A MIXED MODEL FOR THE COMPLETE THREE-WAY LAYOUT WITH TWO RANDOM-EFFECTS FACTORS[1]

By J. P. Imhof[2]

University of California, Berkeley

1. Summary. In the present paper the Mixed Model developed by Scheffé [10] for the complete two-way layout is extended to the complete three-way layout with two random-effects factors. The model involves three basic covariance matrices of unknown parameters in addition to the error variance and fixed effects. Assuming normality, tests of the usual statistical hypotheses, except that of no fixed main effects, are derived from the analysis of variance table. Those of no interaction between the fixed-effects and a random-effects factor are applicable only under a simplifying assumption. A reduced form of the model is derived which involves sets of independent identically distributed random vectors. These are used to obtain unbiased estimators of the basic covariance matrices and to construct a T^2 test of the hypothesis of no fixed main effects. This test involves nonoptimum estimators of the effects, but this is shown to result in general only in a small loss of power. Individual and simultaneous confidence intervals for the fixed main effects are obtained in terms of these nonoptimum estimators.

2. Introduction. In analysis of variance problems involving both fixed-effects (or Model I) and random-effects (or Model II) factors, various Mixed Models have been proposed in recent years. Of those, certain arise as particular cases of very general models for which knowledge about distribution theory is at present only fragmentary [3], [12], [13], while another approach consists in setting up a "normal theory" model with sufficient assumptions so that, in particular, exact tests of the standard hypotheses can be derived. This has been done by Scheffé [10] in the case of the complete two-way layout. His model and the method of analysis derived from it can easily be extended to complete layouts of order higher than two, as long as there is only one random-effects factor. On the other hand, the analysis becomes considerably more intricate when the number of such factors is increased. We develop it for the case of three factors, two of which are Model II.

Matrices of the type

$$S = ((s_{ii'})), \qquad s_{ii'} = b + \delta_{ii'}(a - b), \qquad a \geqq b, \tag{2.1}$$

will frequently occur. Here $((s_{ii'}))$ denotes the matrix having elements $s_{ii'}$, $i = 1, \cdots, n$ refers to the row, $i' = 1, \cdots, n$ to the column and $\delta_{ii'}$ is the

---

Received January 19, 1959; revised March 16, 1960.

[1] This paper was prepared with the partial support of the Office of Naval Research (Nonr-222-43). This paper in whole or in part may be reproduced for any purpose of the United States Government. The paper is based on the author's doctoral dissertation.

[2] Now at the University of Geneva.

Kronecker δ. The canonical reduction of S can be performed by using an orthogonal matrix having elements $n^{-\frac{1}{2}}$ in the first row. In Section 8 it will be advantageous to use for this reduction a matrix with elements having as few different numerical values as possible.

LEMMA 1. *The symmetric matrix P of size n, with elements $p_{ii'} = n^{-\frac{1}{2}}$ if either $i = 1$ or $i' = 1$ and $p_{ii'} = \delta_{ii'} - (n-1)^{-1}(n^{-\frac{1}{2}} + 1)$ if $i > 1$ and $i' > 1$, is orthogonal. In particular, one has*

$$(2.2) \qquad \textstyle\sum_{i>1} p_{ii'}p_{ii''} = \delta_{i'i''} - n^{-1}; \qquad \sum_{i'=1}^{n} p_{ii'} = 0, \quad i = 2, \cdots, n.$$

The following further conventions are made regarding the notation used: Vectors are column vectors. Matrices will always be square. The transpose of A is written A', and $\mathrm{tr}A$ is the trace of A. We write $A = ((A_{ii'}))$ for a matrix A partitioned into submatrices $A_{ii'}$. Also, $((\text{diag. } A_1, \cdots, A_I))$ is the matrix $A = ((A_{ii'}))$, $i, i' = 1, \cdots, I$, where $A_{ii'} = \delta_{ii'}A_i$. If X is a random vector, EX is its expectation and Σ_X its covariance matrix. The vector X is $N(\beta, \Sigma_X)$ means X is normal with expectation β and covariance matrix Σ_X. The symbol U is reserved exclusively for the identity matrix. Finally, a dot substituted for a subscript indicates that the average has been taken over all permissible values of the subscript.

3. The model. Basic formulas. Consider a complete three-way layout design involving the factors A, B and C. A is a fixed-effects factor appearing in the experiment at levels $i = 1, \cdots, I$. B and C are random-effects factors. The levels of B and C at which the experiment is carried out are selected at random and independently as regards B and C from two conceptual infinite populations which we represent as two abstract spaces V and W. The selected levels are labeled v_j $(j = 1, \cdots, J)$ and w_k $(k = 1, \cdots, K)$ respectively. For each of the IJK combinations of levels (i, j, k), L replications are performed. The observed responses are labeled y_{ijkl}. When no further indication is given, subscripts i, j, k, l range over the values 1 to I, 1 to J, 1 to K and 1 to L respectively. It is always assumed that $I, J, K > 1$ and $L \geqq 1$. The case $L = 1$ is formally included, but certain of the results are then obviously meaningless. In later sections, additional restrictions are imposed on I, J and K.

An example may be obtained by extending the illustration given in [10] as follows: Think of an experiment in which K batches of material are used by each of J workers on each of I different makes of machines, the output of the jth worker with the kth batch on the ith machine being determined separately over L experimental periods, to yield the observed outputs y_{ijkl}. Here the workers and batches selected for the experiment are considered to be chosen randomly from the idealized infinite populations of workers and batches that might have been used in the experiment, and the replications are supposed to be carried out in such a way that they do not interact with the other factors, in particular not with the factor "worker".

We assume that the response (in our example, the output) y_{ijkl} has the structure

$$y_{ijkl} = m(i, v_j, w_k) + e_{ijkl}, \tag{3.1}$$

where the e_{ijkl} are "errors" and the $m(i, v_j, w_k)$ are "true cell means." The random selection of v_j in V and, independently, of w_k in W justifies, in our view, the fundamental assumption that $m(i, v_j, w_k)$ is distributed for all j and k like a basic random variable $m(i, v, w)$ and that $m(i, v_j, w_k)$ is independent of $m(i', v_{j'}, w_{k'})$ when both $j \neq j'$ and $k \neq k'$. There is nothing however to justify assuming independence when $j = j'$ or $k = k'$. One can think of the two infinite populations from which the levels of factors B and C are drawn as corresponding to two probability distributions $\mathcal{P}_V$ and $\mathcal{P}_W$ on V and W, so that $(V, \mathcal{P}_V)$ and $(W, \mathcal{P}_W)$ are probability spaces and the distribution of the random variable $m(i, v, w)$ is that of the real-valued function $m(i, v, w)$ on the product space $(V \times W, \mathcal{P}_V \times \mathcal{P}_W)$. One can then define the random variables

$$\begin{aligned} m(i, v, \cdot) &= \int m(i, v, w)\, d\mathcal{P}_W(w), \\ m(i, \cdot, w) &= \int m(i, v, w)\, d\mathcal{P}_V(v). \end{aligned} \tag{3.2}$$

The first of these would have, in our example, the interpretation of true mean of a randomly selected worker labeled v when he uses machine i, averaged over the population of batches. A similar interpretation can be given for $m(i, \cdot, w)$. The second moment structure of the model depends essentially on three basic covariance matrices having elements

$$\begin{aligned} \sigma_{ii'} &= \text{Cov}\,\{m(i, v, w), m(i', v, w)\}, \\ \nu_{ii'} &= \text{Cov}\,\{m(i, v, \cdot), m(i', v, \cdot)\}, \\ \tau_{ii'} &= \text{Cov}\,\{m(i, \cdot, w), m(i', \cdot, w)\}, \end{aligned} \tag{3.3}$$

and on the linear combinations

$$\rho_{ii'} = \sigma_{ii'} - \nu_{ii'} - \tau_{ii'}. \tag{3.4}$$

Assume $\sigma_{ii} < \infty$, all i. The relation $\rho_{ii} \geqq 0$ obtained in Section 6 implies then finiteness of the ν_{ii}'s and τ_{ii}'s also.

The assumptions made so far are, we believe, realistic: They express what is implied by the random selection of the levels of B and, independently, of C from the two conceptual infinite populations V and W described above. Further assumptions which are needed for computing $E(MS)$'s (expected mean squares) and finding unbiased estimators are less satisfactory: The errors e_{ijkl} are assumed to be pairwise uncorrelated and to have zero means and a common variance σ_e^2. Furthermore, the e_{ijkl} are assumed to be uncorrelated with the $m(i', v_{j'}, w_{k'})$ for all i, i', j, j', k, k' and l. In the particular case $I = 1$, the model coincides with the one described in Section 7.4 of [11] for the complete two-way layout under Model II.

The exact distribution of test criteria, and exact confidence intervals for various parameters are derived in later sections under an additional normality assumption, namely that the random variables e_{ijkl}, $m(i, v, w)$, $m(i, v, \cdot)$ and $m(i, \cdot, w)$ have joint normal distributions. For instance, V and W could be two real lines, $\mathcal{P}_V$ and $\mathcal{P}_W$ be two independent normal distributions on them and the functions $m(i, v, w)$ be for each i linear functions of v and w. On the other hand, the case $m(i, v, w) = vw$ shows that joint normality of the $m(i, v, \cdot)$ and $m(i, \cdot, w)$ does not imply that of the $m(i, v, w)$. It can also be verified that joint normality of the $m(i, v, w)$ does not imply that of the $m(i, v, \cdot)$ and $m(i, \cdot, w)$.

We define main effects and interactions in a natural and conventional way by letting

$$
\begin{aligned}
&\mu = m(\cdot, \cdot, \cdot), \qquad \alpha_i^A = m(i, \cdot, \cdot) - m(\cdot, \cdot, \cdot), \\
&a^B(v) = m(\cdot, v, \cdot) - m(\cdot, \cdot, \cdot), \\
&a_i^{AB}(v) = m(i, v, \cdot) - m(i, \cdot, \cdot) - m(\cdot, v, \cdot) + m(\cdot, \cdot, \cdot), \\
&a^{BC}(v, w) = m(\cdot, v, w) - m(\cdot, v, \cdot) - m(\cdot, \cdot, w) + m(\cdot, \cdot, \cdot), \\
&a_i^{ABC}(v, w) = m(i, v, w) - m(i, v, \cdot) - \cdots + m(i, \cdot, \cdot) \\
&\qquad\qquad + \cdots - m(\cdot, \cdot, \cdot).
\end{aligned}
\tag{3.5}
$$

In those formulas a dot substituted for v or w means that expected value has been taken with respect to $\mathcal{P}_V$ or $\mathcal{P}_W$, as in (3.2), while a dot in place of i means the arithmetic average has been taken over the values $i = 1, \cdots, I$. Thus μ is the general mean, α_i^A is the main effect of factor A at level i, $a^B(v)$ is the main effect of factor B at level v, etc. Except for μ and the α_i^A, all main effects and interactions are obviously random and one finds at once from their definition

$$
Ea^B(v) = Ea_i^{AB}(v) = Ea^{BC}(v, w) = Ea_i^{ABC}(v, w) = 0. \tag{3.6}
$$

We have omitted writing $a^C(w)$ and $a_i^{AC}(w)$ because their definitions are similar to those for $a^B(v)$ and $a_i^{AB}(v)$. As a general rule, when considering variance components, sums of squares, estimators, etc., "C" can be treated like "B" and "AC" like "AB" by substituting k for j, $\tau_{ii'}$ for $\nu_{ii'}$, $m(i, \cdot, w)$ for $m(i, v, \cdot)$, etc.

Next we define variance components by using the analogy with a "finite model" described in [10]. This leads to the natural definitions

$$
\begin{aligned}
&\sigma_B^2 = \operatorname{Var} a^B(v), \qquad \sigma^2 = (I - 1)^{-1} \sum_i \operatorname{Var} a_i^{AB}(v), \\
&\sigma_{BC}^2 = \operatorname{Var} a^{BC}(v, w), \qquad \sigma_{ABC}^2 = (I - 1)^{-1} \sum_i \operatorname{Var} a_i^{ABC}(v, w).
\end{aligned}
\tag{3.7}
$$

As usual, let

$$(3.8)\qquad \sigma_A^2 = (I-1)^{-1}\sum_i (\alpha_i^A)^2.$$

In terms of the basic parameters (3.3) and of their linear combinations (3.4) the variance components become

$$(3.9)\qquad \begin{aligned} \sigma_B^2 &= \nu_{\cdot\cdot}, & \sigma_{AB}^2 &= (I-1)^{-1}\sum_i (\nu_{ii} - \nu_{\cdot\cdot}), \\ \sigma_{BC}^2 &= \rho_{\cdot\cdot}, & \sigma_{ABC}^2 &= (I-1)^{-1}\sum_i (\rho_{ii} - \rho_{\cdot\cdot}). \end{aligned}$$

These relations are best derived from (3.15) below.

For the levels v_j $(j = 1, \cdots, J)$ and $w_k (k = 1, \cdots, K)$ of v and w selected in the experiment, equations (3.5) identically give

$$m(i, v_j, w_k) = \mu + \alpha_i^A + a^B(v_j) + a^C(w_k) + a_i^{AB}(v_j) + \cdots + a_i^{ABC}(v_j, w_k).$$

This notation being cumbersome, we write simply

$$(3.10)\qquad m_{ijk} = \mu + \alpha_i^A + a_j^B + a_k^C + a_{ij}^{AB} + a_{ik}^{AC} + a_{jk}^{BC} + a_{ijk}^{ABC},$$

and then (3.1) becomes

$$(3.11)\qquad y_{ijkl} = m_{ijk} + e_{ijkl}.$$

For all j and k, the a_j^B are identically distributed like $a^B(v)$, $\cdots$, the a_{ijk}^{ABC} are identically distributed like $a_i^{ABC}(v, w)$, and (3.5) implies that

$$(3.12)\qquad \alpha_\cdot^A = a_{\cdot j}^{AB} = a_{\cdot k}^{AC} = a_{\cdot jk}^{ABC} = 0, \qquad \text{all } j, k.$$

The main effects and interactions entering (3.10) are independent, except for the three pairs in (3.13) for which one easily finds

$$(3.13)\qquad \begin{gathered} \operatorname{Cov}(a_j^B, a_{ij}^{AB}) = \nu_{i\cdot} - \nu_{\cdot\cdot}, \qquad \operatorname{Cov}(a_k^C, a_{ik}^{AC}) = \tau_{i\cdot} - \tau_{\cdot\cdot}, \\ \operatorname{Cov}(a_{jk}^{BC}, a_{ijk}^{ABC}) = \rho_{i\cdot} - \rho_{\cdot\cdot}. \end{gathered}$$

Further covariances will be needed. Let

$$(3.14)\qquad \begin{aligned} \overset{0}{\sigma}_{ii'} &= \sigma_{ii'} - \sigma_{i\cdot} - \sigma_{\cdot i'} + \sigma_{\cdot\cdot}, & \overset{0}{\nu}_{ii'} &= \nu_{ii'} - \nu_{i\cdot} - \nu_{\cdot i'} + \nu_{\cdot\cdot}, \\ \overset{0}{\tau}_{ii'} &= \tau_{ii'} - \tau_{i\cdot} - \tau_{\cdot i'} + \tau_{\cdot\cdot}, & \overset{0}{\rho}_{ii'} &= \rho_{ii'} - \rho_{i\cdot} - \rho_{\cdot i'} + \rho_{\cdot\cdot}. \end{aligned}$$

One finds

$$(3.15)\qquad \begin{gathered} \operatorname{Cov}(a_{ij}^{AB}, a_{i'j'}^{AB}) = \delta_{jj'}\overset{0}{\nu}_{ii'}, \qquad \operatorname{Cov}(a_{ik}^{AC}, a_{i'k'}^{AC}) = \delta_{kk'}\overset{0}{\tau}_{ii'}, \\ \operatorname{Cov}(a_{jk}^{BC}, a_{j'k'}^{BC}) = \delta_{jj'}\delta_{kk'}\rho_{\cdot\cdot}, \qquad \operatorname{Cov}(a_{ijk}^{ABC}, a_{i'j'k'}^{ABC}) = \delta_{jj'}\delta_{kk'}\overset{0}{\rho}_{ii'}, \\ \operatorname{Cov}(a_j^B, a_{j'k'}^{BC}) = \operatorname{Cov}(a_k^C, a_{j'k'}^{BC}) = \operatorname{Cov}(a_{ij}^{AB}, a_{i'k'}^{AC}) \\ = \operatorname{Cov}(a_{ij}^{AB}, a_{i'j'k'}^{ABC}) = \operatorname{Cov}(a_{ik}^{AC}, a_{i'j'k'}^{ABC}) = 0, \\ \text{all } i, i', j, j', k, k'. \end{gathered}$$

The derivation of these relations is routine, based mainly on equalities like

$$(3.16)\quad \text{Cov}(m_{ijk}, m_{i'j'k}) \doteq \text{Cov}\{m(i, \cdot, w), m(i', \cdot, w)\} = \tau_{ii'}, \quad \text{for } j \neq j'$$

This follows from

$$E[m(i, v_j, w_k)m(i', v_{j'}, w_k)] = E\{E[m(i, v_j, w_k)m(i', v_{j'}, w_k) \mid w_k]\}$$
$$E[m(i, \cdot, w_k)m(i', \cdot, w_k)]$$

when $j \neq j'$.

4. The analysis of variance table. Immediate consequences. In order to obtain point estimators of the variance components and test criteria for the usual statistical hypotheses, we proceed in the usual fashion, which consists in writing the SS's (sums of squares) that one considers in the corresponding pure Model I complete three-way layout and computing the $E(MS)$'s, the numbers of d.f. (degrees of freedom) being the ranks of the quadratic forms defining the SS's. Using (3.10), (3.11), (3.12) one has,

$$(4.1)\quad \begin{aligned} SS_A &= JKL\sum_i (y_{i\cdots} - y_{\cdots\cdot})^2 \\ &= JKL\sum_i (\alpha_i^A + a_{i\cdot}^{AB} + a_{i\cdot}^{AC} + a_{i\cdot\cdot}^{ABC} + e_{i\cdots} - e_{\cdots\cdot})^2 \end{aligned}$$

and so forth, the well-known expressions for the SS's yielding

$$(4.2)\quad \begin{aligned} SS_B &= IKL\sum_j (a_j^B - a_\cdot^B + a_{j\cdot}^{BC} - a_{\cdot\cdot}^{BC} + e_{\cdot j\cdot\cdot} - e_{\cdots\cdot})^2, \\ SS_{AB} &= KL\sum_i\sum_j (a_{ij}^{AB} - a_{i\cdot}^{AB} + a_{ij\cdot}^{ABC} - a_{i\cdot\cdot}^{ABC} + e_{ij\cdot\cdot} - e_{i\cdots} - e_{\cdot j\cdot\cdot} \\ &\quad + e_{\cdots\cdot})^2, \\ SS_{BC} &= IL\sum_j\sum_k (a_{jk}^{BC} - a_{j\cdot}^{BC} - a_{\cdot k}^{BC} + a_{\cdot\cdot}^{BC} + e_{\cdot jk\cdot} - e_{\cdot j\cdot\cdot} - e_{\cdot\cdot k\cdot} \\ &\quad + e_{\cdots\cdot})^2, \\ SS_{ABC} &= L\sum_i\sum_j\sum_k (a_{ijk}^{ABC} - a_{ij\cdot}^{ABC} - a_{i\cdot k}^{ABC} + a_{i\cdot\cdot}^{ABC} + e_{ijk\cdot} - e_{ij\cdot\cdot} \\ &\quad - \cdots - e_{\cdots\cdot})^2, \\ SS_e &= \sum_i\sum_j\sum_k\sum_l (e_{ijkl} - e_{ijk\cdot} - \cdots + e_{ij\cdot\cdot} + \cdots - e_{i\cdots} \\ &\quad - \cdots + e_{\cdots\cdot})^2. \end{aligned}$$

Consider now the computation of $E(MS_A)$. From (3.7), (3.8), (3.15) and the assumptions below (3.4) it follows that

$$\begin{aligned} ESS_A &= JKL\sum_i [(\alpha_i^A)^2 + E(a_{i\cdot}^{AB})^2 + E(a_{i\cdot}^{AC})^2 + E(a_{i\cdot\cdot}^{ABC})^2 + E(e_{i\cdots} - e_{\cdots\cdot})^2] \\ &= (I-1)JKL[\sigma_A^2 + J^{-1}\sigma_{AB}^2 + K^{-1}\sigma_{AC}^2 + (JK)^{-1}\sigma_{ABC}^2 + (JKL)^{-1}\sigma_e^2]. \end{aligned}$$

Proceeding in the same fashion for other SS's leads to the following analysis of variance table:

(4.3)

SS	d.f.	$E(MS)$
A	$I-1$	$JKL\sigma_A^2 + KL\sigma_{AB}^2 + JL\sigma_{AC}^2 + L\sigma_{ABC}^2 + \sigma_e^2$
B	$J-1$	$IKL\sigma_B^2 + IL\sigma_{BC}^2 + \sigma_e^2$
AB	$(I-1)(J-1)$	$KL\sigma_{AB}^2 + L\sigma_{ABC}^2 + \sigma_e^2$
BC	$(J-1)(K-1)$	$IL\sigma_{BC}^2 + \sigma_e^2$
ABC	$(I-1)(J-1)(K-1)$	$L\sigma_{ABC}^2 + \sigma_e^2$
error	$IJK(L-1)$	σ_e^2

This coincides with the table one would obtain by applying the rules of Bennett and Franklin [1] for writing down $E(MS)$'s. The table (4.3) shows that the variance components (3.7) admit the unbiased estimators

$$(4.4)\quad \begin{aligned} \hat{\sigma}_B^2 &= (IKL)^{-1}(MS_B - MS_{BC}), & \hat{\sigma}_{BC}^2 &= (IL)^{-1}(MS_{BC} - MS_e), \\ \hat{\sigma}_{AB}^2 &= (KL)^{-1}(MS_{AB} - MS_{ABC}), & \hat{\sigma}_{ABC}^2 &= L^{-1}(MS_{ABC} - MS_e), \end{aligned}$$

while as usual σ_e^2 has the unbiased estimator $\hat{\sigma}_e^2 = MS_e$. Here and in further sections, the caret is used to denote estimators which are unbiased, but are not in general maximum likelihood estimators.

Natural hypotheses to consider are

$$H_A\colon\ \sigma_A^2 = 0,\ H_B\colon\ \sigma_B^2 = 0,\ H_{BC}\colon\ \sigma_{BC}^2 = 0,\ H_{AB}\colon\ \sigma_{AB}^2 = 0,\ H_{ABC}\colon\ \sigma_{ABC}^2 = 0.$$

The hypothesis H_A will be considered in Section 8. For the other hypotheses (i.e., those relative to random effects) the table of $E(MS)$'s suggests using the criteria MS_B/MS_{BC}, MS_{BC}/MS_e, MS_{AB}/MS_{ABC}, MS_{ABC}/MS_e respectively. In order to get some insight into the meaning of the different hypotheses considered, notice that $H_B \Leftrightarrow \nu_{\cdot\cdot} = 0$ or, using (3.7), $H_B \Leftrightarrow m(\cdot, v, \cdot)$ has a degenerate distribution, $m(\cdot, v, \cdot) = c$, a.s. (i.e., with probability one). This corresponds exactly to the intuitive idea of no main effect due to factor B. Similarly, one can write

$$(4.5)\quad \begin{aligned} &H_{BC} \Leftrightarrow \rho_{\cdot\cdot} = 0 \Leftrightarrow m(\cdot, v, w) = m(\cdot, v, \cdot) + m(\cdot, \cdot, w) + c, \text{ a.s.}, \\ &H_{AB} \Leftrightarrow \nu_{ii'} = \sigma_B^2, \quad \text{all } i, i' \Leftrightarrow m(i, v, \cdot) = m(\cdot, v, \cdot) + c_i, \text{ a.s.}, \\ &H_{ABC} \Leftrightarrow \rho_{ii'} = \sigma_{BC}^2, \quad \text{all } i, i' \Leftrightarrow m(i, v, w) - m(i, v, \cdot) - m(i, \cdot, w) \\ &\qquad = m(\cdot, v, w) - m(\cdot, v, \cdot) - m(\cdot, \cdot, w) + c_i, \text{ a.s.} \end{aligned}$$

Consider first testing H_B and H_{BC}. Letting $v_{jk} = a_{jk}^{BC} + e_{\cdot jk\cdot}$ gives

$$SS_B = IKL \sum_j (a_j^B - a_\cdot^B + v_{j\cdot} - v_{\cdot\cdot})^2,$$

$$SS_{BC} = IL \sum_j \sum_k (v_{jk} - v_{j\cdot} - v_{\cdot k} + v_{\cdot\cdot})^2.$$

According to (3.15) the variables $\{a_j^B, v_{jk}\}$ are mutually independent and $\operatorname{Var} v_{jk} = \sigma_{BC}^2 + (IL)^{-1}\sigma_e^2$. Using the familiar results of Model I theory one finds that when the normality assumptions described above (3.5) are made,

$$\frac{MS_B}{MS_{BC}} = \frac{IKL\sigma_B^2 + IL\sigma_{BC}^2 + \sigma_e^2}{IL\sigma_{BC}^2 + \sigma_e^2} F_{J-1,(J-1)(K-1)}$$

where $F_{m,n}$ is an F-variable with m and n d.f. Thus we reject H_B at the α level of significance if MS_B/MS_{BC} is larger than the upper α-point of this distribution. In the same fashion,

$$\frac{MS_{BC}}{MS_e} = \frac{IL\sigma_{BC}^2 + \sigma_e^2}{\sigma_e^2} F_{(J-1)(K-1),IJK(L-1)} .$$

In this case one can test the more realistic assumption $H_{BC}^{(\theta)}$: $\sigma_{BC}^2/\sigma_e^2 < \theta$ by rejecting it if $(IL\theta + 1)^{-1}MS_{BC}/MS_e$ exceeds the upper α-point of the F-distribution with the above numbers of d.f.

5. The hypotheses H_{AB} and H_{ABC}. We investigate in this section the distributions of the ratios MS_{AB}/MS_{ABC} and MS_{ABC}/MS_e suggested by (4.3) for testing H_{AB} and H_{ABC}. The normality assumptions are made throughout. Let

$$(5.1)\quad b_{ijk} = a_{ij}^{AB} + a_{ijk}^{ABC} + e_{ijk\cdot}, \qquad c_{ijk} = b_{ijk} - b_{i\cdot k} - b_{\cdot jk} + b_{\cdot\cdot k} .$$

Then

$$(5.2)\qquad SS_{AB} = KL\sum_i \sum_j c_{ij\cdot}^2 , \qquad S_{ABC} = L\sum_i \sum_j \sum_k (c_{ijk} - c_{ij\cdot}) .$$

Using (3.15) gives

$$(5.3)\qquad \mathrm{Cov}(b_{ijk}, b_{i'j'k'}) = \delta_{jj'}[\nu_{ii'}^0 + \delta_{kk'}(\rho_{ii'}^0 + \delta_{ii'}L^{-1}\sigma_e^2)].$$

Then, noticing e.g., that $\sigma_{i\cdot}^0 = \sigma_{\cdot i'}^0 = 0$ one finds

$$(5.4)\quad \mathrm{Cov}(c_{ijk}, c_{i'j'k'}) = (\delta_{jj'} - J^{-1})\{\nu_{ii'}^0 + \delta_{kk'}[\rho_{ii'}^0 + (\delta_{ii'} - I^{-1})L^{-1}\sigma_e^2]\},$$

from which it follows that $\mathrm{Cov}(c_{ijk} - c_{ij\cdot}, c_{i'j'\cdot}) = 0$. Hence SS_{AB} and SS_{ABC} are statistically independent and we only need investigate their distributions separately. For this purpose a well-known result [2] relative to quadratic forms in normal variables is used: If the vector X is $N(0, \Sigma)$, the quadratic form $X'QX$ has the distribution of $\sum_r \lambda_r\chi_{(r),1}^2$, where the χ^2 variables each with one d.f., are independent and the coefficients λ_r are the nonzero latent roots of the matrix ΣQ. Consider first SS_{AB}. According to (5.1) it can be written $SS_{AB} = KL(X'X)$, where X is the vector

$$X = (x_{11}, \cdots, x_{1J}, \cdots, x_{i1}, \cdots, x_{iJ}, \cdots, x_{I1}, \cdots, x_{IJ})', x_{ij} = c_{ij\cdot} .$$

The elements of its covariance matrix Σ are found from (5.4) to be given by

$$\mathrm{Cov}(x_{ij}, x_{i'j'}) = (\delta_{jj'} - J^{-1})\{\nu_{ii'}^0 + K^{-1}[\rho_{ii'}^0 + (\delta_{ii'} - I^{-1})L^{-1}\sigma_e^2]\}.$$

Thus one can write $\Sigma = ((\Sigma_{ii'}))$, where each submatrix $\Sigma_{ii'}$ of size J has the structure (2.1) and has row sums equal to zero. Let $P^* = ((\text{diag. } P, \cdots, P))$, the I diagonal blocks P of size J being as in Lemma 1. The nonzero latent roots of Σ, which equal those of $P^*\Sigma P^*$, are then found to be the nonzero latent roots of the matrix of size $I(J - 1)$, $M^* = ((M_{ii'}))$, where the $M_{ii'}$ are diagonal matrices having their $J - 1$ diagonal elements equal to $m_{ii'} = \nu^0_{ii'} + K^{-1}\rho^0_{ii'} + (IKL)^{-1}(I\delta_{ii'} - 1)\sigma_e^2$. Letting $M = ((m_{ii'}))$, one verifies that $|M^* - \lambda U| = |M - \lambda U|^{J-1}$ (where the identity matrix U has in each case the proper size), and so the nonzero latent roots of M^* are, each with order of multiplicity $J - 1$, those of M. Now, substituting in $|M - \lambda U|$ the sum of the columns (all components of which equal to $-\lambda$) for column I, and then row i minus row I for row i ($i = 1, \cdots, I - 1$) and developing the resulting determinant in terms of elements of the last row one finally obtains that

$$SS_{AB} = \sum_{r=1}^{I-1} \epsilon_r \chi^2_{(r),J-1}, \tag{5.5}$$

where the $I - 1$ variables $\chi^2_{(r),J-1}$ have independent χ^2 distributions with $J - 1$ d.f. each and $\epsilon_1, \cdots, \epsilon_{I-1}$ are the latent roots of the matrix of size $I - 1$,

$$C = A + B + \sigma_e^2 U, \tag{5.6}$$

where

$$A = KL((\nu^0_{rr'} - \nu^0_{Ir'})), \qquad B = L((\rho^0_{rr'} - \rho^0_{Ir'})), \qquad r, r' = 1, \cdots, I - 1. \tag{5.7}$$

As a general rule we shall substitute subscripts r, r' for the subscripts i, i' whenever the range of values is 1 to $I - 1$ instead of 1 to I. One has $\sum \epsilon_r = tr\, C = (I - 1)(KL\sigma^2_{AB} + L\sigma^2_{ABC} + \sigma_e^2)$. More can, in fact, be said about the ϵ's: Consider the matrix of size I, $H = ((KL\nu^0_{ii'} + L\rho^0_{ii'}))$. It is a covariance matrix (of the vector with components $(JKL)^{\frac{1}{2}}(a^{AB}_{i.} + a^{ABC}_{i..})$), hence its latent roots $\mu_1, \cdots, \mu_I$ are $\geqq 0$. Thus the latent roots $\mu_i^+ = \mu_i + \sigma_e^2$ of $H^+ = H + \sigma_e^2 U$ are $\geqq \sigma_e^2$. But performing the same column and row operations as above (5.5) one finds that

$$|H^+ - \mu^+ U| = (\sigma_e^2 - \mu^+)|C - \mu^+ U|,$$

where C is given by (5.6). In other words $I - 1$ of the latent roots of H^+ coincide with those of C, while the last one equals σ_e^2. Thus

LEMMA 3. *$\epsilon_r \geqq \sigma_e^2$, $r = 1, \cdots, I - 1$, and* $\epsilon_. = KL\sigma^2_{AB} + L\sigma^2_{ABC} + \sigma_e^2$.

Next consider SS_{ABC}. Define the vector

$$X^* = (c_{111}, \cdots, c_{11K}, c_{121}, \cdots, c_{12K}, \cdots, c_{1JK}, c_{211}, \cdots, c_{IJK})'. \tag{5.8}$$

Then $SS_{ABC} = KL\, X^{*\prime}Q^*X^*$ where $Q^* = ((\text{diag. } Q, \cdots, Q))$ and the IJ diagonal blocks Q of size K are given by $Q = ((\delta_{kk'} - 1))$. The method used above to reduce the computation of the latent roots of Σ can be applied now to Σ^*Q^*, where the elements of the covariance matrix Σ^* of X^* are given by (5.4). One finds that

$$SS_{ABC} = \sum_{r=1}^{I-1} \epsilon_r' \chi^2_{(r),(J-1)(K-1)}, \tag{5.9}$$

where the $I - 1$ variables $\chi^2_{(r),(J-1)(K-1)}$ have independent χ^2 distributions with $(J-1)(K-1)$ d.f. each and $\epsilon_1', \cdots, \epsilon_{I-1}'$ are the latent roots of the matrix of size I − 1,

$$D = B + \sigma_e^2 U, \tag{5.10}$$

where B is given by (5.7). The analogue of Lemma 3 is here

LEMMA 4. *$\epsilon_r' \geqq \sigma_e^2$, $r = 1, \cdots, I-1$ and $\epsilon_.' = L\sigma_{ABC}^2 + \sigma_e^2$.*

Consider the test of the hypotheses H_{AB}, based on the criterion MS_{AB}/MS_{ABC}. Under H_{AB}, (3.14) and (4.5) show that the matrix A of (5.7) reduces to a zero matrix and so, comparing (5.6) and (5.10), one has $\epsilon_r = \epsilon_.'$. Hence MS_{AB}/MS_{ABC} has the distribution of

$$(K-1)\frac{\sum_r \epsilon_r' \chi^2_{(r),J-1}}{\sum_r \epsilon_r' \chi^2_{(r),(J-1)(K-1)}}, \tag{5.11}$$

where all χ^2 variables are independent. This distribution is simple only if $\epsilon_1' = \cdots = \epsilon_{I-1}' = \epsilon_.'$, and it is easily verified that this is the case if, and only if, the matrix $((\rho_{ii'}))$ has the structure (2.1). Then MS_{AB}/MS_{ABC} has the F-distribution with $(I-1)$ $(J-1)$ and $(I-1)(J-1)(K-1)$ d.f. Letting $b_i(v, w) = a_i^{ABC}(v, w) + a^{BC}(v, w)$, one has $\rho_{ii'} = \text{Cov}\,[b_i(v, w), b_{i'}(v, w)]$. This does not help much in giving the above restriction on the $\rho_{ii'}$ a simple physical significance. However, the stronger assumption that all three covariance matrices (3.3) have the structure (2.1) carries more intuitive meaning.

The situation is simpler when testing H_{ABC}. Under this hypothesis, B of (5.7) is a zero matrix so that the criterion MS_{ABC}/MS_e has under the hypothesis the F-distribution with $(I-1)(J-1)(K-1)$ and $IJK(L-1)$ d.f. Concerning the power of the test, one might remark as follows: Specifying a value for σ_{ABC}^2 and σ_e^2 does not specify a unique alternative but a subclass, say $\mathcal{C}(\sigma_{ABC}^2, \sigma_e^2)$, of alternatives. Among those, one might intuitively feel that the hardest ones to distinguish from the hypothesis H_{ABC}: $\rho_{ii'} = \sigma_{BC}^2$ for all i, i' are the ones for which $((\rho_{ii'}))$ has the structure (2.1). When such is the case, $\epsilon_1' = \cdots = \epsilon_{I-1}' = L\sigma_{ABC}^2 + \sigma_e^2$, so that $(L\sigma_{ABC}^2 + \sigma_e^2)^{-1}MS_{ABC}/MS_e$ has the F-distribution and the power is immediately computable. Another reason why the power against those particular alternatives can be expected to be a lower bound (or at least nearly so) of power values for all alternatives in $\mathcal{C}(\sigma_{ABC}^2, \sigma_e^2)$ is that the cumulants of (5.9), obtainable from formula (2.3) of [2], are minimum when $\epsilon_1' = \cdots = \epsilon_{I-1}'$. The cumulants being positive, the same is true for the moments. Now SS_{ABC} is $\geqq 0$, its mean is fixed for fixed σ_{ABC}^2 and σ_e^2, and one therefore expects its distribution to have least mass in the tail when the moments are smallest. The same should then also be true for the distribution of MS_{ABC}/MS_e, which indeed (at least for small enough values of the level of significance) means that the power is lowest when $((\rho_{ii'}))$ has the structure (2.1).

6. Reduced form of the model. Sufficient statistics. By means of an orthogonal transformation, we obtain in this section certain sets of uncorrelated identically distributed vectors. Assuming normality, they yield sufficient statistics for the parameters of the model. Introduce the vector of observations, ordered as follows

$$(6.1)\quad Y = (y_{1111}, \cdots, y_{111L}, y_{1121}, \cdots, y_{112L}, \cdots, y_{11KL}, y_{1211}, \cdots, y_{12KL}, \cdots, y_{IJKL})'.$$

When normality is assumed, its probability density is

$$(6.2)\quad p(Y) = \text{const.}\, | \Sigma_Y |^{-\frac{1}{2}} \exp\{-\tfrac{1}{2}(Y - EY)'\Sigma_Y^{-1}(Y - EY)\},$$

where, putting $\beta_i = \mu + \alpha_i$,

$$EY = (\beta_1, \cdots, \beta_1, \beta_2, \cdots, \beta_2, \cdots, \beta_I, \cdots, \beta_I)',$$

each β_i being repeated JKL times. The elements of Σ_Y are given by

$$(6.3)\quad \text{Cov}(y_{ijkl}, y_{i'j'k'l'}) = \begin{cases} \sigma_{ii'} + \delta_{ii'}\delta_{ll'}\sigma_e^2 & \text{if } j = j',\ k = k', \\ \nu_{ii'} & \text{if } j = j',\ k \neq k', \\ \tau_{ii'} & \text{if } j \neq j',\ k = k', \\ 0 & \text{if } j \neq j',\ k \neq k'. \end{cases}$$

Write $\Sigma_Y = \Lambda + \sigma_e^2 U$ and partition Λ into $\Lambda = ((\Lambda_{ii'}))$.
The submatrices $\Lambda_{ii'}$ of size JKL are then given by

$$(6.4)\quad \Lambda_{ii'} = \left[\begin{array}{cccc|cccc} F_{ii'} & G_{ii'} & \cdots & G_{ii'} & H_{ii'} & 0 & \cdots & 0 \\ G_{ii'} & F_{ii'} & \cdots & G_{ii'} & 0 & H_{ii'} & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdots & \cdot \\ G_{ii'} & G_{ii'} & \cdots & F_{ii'} & 0 & 0 & \cdots & H_{ii'} \\ \hline H_{ii'} & 0 & \cdots & 0 & F_{ii'} & G_{ii'} & \cdots & G_{ii'} \\ 0 & H_{ii'} & \cdots & 0 & G_{ii'} & F_{ii'} & \cdots & G_{ii'} \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & H_{ii'} & G_{ii'} & G_{ii'} & \cdots & F_{ii'} \end{array}\right] \begin{array}{l} k=1 \\ k=2 \\ \qquad j=1 \\ k=K \\ k=1 \\ k=2 \\ \qquad j=2 \\ k=K \end{array}$$

where each of the submatrices $F_{ii'}$, $G_{ii'}$, $H_{ii'}$ of size L has all its elements equal to $\sigma_{ii'}$, $\nu_{ii'}$, $\tau_{ii'}$ respectively and where we have written only the upper left $2KL \times 2KL$ corner of $\Lambda_{ii'}$, from which the structure of the whole matrix is clear.

We shall reduce the exponent in (6.2) to a simple form by applying successively three symmetric orthogonal transformations based on the matrix defined in Lemma 1. More precisely, the matrices P_1, P_2, and P below are defined like the matrix P of Lemma 1, with n taking on the values L, K and J respectively.

First, let

$$Z = P_1^* Y, \qquad P_1^* = ((\text{diag.}\, P_1, \cdots, P_1)),$$

where P_1^* of size $IJKL$ consists of IJK diagonal blocks P_1 of size L. Then

$$z_{ijkl} = L^{\frac{1}{2}} y_{ijk\cdot}, \qquad Ez_{ijkl} = \delta_{1l} L^{\frac{1}{2}} \beta_i, \qquad \Sigma_Z = P_1^* \Lambda P_1^* + \sigma_e^2 U.$$

The matrix $P_1^* \Lambda P_1^*$ has the same structure as Λ, except that e.g., $F_{ii'}$ has to be replaced by $P_1 F_{ii'} P_1$, the only nonzero element of which is the 1,1-element, which equals $L\sigma_{ii'}$. Therefore,

$$(6.5) \quad \begin{aligned} (Y - EY)'\Sigma_Y^{-1}(Y - EY) &= (Z^{(1)} - EZ^{(1)})'\Sigma_{Z^{(1)}}^{-1}(Z^{(1)} - EZ^{(1)}) \\ &\quad + \sigma_e^{-2} \sum_i \sum_j \sum_k \sum_{l>1} z_{ijkl}^2, \end{aligned}$$

where the vector $Z^{(1)}$ of dimension IJK has components $z_{ijk}^{(1)} = z_{ijk1}$ which we order as in (5.8). Writing $\Sigma_{Z^{(1)}} = L((\Phi_{ii'})) + \sigma_e^2 U$ (U is now of size IJK and each $\Phi_{ii'}$ of size JK), one gets for $\Phi_{ii'}$ a matrix like (6.4), but with entries $\sigma_{ii'}$, $\nu_{ii'}$ and $\tau_{ii'}$ instead of $F_{ii'}$, $G_{ii'}$ and $H_{ii'}$ respectively. Next, let

$$Z^{(2)} = P_2^* Z^{(1)}, \qquad P_2^* = ((\text{diag. } P_2, \cdots, P_2)), \qquad P_2 = ((p_{kk'}^{(2)})),$$

where P_2^* of size IJK has IJ diagonal blocks P_2 of size K. Then

$$(6.6) \quad \begin{aligned} z_{ij1}^{(2)} &= K^{\frac{1}{2}} z_{ij\cdot}^{(1)} = (KL)^{\frac{1}{2}} y_{ij\cdot\cdot}, \qquad & z_{ijk}^{(2)} &= L^{\frac{1}{2}} \sum_k p_{kk'}^{(2)} y_{ijk'\cdot} \quad \text{for } k > 1, \\ Ez_{ij1}^{(2)} &= (KL)^{\frac{1}{2}} \beta_i, \qquad & Ez_{ijk}^{(2)} &= 0 \quad \text{for } k > 1. \end{aligned}$$

Furthermore, writing $\Sigma_{Z^{(2)}} = L((\Psi_{ii'})) + \sigma_e^2 U$, one finds that if $\Psi_{ii'}$ is in turn partitioned into J^2 submatrices of size K, then it has the structure (2.1) with $n = J$ and a and b replaced respectively by the submatrices of size K

$$\begin{bmatrix} \sigma_{ii'} + (K-1)\nu_{ii'} & 0 & & 0 \\ 0 & \sigma_{ii'} - \nu_{ii'} & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & & \sigma_{ii'} - \nu_{ii'} \end{bmatrix} \text{ and } \begin{bmatrix} \tau_{ii'} & 0 & \cdots & 0 \\ 0 & \tau_{ii'} & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & \tau_{ii'} \end{bmatrix}.$$

Finally, let

$$X = P_3^* Z^{(2)}, \qquad P_3^* = ((\text{diag. } P_3, \cdots, P_3)), \qquad P_3 = ((P_{jj'}^*)),$$
$$P_{jj'}^* = ((\text{diag. } p_{jj'}, \cdots, p_{jj'})) \quad \text{with} \quad ((p_{jj'})) = P,$$

where P_3^* consists of I diagonal blocks P_2 of size JK, $P_{jj'}^*$ is of size K and P of size J is as in Lemma 1. Then

$$(6.7) \quad \begin{aligned} x_{i11} &= J^{\frac{1}{2}} z_{i\cdot 1}^{(2)} = (JKL)^{\frac{1}{2}} y_{i\cdots}, \qquad & x_{ijk} &= \sum_j p_{jj'} z_{ij'k}^{(2)} \quad \text{for } (j,k) \neq (1,1), \\ Ex_{i11} &= (JKL)^{\frac{1}{2}} \beta_i, \qquad & Ex_{ijk} &= 0 \quad \text{for } (j,k) \neq (1,1). \end{aligned}$$

Writing

$$(6.8) \qquad \Sigma_X = L((\Delta_{ii'})),$$

one finds that $\Delta_{ii'} = P_3 \Psi_{ii'} P_3$ can be written

$$(6.9) \qquad \Delta_{ii'} = \begin{bmatrix} A_{ii'} & 0 & \cdots & 0 \\ 0 & B_{ii'} & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & B_{ii'} \end{bmatrix} \begin{matrix} (j=1) \\ (j=2), \\ \\ (j=J) \end{matrix}$$

where

$$(6.10)\qquad A_{ii'} = \begin{bmatrix} r_{ii'} & 0 & \cdots & 0 \\ 0 & v_{ii'} & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & v_{ii'} \end{bmatrix}, \qquad B_{ii'} = \begin{bmatrix} u_{ii'} & 0 & \cdots & 0 \\ 0 & w_{ii'} & \cdots & 0 \\ 0 & \cdot & \cdots & \cdot \\ \cdot & 0 & \cdots & w_{ii'} \end{bmatrix}$$

are submatrices of size K, the elements of which are defined by the relations

$$(6.11)\qquad \sigma_{ii'}^{+} = \sigma_{ii'} + \delta_{ii'}L^{-1}\sigma_e^2$$

and

$$(6.12)\qquad \begin{aligned} r_{ii'} &= \sigma_{ii'}^{+} + (K-1)\nu_{ii'} + (J-1)\tau_{ii'}, \qquad & u_{ii'} &= \sigma_{ii'}^{+} + (K-1)\nu_{ii'} - \tau_{ii'}, \\ v_{ii'} &= \sigma_{ii'}^{+} - \nu_{ii'} + (J-1)\tau_{ii'}, \qquad & w_{ii'} &= \sigma_{ii'}^{+} - \nu_{ii'} - \tau_{ii'}. \end{aligned}$$

Define the matrices of size I,

$$(6.13)\qquad R_0 = ((r_{ii'})), \qquad U_0 = ((u_{ii'})), \qquad V_0 = ((v_{ii'})), \qquad W_0 = ((w_{ii'})).$$

Then

$$(6.14)\qquad R_0 + W_0 = U_0 + V_0 .$$

Define also a set of JK vectors, all of dimension I, as follows:

$$(6.15)\qquad \begin{aligned} R &= L^{-\frac{1}{2}}(x_{111}, \cdots, x_{i11}, \cdots, x_{I11})', & & \\ U_j &= L^{-\frac{1}{2}}(x_{1j1}, \cdots, x_{ij1}, \cdots, x_{Ij1})', & j &> 1, \\ V_k &= L^{-\frac{1}{2}}(x_{11k}, \cdots, x_{i1k}, \cdots, x_{I1k})', & k &> 1, \\ W_{jk} &= L^{-\frac{1}{2}}(x_{1jk}, \cdots, x_{ijk}, \cdots, x_{Ijk})', & j, k &> 1. \end{aligned}$$

The equations (6.8), (6.9) and (6.10) show that these JK vectors are uncorrelated. Their covariance matrices are respectively R_0, U_0, V_0 and W_0 and, according to (6.7),

$$(6.16)\qquad EU_j = EV_k = EW_{jk} = 0, \qquad j, k > 1.$$

The quadratic form in $Z^{(1)}$ in the right-hand member of (6.5) can then be written, as one easily verifies,

$$(6.17)\qquad \begin{aligned} (Z^{(1)} - EZ^{(1)})'\Sigma_{Z^{(1)}}^{-1}(Z^{(1)} - EZ^{(1)}) &= (X - EX)'\Sigma_X^{-1}(X - EX) \\ &= (R - ER)'R_0^{-1}(R - ER) + \sum_{j>1} U_j'U_0^{-1}U_j \\ &\quad + \sum_{k>1} V_k'V_0^{-1}V_k + \sum_{j>1}\sum_{k>1} W_{jk}'W_0^{-1}W_{jk}, \end{aligned}$$

the last three terms of which also equal

$$(6.18)\qquad \operatorname{tr}\,[U_0^{-1}\sum_{j>1} U_jU_j' + V_0^{-1}\sum_{k>1} V_kV_k' + W_0^{-1}\sum_{j>1}\sum_{k>1} W_{jk}W_{jk}'].$$

When normality is assumed it follows from this, (6.5) and the Neyman factorization theorem, that a sufficient statistic for the parameters of the model is

$$T = \{s^2, \hat{\beta}, \hat{U}_0, \hat{V}_0, \hat{W}_0\}, \tag{6.19}$$

where

$$s^2 = [IJK(L-1)]^{-1} \sum_i \sum_j \sum_k \sum_{l>1} z_{ijkl}^2, \qquad \hat{\beta} = (JK)^{-\frac{1}{2}} R,$$

$$\hat{U}_0 = (J-1)^{-1} \sum_{j>1} U_j U_j', \qquad \hat{V}_0 = (K-1)^{-1} \sum_{k>1} V_k V_k', \tag{6.20}$$

$$\hat{W}_0 = [(J-1)(K-1)]^{-1} \sum_{j>1} \sum_{k>1} W_{jk} W_{jk}'.$$

Using the intermediary relations (6.6), (6.7) and applying (2.2), one finds after some straightforward algebra that T can be expressed in terms of the observations as follows

$$s^2 = MS_e, \qquad \hat{\beta}_i = y_{i\cdots},$$

$$\hat{u}_{ii'} = (J-1)^{-1} K \sum_j (y_{ij\cdot\cdot} - y_{i\cdots})(y_{i'j\cdot\cdot} - y_{i'\cdots}),$$

$$\hat{v}_{ii'} = J(K-1)^{-1} \sum_k (y_{i\cdot k\cdot} - y_{i\cdots})(y_{i'\cdot k\cdot} - y_{i'\cdots}), \tag{6.21}$$

$$\hat{w}_{ii'} = (J-1)^{-1}(K-1)^{-1} \sum_j \sum_k (y_{ijk\cdot} - y_{ij\cdot\cdot} - y_{i\cdot k\cdot} + y_{i\cdots})$$

$$(y_{i'jk\cdot} - y_{i'j\cdot\cdot} - y_{i'\cdot k\cdot} + y_{i'\cdots}),$$

where we write $\hat{\beta} = (\hat{\beta}_1, \cdots, \hat{\beta}_I)'$, $\hat{U}_0 = ((\hat{u}_{ii'}))$, $\hat{V}_0 = ((\hat{v}_{ii'}))$, $\hat{W}_0 = ((\hat{w}_{ii'}))$.

7. Unbiased point estimators of the parameters. The results of the previous section enable us to find unbiased point estimators of the basic parameters of the model, namely μ, α_i, σ_e^2, $\sigma_{ii'}$, $\nu_{ii'}$, $\tau_{ii'}$, $i' \leqq i = 1, \cdots, I$. We shall also prove that if $J, K > I$ and normality is assumed, those estimators are minimum variance unbiased.

Unbiased estimators for μ, α_i, σ_e^2 are at once found to be

$$\hat{\mu} = y_{\cdots\cdot}, \qquad \hat{\alpha}_i = y_{i\cdots} - y_{\cdots\cdot}, \qquad \hat{\sigma}_e^2 = s^2 = MS_e. \tag{7.1}$$

Also, as noticed above (6.16), $\Sigma_{U_j} = U_0$, $j > 1$. This, together with (6.16), (6.20) and the similar relations in V and W shows that

$$E\hat{U}_0 = U_0, \qquad E\hat{V}_0 = V_0, \qquad E\hat{W}_0 = W_0. \tag{7.2}$$

Solving the last three equations of (6.12) for the unknown parameters $\sigma_{ii'}$, $\nu_{ii'}$, $\tau_{ii'}$ and substituting in the resulting equalities their estimators s^2, $\hat{u}_{ii'}$, $\hat{v}_{ii'}$ and $\hat{w}_{ii'}$ for σ_e^2, $u_{ii'}$, $v_{ii'}$ and $w_{ii'}$, one obtains the unbiased estimators

$$\hat{\sigma}_{ii'} = (JK)^{-1}[(JK - J - K)\hat{w}_{ii'} + J\hat{u}_{ii'} + K\hat{v}_{ii'}] - \delta_{ii'} L^{-1} s^2, \tag{7.3}$$

$$\hat{\nu}_{ii'} = K^{-1}(\hat{u}_{ii'} - \hat{w}_{ii'}), \qquad \hat{\tau}_{ii'} = J^{-1}(\hat{v}_{ii'} - \hat{w}_{ii'}).$$

Performing some algebra one verifies also that the estimators (4.4) of the components of variance coincide with the estimators one would obtain by substituting $\hat{\sigma}_{ii'}$, $\hat{\nu}_{ii'}$ and $\hat{\tau}_{ii'}$ for $\sigma_{ii'}$, $\nu_{ii'}$ and $\tau_{ii'}$ in the relations (3.9). This remark is needed to conclude that the estimators (4.4) also possess the optimum property to be considered now.

Assume that normality holds and $J, K > I$. Then, the estimators (4.4), (7.1) and (7.3) are minimum variance unbiased. To see this, the multivariate extension of a completeness lemma of Gautschi [4] is needed.

LEMMA 5. *Let Θ be a parameter vector and Y be a random vector in Euclidian space E_n, similarly let Θ_1 and Y_1 be vectors in E_{n_1}. Assume that Y and Y_1 have probability densities (with respect to Lebesgue measure) of the form*

$$p(Y, \Theta) = g(\Theta)h(Y) \exp\{\Theta'Y\},$$

$$p_1(Y_1, \Theta_1, \Theta) = f(\Theta_1, \Theta) \exp\{Y_1'R(\Theta)Y_1 + \Theta_1'Y_1\},$$

where $R(\Theta)$ is a matrix of size n_1. Let the domain $\mathfrak{D}$ of Θ contain a nondegenerate interval in E_n and the domain of Θ_1 be E_{n_1}. Then, the family of product measures on E_{n+n_1} generated by the family of probability densities

$$\mathfrak{J} = \{p(Y, \Theta)p_1(Y_1, \Theta_1, \Theta) : (\Theta, \Theta_1) \in \mathfrak{D} \times E_{n_1}\}$$

is strongly complete, in the sense of Lehmann and Scheffé [7].

The proof of this can be made along exactly the same lines as in the univariate case [4]. Consider now the probability density of the statistic T defined by (6.19). The vectors (6.15) being independent, s^2, $\hat{\beta}$, $\hat{U}_0$, $\hat{V}_0$ and $\hat{W}_0$ are also independent. Now $IJK(L - 1)s^2$ is $\sigma_e^2\chi^2_{IJK(L-1)}$, $\hat{\beta}$ is $N(\beta, J^{-1}K^{-1}R_0)$ and when $J, K > I$, $\hat{U}_0$, $\hat{V}_0$, $\hat{W}_0$ have respectively the $W([J - 1]^{-1}U_0, J - 1)$, $W([K - 1]^{-1}V_0, K-1)$ and $W([J - 1)(K - 1)]^{-1}W_0, (J - 1)(K - 1))$ distributions, where $W(\Sigma, n)$ denotes the Wishart distribution of the matrix $\sum_1^n Y_iY_i'$ for n independent identically distributed vectors $Y_1, \cdots, Y_n$, each $N(0, \Sigma)$. Let -2Θ be a vector in which the components of the matrices $(J - 1)U_0^{-1}$, $(K - 1)V_0^{-1}$, $(J - 1)(K - 1)W_0^{-1}$ and σ_e^{-2} are strung out, Y be a vector in which the components of the matrices $\hat{U}_0$, $\hat{V}_0$, $\hat{W}_0$ and s^2 are correspondingly strung out, $R(\Theta) = -\frac{1}{2}JKR_0^{-1}$, $\Theta_1' = JK\beta'R_0^{-1}$ and $Y_1 = \hat{\beta}$. By writing it out fully, one can then verify that the density p_T of T becomes

$$p_T = p(Y, \Theta)p_1(Y_1, \Theta_1, \Theta), \tag{7.4}$$

where the two factors are of the type considered in Lemma 5. The family of probability measures generated by the densities (7.4) is, therefore, strongly complete and the unbiased estimators (4.4), (7.1) and (7.3), which are functions of T, are minimum variance unbiased ([6], Theorem 5.1).

When normality is assumed, the variances of the estimators $\hat{\sigma}_e^2$ and $\hat{\alpha}_i$ can be estimated unbiasedly. One verifies at once that Var $\hat{\sigma}_e^2$ has the unbiased estimator $2\hat{\sigma}_e^4/(\nu_e + 2)$, where $\nu_e = IJK(L - 1)$. One can show that an unbiased estimator of the variance of $\hat{\alpha}_i$ is

$$\hat{\sigma}^2_{\hat{\alpha}_i} = J^{-1}(J-1)^{-1}R_i^2 + K^{-1}(K-1)^{-1}S_i^2 - (JK)^{-1}(J-1)^{-1}(K-1)^{-1}T_i^2, \tag{7.5}$$

where

$$R_i^2 = \sum_j (y_{ij\cdot\cdot} - y_{i\cdot\cdot\cdot} - y_{\cdot j\cdot\cdot} + y_{\cdot\cdot\cdot\cdot})^2,$$

$$S_i^2 = \sum_k (y_{i\cdot k\cdot} - y_{i\cdot\cdot\cdot} - y_{\cdot\cdot k\cdot} + y_{\cdot\cdot\cdot\cdot})^2,$$

$$T_i^2 = \sum_j \sum_k (y_{ijk\cdot} - y_{ij\cdot\cdot} - y_{i\cdot k\cdot} - y_{\cdot jk\cdot} + y_{i\cdot\cdot\cdot} + y_{\cdot j\cdot\cdot} + y_{\cdot\cdot k\cdot} - y_{\cdot\cdot\cdot\cdot})^2.$$

8. A test for the hypothesis H_A. In a practical situation to which the model equation (3.1) and our basic assumptions, including the normality assumption, can be applied, the statistical hypothesis of most interest is likely to be that of no fixed main effects, namely $H_A: \alpha_1 = \cdots = \alpha_I = 0$. In the model involving only two factors, Scheffé [10] shows that a T^2 statistic can be constructed for testing H_A. The extension of his procedure to the present model would require that the JK vectors $(y_{1jk\cdot}, \cdots, y_{Ijk\cdot})'$, $j = 1, \cdots, J$, $k = 1, \cdots, K$ be independent. As (6.3) shows this does not hold true unless $\nu_{ii'} = \tau_{ii'} = 0$, all i, i', an additional assumption which cannot be justified with the present model. The likelihood ratio principle is here of no help either, as already remarked by Wilks ([14], p. 259) in a simpler situation. This is due to the fact that the covariance matrix defined by (6.3) is not diagonal and that H_A does not completely specify Ey_{ijkl}. One can avoid this indeterminacy by following a suggestion of Hsu [5], namely by introducing the differences

$$d_{rjk} = y_{rjk\cdot} - y_{Ijk\cdot}, \qquad r = 1, \cdots, I-1, \qquad \text{all } j, k. \tag{8.1}$$

Using a notation analogous to that of Section 6, let

$$x^*_{rjk} = x_{rjk} - x_{Ijk}, \qquad r = 1, \cdots, I-1, \qquad \text{all } j, k. \tag{8.2}$$

We define vectors R^*, U^*_j, V^*_k, W^*_{jk} $(j, k > 1)$ by relations similar to (6.15) but with x^*_{rjk} substituted for x_{rjk}, e.g.,

$$R^* = L^{-\frac{1}{2}}(x^*_{111}, \cdots, x^*_{r11}, \cdots, x^*_{I-1,11})'. \tag{8.3}$$

These JK vectors are thus of dimension $I - 1$ only; like the vectors (6.15), they are independent. Their covariance matrices are respectively

$$R^*_0 = ((r^*_{rr'})) = ((r_{rr'} - r_{rI} - r_{Ir'} + r_{II})), \quad r, r' = 1, \cdots, I-1, \tag{8.4}$$

and U^*_0, V^*_0, W^*_0 whose elements $u^*_{rr'}$, $v^*_{rr'}$, and $w^*_{rr'}$ are similarly defined in terms of the $u_{ii'}$, $v_{ii'}$ and $w_{ii'}$ of (6.12). Let also

$$\beta^* = (\beta_1 - \beta_I, \cdots, \beta_{I-1} - \beta_I)', \qquad \hat{\beta}^* = (JK)^{-\frac{1}{2}}R^*, \tag{8.5}$$

so that $E\hat{\beta}^* = \beta^*$, and define as in (6.20) the matrices

$$\hat{U}_0^* = (J-1)^{-1}\sum_{j>1} U_j^* U_j^{*\prime}, \qquad \hat{V}_0^* = (K-1)^{-1}\sum_{k>1} V_k^* V_k^{*\prime}, \tag{8.6}$$
$$\hat{W}_0^* = (J-1)^{-1}(K-1)^{-1}\sum_{j>1}\sum_{k>1} W_{jk}^* W_{jk}^{*\prime}.$$

Formulas (6.14) and (7.2) now become

$$R_0^* + W_0^* = U_0^* + V_0^*, \tag{8.7}$$

$$E\hat{U}_0^* = U_0^*, \qquad E\hat{V}_0^* = V_0^*, \qquad E\hat{W}_0^* = W_0^*. \tag{8.8}$$

The elements $\hat{u}_{rr'}^*$, $\hat{v}_{rr'}^*$ and $\hat{w}_{rr'}^*$ of the matrices (8.6) can be computed from the observations by using relations analogous to (6.21) but with d_{rjk} substituted for y_{rjk}. Alternatively

$$\hat{u}_{rr'}^* = \hat{u}_{rr'} - \hat{u}_{rI} - \hat{u}_{Ir'} + \hat{u}_{II}, \quad r, r' = 1, \cdots, I-1,$$

with similar relations in v and w.

If the covariance matrix R_0^* were known, one would test the hypothesis H_A by using the criterion

$$T_1^2 = R^{*\prime}(R_0^*)^{-1}R^* \tag{8.9}$$

which has the noncentral χ^2 distribution with $I - 1$ d.f. and noncentrality parameter

$$\delta_1^2 = JK\beta^{*\prime}(R_0^*)^{-1}\beta^*, \tag{8.10}$$

which reduces to zero under H_A.

When R_0^* is unknown, one might think of using instead of T_1^2 the criterion $T_2 = R^{*\prime}(\hat{R}_0^*)^{-1}R^*$, where $\hat{R}_0^*$ is the unbiased estimate of R_0^* based on the sufficient statistic T of (6.19), i.e., by (8.7), (8.8), $\hat{R}_0^* = \hat{U}_0^* + \hat{V}_0^* - \hat{W}_0^*$. Although R_0^*, $\hat{U}_0^*$, $\hat{V}_0^*$, $\hat{W}_0^*$ are mutually independent, the former with multivariate normal and the latter three with Wishart distributions, it does not seem possible to obtain the distribution of T_2. The case $I = 2$ easily shows that it does, under H_A, depend on nuisance parameters, and unlike a T^2 statistic, is not nonnegative. However, when both J and K tend to infinity, one verifies that the limiting distribution of T_2 under H_A is the χ_{I-1}^2 distribution. Hence for large values of both J and K, a satisfactory test of H_A at the α level of significance consists in rejecting the hypothesis if T_2 exceeds the upper α-point of the χ_{I-1}^2 distribution. As is well known, one does not need to compute $(\hat{R}_0^*)^{-1}$ in order to evaluate T_2, but can use a formula similar to (8.21) below.

Consider now the case where J and K are not large enough to justify the use of the χ^2 approximation. Assume, however, that $J, K \geqq I$ (in Section 7, where we had vectors of dimension I, we assumed $J, K > I$. Here the vectors have dimension $I - 1$, so we need only $J, K \geqq I$). The fact that several unknown covariance matrices are involved in our model suggests trying to apply a device similar to the one proposed by Scheffé [8] for solving the Behrens-Fisher problem: Instead of using the estimate $\hat{R}_0^*$ of R_0^* in the definition of T_2, one would like to use one which has a Wishart distribution. This would require finding independent identically distributed vectors $S_1, \cdots, S_n$, each a linear combination of the vectors of observations $(d_{ijk}, \cdots, d_{I-1jk})'$, $j = 1, \cdots, J$, $k = 1, \cdots, K$, and each having mean zero and covariance matrix R_0^*. One would wish to do so for n as close as possible to $JK - 1$, which is as much as one can achieve when said vectors of observations are independent identically distributed. But $S_1, \cdots, S_n$ would then also have to be linear combinations of the vectors R^*, U_j^*, V_k^* and W_{jk}^*, $j, k > 1$. Now because of the minus sign in $R_0^* = U_0^* + V_0^* - W_0^*$, it is clear that no such linear combination except R^* itself has covariance matrix R_0^*.

It appears that the only way to construct a test criterion which has under H_A a distribution free of nuisance parameters consists in looking not at the minimum variance unbiased estimate R^* of $(JK)^{\frac{1}{2}}\beta^*$, but at another unbiased estimate of it, namely

$$(8.11) \qquad (JK)^{\frac{1}{2}}S = R^* + [(J-1)(K-1)]^{-\frac{1}{2}}\sum_{j>1}\sum_{k>1} W_{jk}^*.$$

Although this will result in a loss of power of the test obtained below as compared with the "ideal" test described above, this is the price one has to pay for allowing three unknown covariance matrices in the model. The vector $(JK)^{\frac{1}{2}}S$ has the $N((JK)^{\frac{1}{2}}\beta^*, R_0^* + W_0^*)$ distribution. Let

$$(8.12) \qquad H = (M-1)^{-1}\sum_{m=2}^{M} (U_m^* + V_m^*)(U_m^* + V_m^*)',$$

where $M = \min(J, K)$, then independence of the vectors R^*, U_j^*, V_k^*, W_{jk}^* together with (8.7) shows at once that a T^2 criterion for testing H_A is

$$(8.13) \qquad T^2 = JK \cdot S'H^{-1}S.$$

More precisely, $\mathfrak{F} = (M-1)^{-1}(I-1)^{-1}(M-I+1)T^2$ has under H_A the F-distribution with $I - 1$ and $M - I + 1$ d.f. Under alternatives it has the corresponding noncentral F-distribution (defined in [10], formula (82)) with noncentrality parameter

$$(8.14) \qquad \delta^2 = JK\beta^{*\prime}(R_0^* + W_0^*)^{-1}\beta^*.$$

The test consists in rejecting H_A if $\mathfrak{F} > F_\alpha$, the upper α-point of the F-distribution with the above numbers of d.f., α being the level of significance. The "ideal" test we were imagining above would have a distribution with noncentrality parameter δ_1^2 given by (8.10). Going back to formulas (6.12) we see that if the differences $\sigma_{ii'}^+ - \nu_{ii'} - \tau_{ii'}$ are small compared to $K\nu_{ii'} + J\tau_{ii'}$, then δ^2 should not be appreciably smaller than δ_1^2 and so the loss in power due to the larger variance of $(JK)^{\frac{1}{2}}S$ should not be too considerable. Some better insight into this will be obtained in connection with (8.23). However, if $M - I + 1$ is small, say it equals only 2 or 3, then the drastic curtailing in the number of d.f. "for error" as compared with the "ideal" number $JK - I + 1$ will make the test a rather poor one.

We express now the criterion T^2 in terms of the observations y_{ijkl}. Using (6.6), (6.7) and the values of $p_{ii'}$ in Lemma 1 with the appropriate values for n yields

$$(8.15)\quad z_{ijk}^{(2)} = L^{\frac{1}{2}}[y_{ijk\cdot} - (K^{\frac{1}{2}} - 1)^{-1}(K^{\frac{1}{2}}y_{ij\cdot\cdot} - y_{ij1\cdot})] \quad \text{for } k > 1, \qquad \text{all } i, j.$$

Similarly one has

$$(8.16)\quad x_{ijk} = z_{ijk}^{(2)} - (J^{\frac{1}{2}} - 1)^{-1}(J^{\frac{1}{2}}z_{i\cdot k}^{(2)} - z_{i1k}^{(2)}) \quad \text{for } (j, k) \neq (1, 1), \qquad \text{all } i.$$

The above two formulas and (6.7) easily yield

$$(8.17)\quad \begin{aligned} x_{i11} &= (JKL)^{\frac{1}{2}}y_{i\cdots}, \\ x_{i1k} &= (JL)^{\frac{1}{2}}[y_{i\cdot k\cdot} - (K^{\frac{1}{2}} - 1)^{-1}(K^{\frac{1}{2}}y_{i\cdots} - y_{i\cdot 1\cdot})], && k > 1, \\ x_{ij1} &= (KL)^{\frac{1}{2}}[y_{ij\cdot\cdot} - (J^{\frac{1}{2}} - 1)^{-1}(J^{\frac{1}{2}}y_{i\cdots} - y_{i1\cdot\cdot})], && j > 1. \\ x_{ijk} &= L^{\frac{1}{2}}[y_{ijk\cdot} - (J^{\frac{1}{2}} - 1)^{-1}(J^{\frac{1}{2}}y_{i\cdot k\cdot} - y_{i1k\cdot}) \\ &\quad - (K^{\frac{1}{2}} - 1)^{-1}(K^{\frac{1}{2}}y_{ij\cdot\cdot} - y_{ij1\cdot}) \\ &\quad + (J^{\frac{1}{2}} - 1)^{-1}(K^{\frac{1}{2}} - 1)^{-1}((JK)^{\frac{1}{2}}y_{i\cdots} \\ &\qquad - (J^{\frac{1}{2}}y_{i\cdot 1\cdot} - K^{\frac{1}{2}}y_{i1\cdot\cdot} + y_{i11\cdot})], && j, k > 1. \end{aligned}$$

The equations (8.17) also imply

$$\sum_{j>1}\sum_{k>1} x_{ijk} = J^{\frac{1}{2}}\sum_{k>1}(z_{i1k}^{(2)} - z_{i\cdot k}^{(2)}),$$

from which one finds by (8.15)

$$(8.18)\qquad \sum_{j>1}\sum_{k>1} x_{ijk} = (JKL)^{\frac{1}{2}}(y_{i11\cdot} - y_{i1\cdot\cdot} - y_{i\cdot 1\cdot} + y_{i\cdots}).$$

Let, for convenience, in (8.11), (8.12),

$$(8.19)\quad \begin{gathered} S = (s_1, \cdots, s_r, \cdots, s_{I-1})', \\ H = (M - 1)^{-1}G, \qquad G = ((g_{rr'})), \qquad g_{rr'} = \sum_{m=2}^{M} f_{rm}f_{r'm}, \end{gathered}$$

where f_{rm} is the rth component of the vector $U_m^* + V_m^*$. Using (8.1), (8.2), (8.3), (8.17), (8.18) shows that one can write, for $r = 1, \cdots, I - 1$ and $m = 2, \cdots, M$,

$$(8.20)\quad \begin{aligned} s_r &= d_{r\cdot\cdot} + [(J - 1)(K - 1)]^{-\frac{1}{2}}(d_{r11} - d_{r1\cdot} - d_{r\cdot 1} + d_{r\cdot\cdot}), \\ f_{rm} &= K^{\frac{1}{2}}\,d_{rm\cdot} + J^{\frac{1}{2}}\,d_{r\cdot m} - K^{\frac{1}{2}}(J^{\frac{1}{2}} - 1)^{-1}(J^{\frac{1}{2}}\,d_{r\cdot\cdot} - d_{r1\cdot}) \\ &\qquad - J^{\frac{1}{2}}(K^{\frac{1}{2}} - 1)^{-1}(K^{\frac{1}{2}}\,d_{r\cdot\cdot} - d_{r\cdot 1}). \end{aligned}$$

To compute T^2 it is not necessary to invert the matrix G: One can use instead of (8.13) the relation

$$T^2 = JK(M-1)\left[\frac{|((g_{rr'} + s_r s_{r'}))|}{|G|} - 1\right]. \tag{8.21}$$

There is in general no shortcut available for the computation of T^2: One has to compute separately the s_r and f_{rm} from (8.20), then use (8.19) to compute T^2 from (8.21). If, however, $J = K = M$, one verifies that

$$g_{rr'} = M\sum_{m=1}^{M} (d_{rm\cdot} + d_{r\cdot m} - 2d_{r\cdot\cdot})(d_{r'm\cdot} + d_{r'\cdot m} - 2d_{r'\cdot\cdot}),$$

so that the computational work required is considerably reduced.

There is in principle no difficulty in evaluating the power of the test against a specific alternative. All that is needed is the value of the noncentrality parameter δ^2 of (8.14). On the other hand, specifying an alternative requires specifying the value of $R_0^* + W_0^* = U_0^* + V_0^*$. This in general would have to be estimated from the data by using the estimates (8.6). Unless $J = K = M$, in which case $\hat{U}_0^* + \hat{V}_0^* = (M-1)^{-1}G$, it might require a prohibitive amount of additional computations. One might then be satisfied to determine the power of the test against a "simplified" alternative, namely one obtained by assuming that the three basic covariance matrices (3.3) all have the structure (2.1). From (6.11), (6.12) and the relations analogous to (8.4), the r, r'-element of $U_0^* + V_0^*$ is found to be, under this simplifying assumption,

$$\gamma_{rr'}^* = (1 + \delta_{rr'})(L^{-1}\sigma_e^2 + \gamma_0 - \gamma_1),$$

where $\gamma_0 = 2\rho_{ii} + K\nu_{ii} + J\tau_{ii}$, $\gamma_1 = 2\rho_{ii'} + K\nu_{ii'} + J\tau_{ii'}$, $i \neq i'$. Using (3.9) this becomes

$$\gamma_{rr'}^* = (1 + \delta_{rr'})(K\sigma_{AB}^2 + J\sigma_{AC}^2 + 2\sigma_{ABC}^2 + L^{-1}\sigma_e^2).$$

The matrix $(((\gamma^*)^{rr'})) = ((\gamma_{rr'}^*))^{-1}$ can easily be computed and one finds

$$\delta^2 = JK\sum_r\sum_{r'}\beta_r^*\beta_{r'}^*(\gamma^*)^{rr'} = JK\frac{\sum_r \beta_r^{*2} - I^{-1}(\sum_r \beta_r^*)^2}{K\sigma_{AB}^2 + J\sigma_{AC}^2 + 2\sigma_{ABC}^2 + L^{-1}\sigma_e^2},$$

which in terms of $\alpha_1, \cdots, \alpha_I$ becomes simply

$$\delta^2 = JK(K\sigma_{AB}^2 + J\sigma_{AC}^2 + 2\sigma_{ABC}^2 + L^{-1}\sigma_e^2)^{-1}\sum_i \alpha_i^2. \tag{8.22}$$

When $\Sigma_i\alpha_i^2$ is specified, this value of δ^2 can be very quickly estimated by using (4.4) with

$$\hat{\delta}^2 = JKL(MS_{AB} + MS_{AC} - MS_e)^{-1}\sum_i \alpha_i^2.$$

It is interesting to compare (8.22) with the analogous formula that one obtains when computing the value of the "ideal" noncentrality parameter δ_I^2 of

(8.10) under our simplifying assumption. One finds that

$$\delta_1^2 = (1 + \Delta)\delta^2, \tag{8.23}$$

with

$$\Delta = (K\sigma_{AB}^2 + J\sigma_{AC}^2 + \sigma_{ABC}^2 + L^{-1}\sigma_e^2)^{-1}\sigma_{ABC}^2. \tag{8.24}$$

The conclusion we had arrived at in the discussion below (8.14) can now be put into the following terms: If Δ is small compared to 1, the loss of power introduced in the test by the undesirable last term of (8.11) should not be appreciable. This conclusion is encouraging. In practice, σ_{ABC}^2 is often dominated by one of σ_{AB}^2, σ_{AC}^2, which according to (8.24) should make Δ satisfactorily small. Estimating separately numerator and denominator, one obtains an estimate of Δ, namely

$$\tilde{\Delta} = (MS_{AB} + MS_{AC} - MS_{ABC})^{-1}(MS_{ABC} - MS_e),$$

which is of course not unbiased.

9. Confidence intervals for the fixed main effects. We consider briefly in this section various confidence statements that can be made concerning the parameters $\alpha_1, \cdots, \alpha_I$. Corresponding to the fact that no Hotelling T^2 test of H_A could be constructed on the basis of the estimates $\hat{\alpha}_i$ of the α_i's is the fact here that no confidence interval for α_i can be obtained in the ordinary manner by using the ratio $\hat{\alpha}_i/\hat{\sigma}_{\hat{\alpha}_i}$. In fact, (7.5) shows that $\hat{\sigma}_{\hat{\alpha}_i}^2$ is not even a positive indefinite quadratic form.

By analogy with (8.11), let $m_i = [(J-1)(K-1)L]^{-\frac{1}{2}}\Sigma_{j>1}\Sigma_{k>1}x_{ijk}$, where x_{ijk} is as in (6.15). Then

$$\text{Var}\,(m_i - m.) = w_{ii} - w_{i\cdot} - w_{\cdot i} + w_{\cdot\cdot} = \sigma_{ii}^0 - \nu_{ii}^0 - \tau_{ii}^0 + (I-1)I^{-1}L^{-1}\sigma_e^2.$$

The unbiased estimate $(JK)^{\frac{1}{2}}\tilde{\alpha}_i = (JK)^{\frac{1}{2}}\hat{\alpha}_i + m_i - m.$ of $(JK)^{\frac{1}{2}}\alpha_i$ has variance $2\sigma_{ii}^0 + K\nu_{ii}^0 + J\tau_{ii}^0 + 2(I-1)I^{-1}L^{-1}\sigma_e^2$, as is easily verified. An unbiased estimate of this is $a_{ii} - 2a_{i\cdot} + a_{\cdot\cdot}$ where $a_{ii'}$ is the i, i'-element of the matrix $(M-1)^{-1}\Sigma_{m=2}^M(U_m + V_m)\cdot(U_m + V_m)'$ and where as in Section 8 $M = \min(J, K)$. Furthermore, $a_{ii} - 2a_{i\cdot} + a_{\cdot\cdot}$ is independent of $\tilde{\alpha}_i$. An exact confidence interval for the parameter α_i can therefore be based on the t-distribution with $M - 1$ d.f. of the ratio

$$[(JK)^{\frac{1}{2}}(\hat{\alpha}_i - \alpha_i) + m_i - m.]/[(a_{ii} - 2a_{i\cdot} + a_{\cdot\cdot})^{\frac{1}{2}}]. \tag{9.1}$$

In terms of the observations one has $\hat{\alpha}_i = y_{i\cdots} - y_{\cdots\cdot}$, then from (8.18) $m_i = [JK(J-1)^{-1}(K-1)^{-1}]^{\frac{1}{2}}(y_{i11\cdot} - y_{i1\cdot\cdot} - y_{i\cdot1\cdot} + y_{i\cdots})$, and finally from (8.17) $a_{ii'} = (M-1)^{-1}\sum_{m=2}^M e_{im}e_{i'm}$ where

$$e_{im} = K^{\frac{1}{2}}y_{im\cdot\cdot} + J^{\frac{1}{2}}y_{i\cdot m\cdot} - K^{\frac{1}{2}}(J^{\frac{1}{2}} - 1)^{-1}(J^{\frac{1}{2}}y_{i\cdots} - y_{i1\cdot\cdot}) - J^{\frac{1}{2}}(K^{\frac{1}{2}} - 1)^{-1}(y_{i\cdots} - y_{i\cdot1\cdot}).$$

When $J = K = M$, then $a_{ii} - 2a_{i\cdot} + a_{\cdot\cdot} = (M-1)^{-1}\sum_{m=2}^{M}(e_{im} - e_{\cdot m})^2$ reduces to

$$a_{ii} - 2a_{i\cdot} + a_{\cdot\cdot} = M(M-1)^{-1}\sum_{m=1}^{M}(y_{im\cdot\cdot} + y_{i\cdot m\cdot} - 2y_{i\cdots} - y_{\cdot m\cdot\cdot} - y_{\cdot\cdot m\cdot} + 2y_{\cdots\cdot})^2.$$

For a single difference $\alpha_i - \alpha_{i'}$ one can proceed in a similar manner and base an exact confidence interval on the t-distribution with $M - 1$ d.f. of the ratio

$$(9.2) \qquad [(JK)^{\frac{1}{2}}(y_{i\cdots} - y_{i'\cdots} - \alpha_i + \alpha_{i'}) + m_i - m_{i'}]/[(a_{ii} - 2a_{ii'} + a_{i'i'})^{\frac{1}{2}}]$$

The denominator can be computed from $a_{ii} - 2a_{ii'} + a_{i'i'} = (M-1)^{-1}\sum_{m=2}^{M}(e_{im} - e_{i'm})^2$ which reduces when $J = K = M$ to

$$a_{ii} - 2a_{ii'} + a_{i'i'} = M(M-1)^{-1}\sum_{m=1}^{M}(y_{im\cdot\cdot} + y_{i\cdot m\cdot} - 2y_{i\cdots} - y_{i'm\cdot\cdot} - y_{i'\cdot m\cdot} + 2y_{i'\cdots}).$$

Confidence statements based on (9.1) or (9.2) should be used only if a single statement is made and the particular α_i or $\alpha_i - \alpha_{i'}$ considered has not been suggested by the data. If several confidence statements are desired, Scheffé's method [9] of multiple comparison can be applied when $J, K \geqq I$ in a way similar to that described in [10], but again based on the nonoptimum unbiased estimate (8.11): We estimate a contrast $\theta = \sum_i h_i\alpha_i(\sum_i h_i = 0)$ with $\tilde{\theta} = \sum_{r=1}^{i-1} h_r[\hat{\beta}_r^* + (JK)^{-\frac{1}{2}}m_r^*]$, where $\hat{\beta}_r^* = \hat{\beta}_r - \hat{\beta}_I = d_{r\cdot\cdot}$ and $m_r^* = m_r - m_I$, $r = 1, \cdots, I-1$. The variance of $\tilde{\theta}$ is $\sigma^2(\tilde{\theta}) = (JK)^{-1}\sum_r\sum_{r'} h_r h_{r'}(r_{rr'}^* + w_{rr'}^*)$ and has the unbiased estimate $\hat{\sigma}^2(\tilde{\theta}) = (JK)^{-1}\sum_r\sum_{r'} h_r h_{r'} a_{rr'}^*$, where $a_{rr'}^* = a_{rr'} - a_{rI} - a_{Ir'} + a_{II}$. Then, the probability is $1 - \alpha$ that the totality of contrasts $\theta = \sum h_i\alpha_i$ simultaneously satisfy

$$(9.3) \qquad \tilde{\theta} - S\hat{\sigma}(\tilde{\theta}) \leqq \theta \leqq \tilde{\theta} + S\hat{\sigma}(\tilde{\theta}),$$

where the constant S can be computed from F_α, the upper α-point of the F-distribution with $I - 1$ and $M - I + 1$ d.f., through the relation

$$S^2 = (M-1)(I-1)(M-I+1)^{-1}F_\alpha.$$

The conclusion arrived at in Section 8 that the use of the nonoptimum estimate (8.11) in the T^2 criterion does not in general affect the power of the test too adversely implies here that the confidence intervals (9.3) and those based on (9.1), (9.2) are not considerably lengthened by the necessary introduction of the undesirable m_i in the estimates of the α_i.

10. Acknowledgment. The author wishes to express his gratitude to Professor Henry Scheffé for suggesting the present investigation and for his constant advice and suggestions while it was in progress. Improvements in the presentation of the paper have also resulted from helpful comments made by the referees.

REFERENCES

[1] C. A. Bennett and N. L. Franklin, *Statistical Analysis in Chemistry and the Chemical Industries*, John Wiley and Sons, New York, 1954.

[2] G. E. P. Box, "Some theorems on quadratic forms applied in the study of analysis of variance problems, I. Effect of inequality of variance in the one-way classification," *Ann. Math. Stat.*, Vol. 25 (1954), pp. 290–302.

[3] Jerome Cornfield and John W. Tukey, "Average values of mean squares in factorials," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 907–949.

[4] Werner, Gautschi, "Some remarks on Herbach's paper, 'Optimum nature of the F-test for model II in the balanced case'," *Ann. Math. Stat.*, Vol. 30 (1959), pp. 960–963.

[5] H. L. Hsu, "Notes on Hotelling's generalized T," *Ann. Math. Stat.*, Vol. 9 (1938), pp. 231–243.

[6] E. L. Lehmann and Henry Scheffé, "Completeness, similar regions, and unbiased estimation. Part I," *Sankhyā*, Vol. 10 (1950), pp. 305–340.

[7] E. L. Lehmann and Henry Scheffé, "Completeness, similar regions, and unbiased estimation. Part II," *Sankhyā*, Vol. 15 (1955), pp. 219–236.

[8] Henry Scheffé, "On solutions of the Behrens-Fisher problem, based on the t-distribution," *Ann. Math. Stat.*, Vol. 14 (1943), pp. 35–44.

[9] Henry Scheffé, "A method for judging all contrasts in the analysis of variance," *Biometrika*, Vol. 40 (1953), pp. 87–104.

[10] Henry Scheffé, "A 'mixed-model' for the analysis of variance," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 23–36.

[11] Henry Scheffé, *The Analysis of Variance*, John Wiley and Sons, New York, 1959.

[12] M. B. Wilk and O. Kempthorne, "Fixed, mixed and random models," *J. Amer. Stat. Assn.*, Vol. 50 (1955), pp. 1144–1178.

[13] M. B. Wilk and O. Kempthorne, "Some aspects of the analysis of factorial experiments in a completely randomized design," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 950–985.

[14] S. S. Wilks, "Sample criteria for testing equality of means, equality of variances, and equality of covariances in a normal multivariate distribution," *Ann. Math. Stat.*, Vol. 17 (1946), pp. 257–281.