12$ - $J_{36}/36$
P <sub>2</sub>	(3.2)	$\left( \underline{\mathfrak{l}}_{12} \otimes \underline{\mathfrak{J}}_{3} \right) / 3 - \left( \underline{\mathfrak{l}}_{3} \otimes \underline{\mathfrak{J}}_{12} \right) / 12$
P <sub>3</sub>	(3.2)	$\left(\underline{I}_3 \otimes \underline{J}_4 \otimes \underline{I}_3\right)/4 - \left(\underline{I}_3 \otimes \underline{J}_{12}\right)/12$
P <sub>4</sub>	(3.2)	$ \underbrace{I_{36} - (I_{12} \otimes J_3)/3 - (I_3 \otimes J_4 \otimes I_3)/4 + (I_3 \otimes J_{12})/12 }_{(12)} $
<sup>m</sup> 1	(3.29)	2
<sup>m</sup> 2	(3.29)	9
m <sub>3</sub>	(3.29)	6
m <sub>4</sub>	(3.29)	18
$\delta_1$	(3.9)	$12\sigma_{\alpha}^2 + 3\sigma_{\beta}^2 + 4\sigma_{\gamma}^2 + \sigma_{\beta\gamma}^2$
<sup>8</sup> 2	(3.9)	$3\sigma_{eta}^2 + \sigma_{eta\gamma}^2$
δ <sub>3</sub>	(3.9)	$4\sigma_{\gamma}^2 + \sigma_{\beta\gamma}^2$
8 <sub>4</sub>	(3.9)	$\sigma_{\beta\gamma}^2$
ξ1	(3.24)	35
ξ <sub>2</sub>	(3.24)	47
$\lambda_{\max}$	(3.27)	1
$SS_1 = \omega_1' \omega_1$	Lemma 3.6	1265.96
$SS_2 = \omega'_2 \omega_2$	Lemma 3.6	332.313
$SS_3 = \omega'_3 \omega_3$	Lemma 3.6	733.949
$SS_4 = \omega'_4 \omega_4$	Lemma 3.6	668.634
$SS_E^{(2)}$	Lemma 3.6	246.245

Some key quantities used in the development of the exact tests for the numerical example

Table 3

Đ.

Ŕ

TO I

Expected mean square values for the nun	ierical example
---	-----------------

Mean square	Expected value
$MS_1 = SS_1/2$	$12\sigma_{\alpha}^{2}+3\sigma_{\beta}^{2}+4\sigma_{\gamma}^{2}+\sigma_{\beta\gamma}^{2}+\sigma_{\epsilon}^{2}$
$MS_2 = SS_2/9$	$3\sigma_{\beta}^2 + \sigma_{\beta\gamma}^2 + \sigma_{\epsilon}^2$
$MS_3 = SS_3/6$	$4\sigma_{\gamma}^2 + \sigma_{\beta\gamma}^2 + \sigma_{\epsilon}^2$
$\mathrm{MS}_4 = \mathrm{SS}_4/18$	$\sigma^2_{\beta\gamma} + \sigma^2_\epsilon$
$MS_{E}^{(2)} = SS_{E}^{(2)}/47$	$\sigma_{\epsilon}^2$

.

÷.

E L IMED