The Annals of Statistics
1975, Vol. 3, No. 2, 314-333

DIRECT PRODUCTS AND LINEAR MODELS FOR COMPLETE FACTORIAL TABLES[1]

BY SHELBY J. HABERMAN

University of Chicago and University of California, Berkeley

Direct products are used in some problems involving balanced factorial models for analysis of variance. Problems considered include demonstration that least squares estimators are minimum variance unbiased linear estimators when mixed models are considered, determination of the covariance structure of least squares estimators, calculation of expected values of sums of squares and mean squares, and determination of the distribution of sums of squares and least squares estimators under the normality assumption.

1. Introduction. Direct products provide a powerful tool for the examination of linear models for balanced factorial tables of the form $Y = \{Y_{i_1 \cdots i_d} : i_j \in I_j, j = 1, \cdots, d\}$. These operations have value in characterization of classes of linear models, in computation of least squares estimators, and in investigation of the behavior of models with random effects.

In Section 2 of this paper the basic properties of direct products are reviewed. Since properties of linear manifolds and of orthogonal projections have considerable importance in the analysis of variance (see Kruskal (1961, 1968)), emphasis will be placed on the properties of direct products of linear manifolds and of direct products of orthogonal projections.

In Section 3 familiar models for analysis of variance are described in terms of direct products. Balanced factorial designs are considered in which nesting may be present and in which the observation vector has a permutation-invariant covariance operator. In these models, the least squares estimator is shown to be the minimum variance unbiased linear estimator, and the distributional properties of least squares estimators and mean squares are examined. Results in this section concerning expected mean squares are related to work by Kempthorne and Wilk (1955, 1956), Cornfield and Tukey (1956), Scheffé (1956a, 1956b), and Zyskind (1962). Derivations for general balanced factorial designs are more explicit and somewhat simpler, however, than those presented in these papers.

In Section 4 results of Section 3 are applied to a generalized version of the pigeonhole model of Cornfield and Tukey (1956). An explicit proof is provided for the general results presented in that paper.

2. Properties of direct products. Direct products are closely related to tensor

Received April 1973; revised March 1974.

[1] This research was carried out in the Department of Statistics, University of Chicago, under partial support by Research Grant No. NSF GP 32037x from the Division of the Mathematical, Physical and Engineering Sciences of the National Science Foundation.

AMS 1970 *subject classifications.* Primary 62J10, 62K15; Secondary 62E15, 62F10.

Key words and phrases. Random effects, linear models, direct products, Gauss-Markov estimation, expected mean squares, analysis of variance, factorial tables.

products, Kronecker products, and outer products (see Good (1958), Halmos (1958, pages 40, 41, and 174), and Eaton (1970)). For the purpose of this paper, a direct product may be defined by considering an index set I which is a Cartesian product of sets I_j, where j is in a finite nonempty set B. Thus $I = \prod_{j \in B} I_j$. Alternatively, I may be written as $\prod \{I_j : j \in B\}$. Similar conventions will be used with other product and summation notation in this paper. Normally I_j will be the set $\bar{r}_j$ of integers i such that $1 \leqq i \leqq r_j$, where r_j is a positive integer. The set B will normally be $\bar{d}$, where d is a positive integer. If $x^{(j)} = \{x_{i_j}^{(j)}\} \in R^{I_j}$, the set of real I_j-tuples, for $j \in B$, then the direct product $x = \bigotimes \{x^{(j)} : j \in B\}$ is defined as the element of R^I such that coordinate $\{i_j : j \in B\}$ of x satisfies the equation

$$x_{\{i_j : j \in B\}} = \prod_{j \in B} x_{i_j}^{(j)} .$$

The operation $\bigotimes$ is a multilinear mapping from $\prod \{R^{I_j} : j \in B\}$ to R^I (see Loomis and Sternberg (1968, pages 306–308)); that is, if $j^* \in B$ and if $x^{(j)} \in R^j$ is fixed for $j \neq j^*$, then $\bigotimes \{x^{(j)} : j \in B\}$ is a linear function of $x^{(j^*)}$. If B_1 and B_2 are nonempty disjoint sets such that $B_1 \cup B_2 = B$, then

$$(2.1) \qquad [\bigotimes_{j \in B_1} x^{(j)}] \otimes [\bigotimes_{j \in B_2} x^{(j)}] = [\bigotimes_{j \in B_2} x^{(j)}] \otimes [\bigotimes_{j \in B_1} x^{(j)}] = \bigotimes_{j \in B} x^{(j)} .$$

If any $x^{(j)}$ is 0, then $\bigotimes \{x^{(j)} : j \in B\}$ is 0. It is sometimes convenient to extend our definition of direct products to empty sets B through the convention that if B is empty, then $\bigotimes \{x^{(j)} : j \in B\}$ is the scalar 1. It should also be noted that the convention is used that if the sets B_k, $k \in C$, are disjoint and have union B and if $i_j \in I_j$ for $j \in B_k$ and $k \in C$, then the coordinate $\{\{i_j : j \in B_k\} : k \in C\}$ is equal to $\{i_j : j \in B\}$.

If Ω_j is a linear manifold in R^{I_j} for each $j \in B$, then the direct product $\bigotimes \{\Omega_j : j \in B\}$ is the linear manifold spanned by the direct products $\bigotimes \{x^{(j)} : j \in B\}$, where $x^{(j)} \in \Omega_j$ for $j \in B$. If B is empty, $\bigotimes \{\Omega_j : j \in B\}$ is the real line R. If B_1 and B_2 are disjoint sets such that $B_1 \cup B_2 = B$, then

$$(2.2) \qquad [\bigotimes_{j \in B_1} \Omega_j] \otimes [\bigotimes_{j \in B_2} \Omega_j] = \bigotimes_{j \in B} \Omega_j .$$

If some Ω_j is the trivial linear manifold $\{0\}$, then $\bigotimes \{\Omega_j : j \in B\} = \{0\}$.

If A_j is a linear transformation from R^{I_j} to R^{I_j} for each $j \in B$, then the direct product $\bigotimes \{A_j : j \in B\}$ is the linear operator with domain and range R^I such that if $x^{(j)} \in R^{I_j}$ for $j \in B$, then

$$(2.3) \qquad [\bigotimes_{j \in B} A_j] \bigotimes_{j \in B} x^{(j)} = \bigotimes_{j \in B} A_j x^{(j)} .$$

The basic properties of direct products are summarized in the following lemmas. Most lemmas and corollaries are similar to results in Halmos (1958) or Haberman (1974), so most proofs are omitted.

LEMMA 1. *Suppose that for each* $j \in B$, $\{x^{(k_j, j)} : k_j \in \bar{p}_j\}$ *is a basis of a linear manifold* $\Omega_j \subset R^{I_j}$. *Then*

$$S = \{\bigotimes_{j \in B} x^{(k_j, j)} : k_j \in \bar{p}_j, j \in B\}$$

is a basis for $\Omega = \bigotimes \{\Omega_j : j \in B\}$, *and* Ω *has dimension* $\prod \{p_j : j \in B\}$.

REMARK 1. If Ω_j is the trivial linear manifold $\{0\}$, then $p_j = 0$, $\bar{p}_j$ is empty, and the basis is empty.

COROLLARY 1. *The space R^I is equal to $\otimes \{R^{I_j} : j \in B\}$.*

LEMMA 2. *If $(\cdot, \cdot)_J$ is the conventional inner product on R for any finite set J, then*

$$(\otimes_{j \in B} x^{(j)}, \otimes_{j \in B} y^{(j)})_I = \prod_{j \in B} (x^{(j)}, y^{(j)})_{I_j} \tag{2.4}$$

for all $x^{(j)}$ and $y^{(j)}$ in R^{I_j}, $j \in B$.

LEMMA 3. *Suppose that Ω_{j1} and Ω_{j2} are linear manifolds in R^{I_j} for $j \in B$. Then*

$$[\otimes_{j \in B} \Omega_{j1}] \cap [\otimes_{j \in B} \Omega_{j2}] = \otimes_{j \in B} (\Omega_{j1} \cap \Omega_{j2}) . \tag{2.5}$$

COROLLARY 2. *Suppose that Ω_{jk}, $k \in C$, are linear manifolds in R^{I_j} for $j \in B$. Then*

$$\bigcap_{k \in C} \otimes_{j \in B} \Omega_{jk} = \otimes_{j \in B} (\bigcap_{k \in C} \Omega_{jk}) . \tag{2.6}$$

COROLLARY 3. *If Ω_{jk}, $k \in C$, are linear manifolds in R^{I_j} for $j \in B$ and if for some $j \in B$,*

$$\bigcap_{k \in C} \Omega_{jk} = \{0\} , \tag{2.7}$$

then

$$\bigcap_{k \in C} \otimes_{j \in B} \Omega_{jk} = \{0\} . \tag{2.8}$$

COROLLARY 4. *If for $j \in B$, Ω_{j1} and Ω_{j2} are linear manifolds in R^{I_j} such that $\Omega_{j1} \subset \Omega_{j2}$, then*

$$\otimes_{j \in B} \Omega_{j1} \subset \otimes_{j \in B} \Omega_{j2} . \tag{2.9}$$

LEMMA 4. *Suppose that for $j \in B$, Ω_j is the direct sum of the pairwise disjoint linear manifolds $\Omega_{jk_j} \subset R^{I_j}$, $k_j \in \bar{s}_j$, $s_j \geqq 1$. If $\Omega = \otimes \{\Omega_j : j \in B\}$, then*

$$\Omega = \oplus \{\otimes_{j \in B} \Omega_{jk_j} : k \in \prod_{j \in B} \bar{s}_j\} . \tag{2.10}$$

REMARK. If V_j, $j \in C$, are linear manifolds in a vector space V such that $V_j \cap V_{j'} = \{0\}$ if $j \neq j'$, then the direct sum of these linear manifolds is defined as

$$\oplus \{V_j : j \in C\} = \{\sum_{j \in C} x^{(j)} : x^{(j)} \in V_j, j \in C\} .$$

The linear manifolds V_j, $j \in C$, are said to be pairwise disjoint (see Hoffman and Kunze (1961, pages 154–155)).

LEMMA 5. *Suppose that Ω_{j1} and Ω_{j2} are linear manifolds in R^{I_j} for $j \in B$, and suppose that for some $j^* \in B$, Ω_{j^*1} and Ω_{j^*2} are orthogonal. Then $\Omega_1 = \otimes \{\Omega_{j1} : j \in B\}$ and $\Omega_2 = \otimes \{\Omega_{j2} : j \in B\}$ are orthogonal.*

COROLLARY 5. *If in Lemma 4, the linear manifolds Ω_{jk_j}, $k_j \in \bar{s}_j$, are pairwise orthogonal for $j \in B$, then Ω is the direct sum of the pairwise orthogonal linear manifolds $\otimes \{\Omega_{jk_j} : j \in B\}$, $k \in \prod \{\bar{s}_j : j \in B\}$.*

LEMMA 6. *Suppose that $\Omega = \otimes \{\Omega_j : j \in B\}$, where Ω_j is a linear manifold in R^{I_j}*

for $j \in B$, and suppose that B_k, $k \in \bar{s}$, $s \geqq 2$, are nonempty pairwise disjoint sets with union B. For $k \in \bar{s}$, $x \in R^I$, and $i \in I$, let

$$T_i x = x_i .$$

Then Ω is the set of x in R^I such that if $i^* \in I(B - B_k) = \prod \{I_j : j \in B - B_k\}$, then

$$\{T_{(i^*,i')}x : i' \in I(B_k)\} \in \bigotimes_{j \in B_k} \Omega_j .$$

REMARK. This lemma is used in the next section to examine constraints in models for analysis of variance. To illustrate the meaning of the lemma, suppose that Ω_1 is the set of $x^{(1)} \in R^{I_1}$ such that

$$\sum_{i_1 \in I_1} x^{(1)}_{i_1} = 0$$

and suppose that Ω_2 is the set of $x^{(2)} \in R^{I_2}$ such that $x^{(2)}_{i_2}$ is constant for $i_2 \in I_2$. Then by letting $B_1 = \{1\}$, $B_2 = \{2\}$, and $s = 2$, one finds that $\Omega_1 \otimes \Omega_2$ consists of those x such that

$$\sum_{i_1 \in I_1} x_{i_1 i_2} = 0 , \qquad i_2 \in I_2 ,$$

and such that for a given $i_1 \in I_1$, $x_{i_1 i_2}$ is constant over $i_2 \in I_2$.

PROOF. Suppose M_k is that set of x such that

$$\{T_{(i^*,i')}x : i' \in I(B_k)\} \in \bigotimes \{\Omega_j : j \in B_k\}$$

for all $i^* \in I(B - B_k)$. Then it is necessary to show that

$$\Omega = \bigcap_{k=1}^{s} M_k .$$

By Corollary 2, it is sufficient to demonstrate that

$$M_k = [\bigotimes \{\Omega_j : j \in B_k\}] \otimes [\bigotimes \{R^{I_j} : j \in B - B_k\}] .$$

To show that M_k is included in $N_k = [\bigotimes \{\Omega_j ; j \in B_k\}] \otimes [\bigotimes \{R^{I_j} : j \in B - B_k\}]$, note that any element x of $\bigotimes \{R^{I_j} : j \in B\}$ may be written as

$$\sum \{\{T_{(i^*,i')}x : i' \in I(B_k)\} \otimes \delta(i^*) : i^* \in I(B - B_k)\} ,$$

where $\delta(i^*) \in \bigotimes \{R^{I_j} : j \in B - B_k\}$ satisfies

$$\begin{aligned} \delta_i(i^*) &= 1 \qquad \text{if} \quad i_j = i_j^* \quad \text{for} \quad j \in B - B_k , \\ &= 0 \qquad \text{otherwise.} \end{aligned}$$

If $x \in M_k$, then each summand is in N_k and therefore $x \in N_k$. If $x \in N_k$, then x can be written as

$$\sum \{a(i^*) \otimes \delta(i^*) : i^* \in I(B - B_k)\} ,$$

where for $i^* \in I(B - B_k)$, $a(i^*) \in \bigotimes \{\Omega_j : j \in B_k\}$. Since

$$\begin{aligned} T_{(i^*,i')}[a(i) \otimes \delta(i)] &= a_i(i) \qquad \text{if} \quad i = i^* , \\ &= 0 \qquad \text{if} \quad i \neq i^* , \end{aligned}$$

for all $i^* \in I(B - B_k)$, one has

$$\{T_{(i^*,i')}x : i' \in I(B_k)\} = a(i^*) \in \bigotimes \{\Omega_j : j \in B_k\} .$$

Thus $x \in M_k$. Therefore $N_k = M_k$. The lemma now follows. □

LEMMA 7. *The following relationships hold if for $j \in B$, A_j and C_j are linear transformations from R^{I_j} to R^{I_j} and $E(I_j)$ is the identity transformation on R^{I_j}:*

(a) $\bigotimes_{j \in B} A_j = 0$ *if for some* $j \in B$, $A_j = 0$.
(b) $\bigotimes_{j \in B} E(I_j) = E(I)$.
(c) $[\bigotimes_{j \in B} A_j][\bigotimes_{j \in B} C_j] = \bigotimes_{j \in B} (A_j C_j)$
(d) $[\bigotimes_{j \in B} A_j]^* = \bigotimes_{j \in B} A_j^*$, *where* A_j^* *is the adjoint of* A_j.
(e) $\mathscr{R}(\bigotimes_{j \in B} A_j) = \bigotimes_{j \in B} \mathscr{R}(A_j)$, *where* $\mathscr{R}(A_j)$ *is the range of* A_j.
(f) $[\bigotimes_{j \in B} A_j]x = \sum_{i \in I} x_i \bigotimes_{j \in B} A_j \delta^{(j,i_j)}$ *for all* $x \in R^I$, *where for* $i_j \in I_j$, $i_j' \in I_j$, *and* $j \in B$, $\delta_{i_j'}^{(j,i_j)}$ *is* 0 *if* $i_j' \neq i_j$ *and* 1 *if* $i_j' = i_j$.

PROOF. Results (a), (b), and (c) are straightforward generalizations of results in Halmos (1958, page 96). To prove (d), note that for x_j and z_j in R^{I_j}, $j \in B$, Lemma 2 implies that

$$\begin{aligned} ([\textstyle\bigotimes_{j \in B} A_j] \bigotimes_{j \in B} x_j, \bigotimes_{j \in B} z_j)_I &= (\textstyle\bigotimes_{j \in B} A_j x_j, \bigotimes_{j \in B} z_j)_I \\ &= \textstyle\prod_{j \in B} (A_j x_j, z_j)_{I_j} \\ &= \textstyle\prod_{j \in B} (x_j, A_j^* z_j)_{I_j} \\ &= (\textstyle\bigotimes_{j \in B} x_j, \bigotimes_{j \in B} A_j^* z_j)_I \\ &= (\textstyle\bigotimes_{j \in B} x_j, [\bigotimes_{j \in B} A_j^*] \bigotimes_{j \in B} z_j)_I . \end{aligned} \tag{2.11}$$

To prove (e), note that if $x_j \in \mathscr{R}(A_j)$ for $j \in B$, then for each $j \in B$, there exists $z_j \in R^{I_j}$ such that $x_j = A_j z_j$. Thus

$$\textstyle\bigotimes_{j \in B} x_j = [\bigotimes_{j \in B} A_j] \bigotimes_{j \in B} z_j \in \mathscr{R}(\bigotimes_{j \in B} A_j) . \tag{2.12}$$

Conversely, if $x \in \mathscr{R}(\bigotimes \{A_j : j \in B\})$, then for some k and $z_j^{(l)}$, $j \in B$, $l \in \bar{k}$,

$$x = [\textstyle\bigotimes_{j \in B} A_j] \sum_{l=1}^{k} \bigotimes_{j \in B} z_j^{(l)} = \sum_{l=1}^{k} \bigotimes_{j \in B} A_j z_j^{(l)} \in \bigotimes_{j \in B} \mathscr{R}(A_j) . \tag{2.13}$$

Thus (e) must hold.

Result (f) follows from the definition of $\bigotimes \{A_j : j \in B\}$ since

$$x = \textstyle\sum_{i \in I} x_i \bigotimes_{j \in B} \delta^{(j,i_j)} . \qquad \square$$

LEMMA 8. *If the orthogonal projection on* Ω_j *is* P_{Ω_j} *for* $j \in B$, *then the orthogonal projection* P_Ω *on* $\Omega = \bigotimes \{\Omega_j : j \in B\}$ *is* $\bigotimes \{P_{\Omega_j} : j \in B\}$.

REMARK. In the case $B = \{1, 2\}$, this lemma is given by Rao and Mitra (1971. page 119).

PROOF. Note that since

$$\begin{aligned} [\textstyle\bigotimes_{j \in B} P_{\Omega_j}]^2 &= \textstyle\bigotimes_{j \in B} P_{\Omega_j} P_{\Omega j} \\ &= \textstyle\bigotimes_{j \in B} P_{\Omega_j} , \\ [\textstyle\bigotimes_{j \in B} P_{\Omega_j}]^* &= \textstyle\bigotimes_{j \in B} P_{\Omega_j}^* \\ &= \textstyle\bigotimes_{j \in B} P_{\Omega_j} , \end{aligned}$$

and

$$\mathscr{R}(\bigotimes_{j \in B} P_{\Omega_j}) = \bigotimes_{j \in B} \mathscr{R}(P_{\Omega_j}) = \bigotimes_{j \in B} \Omega_j ,$$

$\bigotimes \{P_{\Omega_j} : j \in B\}$ is the orthogonal projection on Ω (see Halmos (1958, page 146)). □

LEMMA 9. *If for some subset C of B, X is a random variable on $R^{I(C)}$ with mean μ and covariance operator $\sum$, then for any constant $z \in R^{I(B-C)}$, $X \otimes z$ has mean $\mu \otimes z$ and covariance operator $\|z\|^2_{I(B-C)} \sum \otimes P_{\mathrm{span}\{z\}}$.*

REMARK. As in Kruskal (1961), the covariance operator $\sum$ is the linear transformation such that for u and v in $R^{I(C)}$,

$$\mathrm{Cov}\,[(u, X), (v, X)] = (u, \textstyle\sum v)_{I(C)} .$$

PROOF. If $v \in R^{I(C)}$ and $w \in R^{I(B-C)}$, then $(v \otimes w, X \otimes z)_I$ has expected value

$$\begin{aligned} E(v \otimes w, X \otimes z)_I &= E[(v, X)_{I(C)}(w, z)_{I(B-C)}] \\ &= (v, \mu)_{I(C)}(w, z)_{I(B-C)} \\ &= (v \otimes w, \mu \otimes z)_I \end{aligned} \tag{2.14}$$

and variance

$$\begin{aligned} \mathrm{Var}\,(v \otimes w, X \otimes z)_I &= [(w, z)_{I(B-C)}]^2 \,\mathrm{Var}\,(v, X)_{I(C)} \\ &= [(w, z)_{I(B-C)}]^2 (v, \textstyle\sum v)_{I(C)} . \end{aligned} \tag{2.15}$$

Since

$$P_{\mathrm{span}\{z\}} w = \frac{(w, z)_{I(C)}}{(z, z)_{I(C)}} z , \tag{2.16}$$

it follows that

$$\begin{aligned} \mathrm{Var}\,(v \otimes w, X \otimes z)_I &= \|z\|^2_{I(C)} (w, P_{\mathrm{span}\{z\}} w)_{I(B-C)} (v, \textstyle\sum v)_{I(C)} \\ &= \|z\|^2_{I(C)} (v \otimes w, [\textstyle\sum \otimes P_{\mathrm{span}\{z\}}] v \otimes w)_I . \end{aligned} \tag{2.17}$$

These results imply that $X \otimes z$ has expected value $\mu \otimes z$ and covariance operator

$$\|z\|^2_{I(C)} \textstyle\sum \otimes P_{\mathrm{span}\{z\}} . \qquad \square$$

3. Direct products and models for analysis of variance. In the models considered in this paper, factors $j \in B$ are present, where B is a nonempty set of d elements. Factors are divided into sets D_k, $k \in \bar{s}$, $s \geqq 1$, such that for $k \in \bar{s}$, all factors $j \in D_k$ are completely crossed. For any $k > 1$ such that $k \leqq s$, a set $S_k \subset \overline{k-1}$ exists such that the factors in D_k are nested within all factors in $N_k = \bigcup \{D_l : l \in S_k\}$ and completely crossed with the factors $C_k = \{D_l : l \in \overline{k-1} - S_k\}$. If $k = 1$, one may define C_k and N_k as the empty set $\varnothing$. It should be noted that if $D_k \subset N_{k'}$ and $D_{k'} \subset N_{k''}$, then $D_k \subset N_{k''}$. The possible factor levels are described by finite index sets I_j, $j \in B$, with $r_j \geqq 1$ elements, where each factor $j \in D_k$, $k \in \bar{s}$, has levels $\{i_{j'} : j' \in N_k \cup \{j\}\} \in I(N_k \cup \{j\})$.

For each $i \in I$, an observation Y_i is taken for which factors $j \in D_k$, $k \in \bar{s}$, are at levels $\{i_{j'} : j' \in N_k \cup \{j\}\}$. This observation is assumed to satisfy the equation

$$Y_i = \textstyle\sum_{A \in \mathscr{A}} \alpha^A_{\{i_j : j \in N(A)\}} + \sum_{G \in \mathscr{G}} e^G_{\{i_j : j \in N(G)\}} , \tag{3.1}$$

where $\mathscr{A}$ and $\mathscr{G}$ are nonempty classes of subsets of B such that if $A \in \mathscr{A} \cup \mathscr{G}$, then for $1 < k \leqq s$, $A \cap D_k \neq \emptyset$ implies that $A \cap N_k = \emptyset$. The set $N(A)$ is defined as the union of A and $\bigcup \{N_k : A \cap D_k \neq \emptyset, 1 < k \leqq s\}$.

The vectors $\alpha^A = \{\alpha_i^A : i \in I(N(A))\}$, $A \in \mathscr{A}$, are fixed effects such that $\alpha^A \in \omega(A)$, where

$$\omega(A) = [\bigotimes_{j \in A} \{h^{(j)}\}^\perp] \otimes [\bigotimes \{R^{I_j} : j \in N(A) - A\}] \tag{3.2}$$

and $h^{(j)} = \{1 : i_j \in I_j\}$ is the unit vector of R^{I_j}. By Lemma 6, $\alpha^A \in \omega(A)$ if and only if α^A satisfies the constraints

$$\sum \{\alpha_i^A : i_{j^*} \in I_{j^*}, i_{j'} = i^*_{j'} \;\; \forall j' \in N(A) - \{j^*\}\} = 0 \tag{3.3}$$

for all $j^* \in A$ and $i^* \in I(N(A) - \{j^*\})$.

If $G \in \mathscr{G}$, then $e^G = \{e_i^G : i \in I(N(G))\}$ is a vector of random effects with mean 0 and covariance operator

$$\begin{aligned} \textstyle\sum_G &= \sum \{b(G, G')P_{\Lambda(G,G')} : G' \subset N(G)\} \\ &= \sum \{b(G, G')[\otimes \{Q_j : j \in G'\}] \\ &\qquad \otimes [\otimes \{P_j : j \in N(G) - G'\}] : G' \subset N(G)\}, \end{aligned} \tag{3.4}$$

where $\{b(G, G') : G' \subset N(G), G \in \mathscr{G}\}$ is included in some set Θ,

$$\Lambda(G, G') = [\otimes \{\{h^{(j)}\}^\perp : j \in G'\}] \otimes [\otimes \{\text{span}\,\{h^{(j)}\} : j \in N(G) - G'\}], \tag{3.5}$$

P_j is the orthogonal projection on span $\{h^{(j)}\}$, and $Q_j = E(I_j) - P_j$ is the orthogonal projection on $\{h^{(j)}\}^\perp$. It should be noted that Corollary 5 implies that the linear manifolds $\Lambda(G, G')$, $G' \subset N(G)$, are mutually orthogonal. By Lemma 8,

$$R^{I(N(G))} = \bigoplus \{\Lambda(G, G') : G' \subset N(G)\}. \tag{3.6}$$

Since $\sum_G$ must be nonnegative definite, $b(G, G') \geqq 0$ for $G' \subset N(G)$, and $b(G, G') > 0$ for $G' \subset N(G)$ if and only if $\sum_G$ is positive definite (see Halmos (1958, pages 153 and 156)). In the case where $b(G, G') = \sigma_G^2$ for $G' \subset N(G)$, it follows from Halmos (1958, pages 147–148) that

$$\textstyle\sum_G = \sigma_G^2 E(N(G)), \tag{3.7}$$

the covariance operator when $\text{Var}(e_i^G) = \sigma_G^2$ for $i \in I(N(G))$ and $\text{Cov}(e_i^G, e_{i'}^G) = 0$ for $i \in I(N(G))$, $i' \in I(N(G))$, $i \neq i'$.

In general, (3.4) is equivalent to the assumption that the covariance structure of e^G is permutation invariant in the sense that for each $G' \subset N(G)$ there exists a $c(G, G')$ such that if $i \in I(N(G))$, $i' \in I(N(G))$, $i_j = i_j'$ for $j \in G'$, and $i_j \neq i_j'$ for $j \in N(G) - G'$, then $\text{Cov}(e_i^G, e_{i'}^G) = c(G, G')$. To prove this assertion, first suppose that (3.4) holds. Since

$$e_i^G = (e^G, \bigotimes_{j \in N(G)} \delta^{(j, i_j)})_{I(N(G))}, \tag{3.8}$$

where $\delta^{(j,i_j)}$ is defined as in Lemma 7 for $j \in B$ and $i_j \in I_j$, Lemma 2 then implies that

$$\begin{aligned}(3.9)\qquad &\operatorname{Cov}(e_i^G, e_{i'}^G) \\ &= (\bigotimes_{j\in N(G)} \delta^{(j,i_j)}, \textstyle\sum_G \bigotimes_{j\in N(G)} \delta^{(j,i_j')})_{I(N(G))} \\ &= \textstyle\sum \{b(G, G'')[\prod \{(\delta^{(j,i_j)}, Q_j\delta^{(j,i_j')})_{I_j} : j \in G''\}] \\ &\qquad \times [\prod \{(\delta^{(j,i_j)}, P_j\delta^{(j,i_j')})_{I_j} : j \in N(G) - G''\}] : G'' \subset N(G)\} .\end{aligned}$$

Since

$$P_j\delta^{(j,i_j)} = \frac{1}{r_j} h^{(j)} ,$$

it follows that if $i_j = i_j'$ for $j \in G'$ and $i_j \neq i_j'$ for $j \in N(G) - G'$, then

$$\begin{aligned}(3.10)\qquad c(G, G') &= \operatorname{Cov}(e_i^G, e_{i'}^G) \\ &= \textstyle\sum \{s(G', G'')b(G, G'')[\prod \{(1 - 1/r_j) : j \in G' \cap G''\}] \\ &\qquad \div [\prod \{r_j : j \in N(G) - G' \cap G''\}] : G'' \subset N(G)\} ,\end{aligned}$$

where $s(G', G'')$ is 1 if $G'' - G'' \cap G'$ has an even number of elements and $s(G', G'')$ is -1 otherwise. Similar calculations may be used to show that whenever $\operatorname{Cov}(e_i^G, e_{i'}^G) = c(G, G')$ for all $i \in I(N(G))$, $i' \in I(N(G))$, $i_j = i_j'$ for $j \in G$ and $i_j \neq i_j'$ for $j \in N(G) - G'$, then (3.4) holds with

$$\begin{aligned}(3.11)\qquad b(G, G') = \textstyle\sum \{&c(G, G'')s(G'', G') \prod \{r_j - 1 : j \in [N(G) - G'] \\ &\cap [N(G) - G'']\} : G'' \subset N(G)\} .\end{aligned}$$

To place the discussion in the remainder of the section on a more concrete level, it is useful to consider the following example.

EXAMPLE 1. Scheffé (1959, pages 276–278) considers an experiment in which factors 1 and 4 are completely crossed, factor 2 is nested within factors 1 and 4, factor 3 is completely crossed with factors 1, 2, and 4, and factor 5, which represents random error, is nested within the remaining four factors. In the model considered, $D_1 = \{1, 4\}$, $D_2 = \{2\}$, $D_3 = \{3\}$, $D_4 = \{5\}$, $S_2 = \{1\}$, $S_3 = \varnothing$, $S_4 = \{1, 2, 3\}$, and Y satisfies

$$\begin{aligned}Y_{i_1i_2i_3i_4} = \alpha^{\varnothing} + \alpha_{i_1}^{\{1\}} + \alpha_{i_3}^{\{3\}} + \alpha_{i_1i_3}^{\{1,3\}} + e_{i_4}^{\{4\}} + e_{i_1i_2i_4}^{\{2\}} + e_{i_1i_4}^{\{1,4\}} + e_{i_3i_4}^{\{3,4\}} \\ + e_{i_1i_2i_3i_4}^{\{2,3\}} + e_{i_1i_3i_4}^{\{1,3,4\}} + e_{i_1i_2i_3i_4i_5}^{\{5\}} ,\end{aligned}$$

where

$$\textstyle\sum_{i_1=1}^{r_1} \alpha_{i_1}^{\{1\}} = \sum_{i_3=1}^{r_3} \alpha_{i_3}^{\{3\}} = \sum_{i_1=1}^{r_1} \alpha_{i_1i_2}^{\{1,3\}} = \sum_{i_3=1}^{r_3} \alpha_{i_1i_3}^{\{1,3\}} = 0 .$$

Thus $\mathscr{A} = \{\varnothing, \{1\}, \{3\}, \{1, 3\}\}$ and $\mathscr{G} = \{\{4\}, \{2\}, \{1, 4\}, \{3, 4\}, \{2, 3\}, \{1, 3, 4\}, \{5\}\}$. The vectors $e^{\{4\}}$, $e^{\{2\}}$, $e^{\{1,4\}}$, $e^{\{3,4\}}$, $e^{\{2,3\}}$, $e^{\{1,3,4\}}$, and $e^{\{5\}}$ all have mean 0, and it is assumed that

$$\begin{aligned}\textstyle\sum_{i_1=1}^{r_1} e_{i_1i_4}^{\{1,4\}} = \sum_{i_3=1}^{r_3} e_{i_3i_4}^{\{3,4\}} = \sum_{i_3=1}^{r_3} e_{i_1i_2i_3i_4}^{\{2,3\}} = \sum_{i_3=1}^{r_3} e_{i_1i_3i_4}^{\{1,3,4\}} = \sum_{i_1=1}^{r_1} e_{i_1i_3i_4}^{\{1,3,4\}} \\ = 0 .\end{aligned}$$

To obtain these constraints, one may assume that $b(G, G') = \sigma_G^2$ if $G \subset \mathscr{G}$ and $\{1, 3\} \cap G \subset G'$ and $b(G, G') = 0$ otherwise. The model in this example belongs to the general class of models for which B is divided into sets B_f and B_r, where the factors $j \in B_f$ are fixed in advance and the factors $j \in B_r$ are regarded as sampled from an infinite population. In such models, any $A \in \mathscr{A}$ satisfies the condition that $N(A) \subset B_f$ and for any $G \in \mathscr{G}$, $N(G) \cap B_r \neq \emptyset$, $b(G, G') = \sigma_G^2$ if $G' \in N(G)$ and $B_f \cap G \subset G'$, and $b(G, G') = 0$ if $G' \subset N(G)$ and $B_f \cap G \not\subset G'$.

3.1. *The mean and covariance operator of Y.* The vector Y may be written as

$$(3.12) \qquad Y = \textstyle\sum_{A \in \mathscr{A}} \alpha^A \otimes [\bigotimes \{h^{(j)} : j \in B - N(A)\}] + \sum_{G \in \mathscr{G}} e^G \otimes [\bigotimes \{h^{(j)} : j \in B - N(G)\}] .$$

By Lemmas 8 and 9, Y has expected value

$$(3.13) \qquad \mu = \textstyle\sum_{A \in \mathscr{A}} \alpha^A \otimes [\bigotimes \{h^{(j)} : j \in B - N(A)\}]$$

and covariance operator

$$(3.14) \qquad \begin{aligned} \Sigma &= \textstyle\sum_{G \in \mathscr{G}} [\prod \{r_j : j \in B - N(G)\}] \Sigma_G \otimes [\bigotimes \{P_j : j \in B - N(G)\}] \\ &= \textstyle\sum_{G \in \mathscr{G}} \sum_{G' \subset N(G)} b(G, G')[\prod \{r_j : j \in B - N(G)\}][\bigotimes \{Q_j : j \in G'\}] \\ &\qquad \otimes [\bigotimes \{P_j : j \in B - G'\}] \\ &= \textstyle\sum_{G' \subset B} d(G') P_{\Delta(G')} , \end{aligned}$$

where

$$(3.15) \qquad d(G') = \textstyle\sum \{b(G, G') \prod \{r_j : j \in B - N(G)\} : G' \subset N(G), G \in \mathscr{G}\}$$

and

$$(3.16) \qquad \Delta(G') = [\textstyle\bigotimes_{j \in G'} \{h^{(j)}\}^\perp] \otimes [\bigotimes \{\text{span}\, \{h^{(j)}\} : j \in B - G'\}] .$$

To describe the vector μ more thoroughly, suppose that if $A \in \mathscr{A} \cup \mathscr{G}$, then

$$(3.17) \qquad \begin{aligned} \Omega(A) &= \omega(A) \otimes [\textstyle\bigotimes \{\text{span}\, \{h^{(j)} : j \in B - N(A)\}] \\ &= \bigoplus \{\Delta(A') : A \subset A' \subset N(A)\} , \end{aligned}$$

where the linear manifolds $\Delta(A')$, $A \subset A' \subset N(A)$, are orthogonal (see (3.2), (3.16), and Corollary 5). Consider the following theorem:

THEOREM 1. *The linear manifolds $\Omega(A)$, $A \in \mathscr{A}$, are mutually orthogonal.*

PROOF. Since the linear manifolds $\Delta(A)$, $A \subset B$, are mutually orthogonal, (3.17) implies that it is only necessary to show that if $A_1 \in \mathscr{A}$, $A_2 \in \mathscr{A}$, and $A_1 \neq A_2$, then there exists no $A' \subset B$ for which $A_i \subset A' \subset N(A_i)$, $i = 1$ or 2. To do so, let k be the largest integer, $1 \leqq k \leqq s$, such that $A_1 \cap D_k \neq A_2 \cap D_k$ and suppose that $A_i \subset A' \subset N(A_i)$, $i = 1$ or 2. Either one of the two sets $A_1 \cap D_k$ and $A_2 \cap D_k$ is empty or neither of the sets is empty. In the first case, suppose $A_2 \cap D_k = \emptyset$. Then in order for A' to contain A_1 and be contained in $N(A_2)$, it must be the case that for some $k' > k$, $D_k \subset N_{k'}$ and $A_2 \cap D_{k'} \neq \emptyset$. However, for such a k', $A_1 \cap D_{k'} = A_2 \cap D_{k'}$ but $A_1 \cap N_{k'} \supset A_1 \cap D_k \neq \emptyset$, a contradiction. If neither $A_1 \cap D_k$ nor $A_2 \cap D_k$ is empty, then $N(A_1) \cap D_k = A_1 \cap D_k$ and $N(A_2) \cap D_k = A_2 \cap D_k$. Thus $A_1 \cap D_k \subset A' \cap D_k \subset A_2 \cap D_k$ and $A_2 \cap D_k \subset A' \cap D_k \subset A_1 \cap D_k$, which implies that $A_1 \cap D_k = A_2 \cap D_k$, a contradiction. □

Given this theorem, it follows that $\mu \in \Omega$, where

$$(3.18) \qquad \begin{aligned} \Omega &= \bigoplus_{A \in \mathscr{A}} \Omega(A) \\ &= \bigoplus \{\Delta(A') : A \subset A' \subset N(A), A \in \mathscr{A}\} . \end{aligned}$$

Conversely, if $\mu \in \Omega$, then μ satisfies (3.13) for some unique $\alpha^A \in \omega(A)$, $A \in \mathscr{A}$.

The covariance operator $\sum$ is always permutation-invariant in the sense that $\operatorname{Cov}(Y_i, Y_{i'}) = c'(G')$ if $G' \subset B$, $i_j = i_j'$ for $j \in G'$, and $i_j \neq i_j'$ for $j \in B - G'$. Whether $\sum$ determines $\{b(G, G')\}$ depends on the size of the set Θ. If $\sum$ is permutation-invariant and $\mu \in \Omega$, then Y satisfies (3.1) for some choice of $\mathscr{G}$ and Θ, for one may let $\mathscr{G} = \{B\}$ and $\sum_B = \sum$.

3.2. *Least squares and Gauss–Markov estimation.* If $\sum$ is assumed to be a multiple of the identity operator $E(I)$, then the least squares estimator $\hat{\mu} = P_\Omega Y$ of μ is well known to be the minimum variance linear unbiased (Gauss–Markov) estimator of μ (see Kruskal (1961)). An important feature of the models examined in this section is that $\hat{\mu}$ has the property that if AY is a linear estimator of μ such that $Ax = x$ for all $x \in \Omega$, and if $\mu \in \Omega$ and $\sum$ satisfies (3.14) for some $\{b(G, G') : G' \subset N(G), G \in \mathscr{G}\} \in \Theta$, then for all $x \in R^I$,

$$(3.19) \qquad \operatorname{Var}(x, \hat{\mu}) \leqq \operatorname{Var}(x, AY) .$$

Following Eaton (1970), $\hat{\mu}$ may be called a Gauss–Markov estimator of μ.

To verify that $\hat{\mu}$ has the desired property, it suffices to show that $\sum \Omega$, the image of Ω under $\sum$, is included in Ω whenever $\sum$ satisfies (3.14) for some $\{b(G, G')\} \in \Theta$. This claim follows since if $x \in \Delta(A')$, $A \subset A' \subset N(A)$, then (3.14) implies that

$$(3.20) \qquad \textstyle\sum x = d(A')x .$$

This observation has been made previously by Kruskal (1968) and Eaton (1970) for the case in which B has a single element 1 and

$$Y_i = \alpha^{\varnothing} + e_i^{\{1\}} ,$$

where $E(e^{\{1\}}) = 0$ and $\sum = \sum_{\{1\}}$ has the permutation-invariance property that $\operatorname{Cov}(Y_i, Y_{i'}) = \rho\sigma^2$ for $i \neq i'$ and $\operatorname{Var}(Y_i) = \sigma^2$ for $i \in I = I_1$.

3.3. *Sampling properties of least squares estimators.* The mean and covariance operator of $\hat{\mu}$ are easily found by reference to Rao (1965, page 438), (3.14), and (3.18). Since

$$(3.21) \qquad P_\Omega = \sum \{P_{\Delta(A')} : A \subset A' \subset N(A), A \in \mathscr{A}\} ,$$

the mean of $\hat{\mu} = P_\Omega Y$ is μ and the covariance operator $P_\Omega \sum P_\Omega$ satisfies

$$(3.22) \qquad P_\Omega \textstyle\sum P_\Omega = \sum \{d(A')P_{\Delta(A')} : A \subset A' \subset N(A), A \in \mathscr{A}\} .$$

To examine least squares estimators of the fixed effects α^A, $A \in \mathscr{A}$, the well-known observation is used that the least squares estimate of a linear functional $(x, \mu)_I$ of μ is $(x, \hat{\mu})_I$, which is equal to $(x, Y)_I$ if $x \in \Omega$ (see Kruskal (1961)).

Since whenever $z \in R^{I(N(A))}$ for some $A \in \mathscr{A}$, (3.13) and (3.16) imply that

$$\begin{aligned}(3.23)\qquad (z, \alpha^A)_{I(N(A))} &= \frac{(z \otimes [\otimes \{h^{(j)} : j \in B - N(A)\}], \alpha^A \otimes [\otimes \{h^{(j)} : j \in B - N(A)\}])_I}{\prod \{r_j : j \in B - N(A)\}} \\ &= (z \otimes [\otimes \{h^{(j)} : j \in B - N(A)\}], P_{\Omega(A)}\mu)_I / \prod \{r_j : j \in B - N(A)\} \\ &= (P_{\omega(A)} z \otimes [\otimes \{h^{(j)} : j \in B - N(A)\}], \mu)_I / \prod \{r_j : j \in B - N(A)\} ,\end{aligned}$$

it follows that the least squares estimator of $(z, \alpha^A)_{I(N(A))}$ is

$$\begin{aligned}(3.24)\qquad (P_{\omega(A)} z \otimes [\otimes \{h^{(j)} : j \in B - N(A)\}], Y)_I / \prod \{r_j : j \in B - N(A)\} &= (P_{\omega(A)} z, \bar{Y}^{N(A)})_{I(N(A))} \\ &= (z, P_{\omega(A)} \bar{Y}^{N(A)})_{I(N(A))} ,\end{aligned}$$

where $\bar{Y}^{N(A)} = \{\bar{Y}_{i'}^{N(A)} : i' \in I(N(A))\}$ is the vector of averages of Y_i such that $i_j = i_j'$ for $j \in N(A)$. In general, if $i' \in I(A)$, $A \subset B$,

$$(3.25)\qquad \bar{Y}_{i'}^A = \sum \{Y_i : i \in \{i'\} \times I(B - A)\} / \prod \{r_j : j \in B - A\} .$$

Given Lemma 7, Lemma 8, (3.2), and (3.24), the least squares estimator $\hat{\alpha}^A$ of α^A may be written

$$\begin{aligned}(3.26)\qquad \hat{\alpha}^A &= P_{\omega(A)} \bar{Y}^{N(A)} \\ &= [\otimes_{j \in A} (E(I_j) - P_j)] \otimes [\otimes \{E(I_j) : j \in N(A) - A\}] \bar{Y}^{N(A)} \\ &= \{\textstyle\sum \{\prod \{(-1) : j \in N(A) - A'\} \bar{Y}_{\{i_j : j \in A'\}}^{A'} : A \subset A' \subset N(A)\} : \\ &\qquad i \in I(N(A))\} .\end{aligned}$$

Thus the least squares estimator of $(z, \alpha^A)_{I(N(A))}$ is $(z, \hat{\alpha}^A)_{I(N(A))}$.

By (3.24) and Lemmas 2 and 8, $(z, \hat{\alpha}^A)_{I(N(A))}$ has expected value $(z, \alpha^A)_{I(N(A))}$ and variance

$$\begin{aligned}(3.27)\qquad &\frac{(P_{\omega(A)} z \otimes [\otimes \{h^{(j)} : j \in B - N(A)\}], \Sigma [P_{\omega(A)} z \otimes [\otimes \{h^{(j)} : j \in B - N(A)\}])_I}{[\prod \{r_j : j \in B - N(A)\}]^2} \\ &= \frac{\sum \{d(A') \| P_{\Lambda(A,A')} z \|_{I(N(A))}^2 : A \subset A' \subset N(A)\}}{\prod \{r_j : j \in B - N(A)\}} .\end{aligned}$$

Since the covariance operator of $\hat{\alpha}^A$ is determined by $\operatorname{Var}(z, \hat{\alpha}^A)$ for $z \in R^{I(N(A))}$, it follows that $\hat{\alpha}^A$ has expected value α^A and covariance operator

$$(3.28)\qquad \operatorname{Cov}(\hat{\alpha}^A) = \frac{\sum \{d(A') P_{\Lambda(A,A')} : A \subset A' \subset N(A)\}}{\prod \{r_j : j \in B - N(A)\}} .$$

In this formula, it should be noted that

$$\begin{aligned}(3.29)\qquad & d(A') / \prod \{r_j : j \in B - N(A)\} \\ &= \sum \{b(G, A') / \prod \{r_j : j \in N(G) - N(A)\} : G \in \mathscr{G}, A' \subset G\} .\end{aligned}$$

If $d(A') = d(A)$ for $A \subset A' \subset N(A)$, then $\operatorname{Cov}(\hat{\alpha}^A)$ is $[d(A) / \prod \{r_j : j \in B - N(A)\}] P_{\omega(A)}$.

To complete the description of the covariance structure of the estimators $\hat{\alpha}^A$,

$A \in \mathscr{A}$, it should be noted that these estimators are uncorrelated. To prove this assertion, note that if $z_1 \in R^{I(N(A_1))}$ and $z_2 \in R^{I(N(A_2))}$, where A_1 and A_2 are in $\mathscr{A}$, then the covariance of $(z_1, \hat{\alpha}^{A_1})$ and $(z_2, \hat{\alpha}^{A_2})$ is

$$\begin{aligned}(P_{\omega(A_1)}z_1 \otimes [\otimes \{h^{(j)} : j \in B - N(A_1)\}],\\ \textstyle\sum (P_{\omega(A_2)}z_2 \otimes [\otimes \{h^{(j)} : j \in B - N(A_2)\}]))_I\\ \div [\prod \{r_j : j \in B - N(A_1)\}][\prod \{r_j : j \in B - N(A_2)\}]\\ = (z_1 \otimes [\otimes \{h^{(j)} : j \in B - N(A_1)\}],\\ P_{\Omega(A_1)} \textstyle\sum P_{\Omega(A_2)}\{z_2 \otimes [\otimes \{h^{(j)} : j \in B - N(A_2)\}]\})_I\\ \div [\prod \{r_j : j \in B - N(A_1)\}][\prod \{r_j : j \in B - N(A_2)\}],\end{aligned} \tag{3.30}$$

which is 0 since Theorem 1, (3.14), and (3.17) imply that $P_{\Omega(A_1)} \sum P_{\dot{\Omega}(A_2)} = 0$. If Y has a multivariate normal distribution, then the estimators $\hat{\alpha}^A$, $A \in \mathscr{A}$, have multivariate normal distributions and are independent.

EXAMPLE 2. To illustrate the sampling results in this section, consider the experiment described in Example 1. Since $N(\varnothing) = \varnothing$, $N(\{1\}) = \{1\}$, $N(\{3\}) = \{3\}$, and $N(\{1, 3\}) = \{1, 3\}$, it is only necessary to compute $d(\varnothing)$, $d(\{1\})$, $d(\{3\})$, and $d(\{1, 3\})$ to describe the covariance structure of the least squares estimators. Since $N(\{2\}) = \{1, 2, 4\}$, $N(\{4\}) = \{4\}$, $N(\{1, 4\}) = \{1, 4\}$, $N(\{3, 4\}) = \{3, 4\}$, $N(\{2, 3\}) = \{1, 2, 3, 4\}$, $N(\{1, 3, 4\}) = \{1, 3, 4\}$, and $N(\{5\}) = \{1, 2, 3, 4, 5\}$, (3.15) implies that

$$d(\varnothing) = r_3 r_5 \sigma^2_{\{2\}} + r_1 r_2 r_3 r_5 \sigma^2_{\{4\}} + \sigma^2_{\{5\}}, \tag{3.31}$$

$$d(\{1\}) = r_3 r_5 \sigma^2_{\{2\}} + r_2 r_3 r_5 \sigma^2_{\{1,4\}} + \sigma^2_{\{5\}}, \tag{3.32}$$

$$d(\{3\}) = r_1 r_2 r_5 \sigma^2_{\{3,4\}} + r_5 \sigma^2_{\{2,3\}} + \sigma^2_{\{5\}}, \tag{3.33}$$

and

$$d(\{1, 3\}) = r_5 \sigma^2_{\{2,3\}} + \sigma^2_{\{5\}}. \tag{3.34}$$

By (3.22), the covariance of $\hat{\mu}$ is

$$d(\varnothing)P_{\Delta(\varnothing)} + d(\{1\})P_{\Delta(\{1\})} + d(\{3\})P_{\Delta(\{3\})} + d(\{1, 3\})P_{\Delta(\{1,3\})},$$

while (3.28) and the fact that $\Lambda(A, A) = \omega(A)$ for $A \in \mathscr{A}$ imply that

$$\operatorname{Var}(\hat{\alpha}^{\varnothing}) = \frac{\sigma^2_{\{2\}}}{r_1 r_2 r_4} + \frac{\sigma^2_{\{4\}}}{r_4} + \frac{\sigma^2_{\{5\}}}{r_1 r_2 r_3 r_4 r_5}, \tag{3.35}$$

$$\operatorname{Cov}(\hat{\alpha}^{\{1\}}) = \left\{\frac{\sigma^2_{\{2\}}}{r_2 r_4} + \frac{\sigma^2_{\{1,4\}}}{r_4} + \frac{\sigma^2_{\{5\}}}{r_2 r_3 r_4 r_5}\right\} P_{\omega(\{1\})}, \tag{3.36}$$

$$\operatorname{Cov}(\hat{\alpha}^{\{2\}}) = \left\{\frac{\sigma^2_{\{3,4\}}}{r_4} + \frac{\sigma^2_{\{2,3\}}}{r_1 r_2 r_4} + \frac{\sigma^2_{\{5\}}}{r_1 r_2 r_4 r_5}\right\} P_{\omega(\{2\})}, \tag{3.37}$$

and

$$\operatorname{Cov}(\hat{\alpha}^{\{1,3\}}) = \left\{\frac{\sigma^2_{\{2,3\}}}{r_2 r_4} + \frac{\sigma^2_{\{1,3,4\}}}{r_4} + \frac{\sigma^2_{\{5\}}}{r_2 r_4 r_5}\right\} P_{\omega(\{1,3\})}. \tag{3.38}$$

Within the general class of models described in Example 1, one has

$$d(A') = \sum \{\sigma_G^2 \prod \{r_j : j \in B - N(G)\} : B_f \cap G \subset A' \subset N(G), G \in \mathscr{G}\} \tag{3.39}$$

for $A \subset A' \subset N(A)$, $A \in \mathscr{A} \cup \mathscr{G}$. Given $A \in \mathscr{A}$, $d(A')$ is constant for all A' such that $A \subset A' \subset N(A)$. To prove this assertion, it is sufficient to show that for $G \in \mathscr{G}$, $B_f \cap G \subset A \subset N(G)$ if and only if $B_f \cap G \subset N(A) \subset N(G)$.

If $B_f \cap G \subset A \subset N(G)$, then the relationship $A \subset N(A)$ implies that $B_f \cap G \subset N(A)$. For each D_k, $1 \leqq k \leqq s$, $A \cap D_k \neq \emptyset$ implies that $G \cap D_k \supset A \cap D_k$ or $G \cap D_k = \emptyset$ and $D_k \subset N_{k'}$ for some k' such that $G \cap D_{k'} \neq \emptyset$. In either case $N(A) \cap D_k = A \cap D_k \subset N(G) \cap D_k$. If $A \cap D_k = \emptyset$ but $D_k \subset N_{k'}$ for some k' such that $A \cap D_{k'} \neq \emptyset$, then $G \cap D_{k'} \neq \emptyset$ or for some k'', $D_{k'} \subset N_{k''}$ and $G \cap D_{k''} \neq \emptyset$. In either case $N(A) \cap D_k = D_k = N(G) \cap D_k$. Thus $B_f \cap G \subset N(A) \subset N(G)$.

On the other hand, if $B_f \cap G \subset N(A) \subset N(G)$, then the relationship $A \subset N(A)$ implies that $A \subset N(G)$. For each D_k, $1 \leqq k \leqq s$, $A \cap D_k \neq \emptyset$ implies that $B_f \cap G \cap D_k \subset A \cap D_k$. If $A \cap D_k = \emptyset$ but $D_k \subset N_{k'}$ for some k' such that $A \cap D_{k'} \neq \emptyset$, then $G \cap D_{k'} \neq \emptyset$ or for some k'', $G \cap D_{k''} \neq \emptyset$ and $D_{k'} \subset N_{k''}$. In either case, $B_f \cap G \cap D_k = \emptyset$. Thus $B_f \cap G \subset A \subset N(G)$.

Since $d(A')$ is constant for $A \subset A' \subset N(A)$ for a given $A \in \mathscr{A}$, (3.17) and (3.22) imply that the covariance operator of $\hat{\mu}$ is

$$\textstyle\sum_{A \in \mathscr{A}} d(A) P_{\Omega(A)} ,$$

and for $A \in \mathscr{A}$,

$$\begin{aligned} \operatorname{Cov}(\hat{\alpha}^A) &= \frac{d(A)}{\prod \{r_j : j \in B - N(A)\}} P_{\omega(A)} \\ &= [\textstyle\sum \{\sigma_G^2 / \prod \{r_j : j \in N(G) - N(A)\} : B_f \cap G \subset A \subset N(G), \\ &\qquad G \in \mathscr{G}\}] P_{\omega(A)} . \end{aligned} \tag{3.40}$$

3.4. *Sampling properties of sums of squares and mean squares.* As Kruskal (1961) notes, the sum of squares for the factors in $A \in \mathscr{A} \cup \mathscr{G}$ is $\mathrm{SS}_A = ||P_{\Omega(A)} Y||_I^2$ and the corresponding mean square MS_A is $\mathrm{SS}_A / \dim \Omega(A)$. If $A \in \mathscr{A}$, then the argument used in (3.26) may be used to show that

$$\mathrm{SS}_A = \prod \{r_j : j \in B - N(A)\} ||\hat{\alpha}^A||^2_{I(N(A))} \tag{3.41}$$

and

$$\mathrm{MS}_A = \prod \{r_j : j \in B - N(A)\} \hat{\sigma}_A^2 , \tag{3.42}$$

where

$$\hat{\sigma}_A^2 = \frac{||\hat{\alpha}^A||^2_{I(N(A))}}{[\prod \{r_j : j \in N(A) - A\}] \prod_{j \in A} (r_j - 1)} . \tag{3.43}$$

If $A \in \mathscr{G}$, $A \notin \mathscr{A}$, it is still the case that

$$\mathrm{SS}_A = \prod \{r_j : j \in B - N(A)\} ||P_{\omega(A)} \bar{Y}^{N(A)}||^2_{I(N(A))} , \tag{3.44}$$

where $P_{\omega(A)} \bar{Y}^{N(A)}$ satisfies the last two equations of (3.26). If

$$\hat{\sigma}_A^2 = \frac{||P_{\omega(A)} \bar{Y}^{N(A)}||^2_{I(N(A))}}{[\prod \{r_j : j \in N(A) - A\}] \prod_{j \in A} (r_j - 1)} , \tag{3.45}$$

then

$$\mathrm{MS}_A = \prod \{r_j : j \in B - N(A)\}\hat{\sigma}_A{}^2 . \tag{3.46}$$

The sum of squares and mean squares may be used to estimate the covariance operators $\mathrm{Cov}(\hat{\alpha}^A)$, $A \in \mathscr{A}$, and to test hypotheses concerning α^A, $A \in \mathscr{A}$, and $\mathrm{Cov}(e^G)$, $G \in \mathscr{G}$. In this section, expected values of SS_A and MS_A are computed, and the distributions of these statistics are determined under the assumption that Y has a multivariate normal distribution.

The following theorem provides expressions for $E(SS_A)$ and $E(\mathrm{MS}_A)$ for $A \in \mathscr{A} \cup \mathscr{G}$. As explained in Section 4, this theorem explicitly proves and generalizes results of Cornfield and Tukey (1956).

THEOREM 2. *If* $A \in \mathscr{A}$, *then*

$$\begin{aligned} E(\mathrm{SS}_A) = \sum \{d(A') \textstyle\prod_{j \in A'} (r_j - 1) : A \subset A' \subset N(A)\} \\ + [\textstyle\prod \{r_j : j \in B - N(A)\}] \|\alpha^A\|^2_{I(N(A))} \end{aligned} \tag{3.47}$$

and

$$\begin{aligned} E(\mathrm{MS}_A) = \sum \left\{\frac{d(A')[\prod \{1 - 1/r_j : j \in A' - A\}]}{\prod \{r_j : j \in N(A) - A'\}} : A \subset A' \subset N(A)\right\} \\ + [\textstyle\prod \{r_j : j \in B - N(A)\}]\sigma_A{}^2 , \end{aligned} \tag{3.48}$$

where

$$\sigma_A{}^2 = \frac{\|\alpha^A\|^2_{I(N(A))}}{[\prod \{r_j : j \in N(A) - A\}] \prod_{j \in A} (r_j - 1)} . \tag{3.49}$$

If $G \in \mathscr{G}$, $G \notin \mathscr{A}$, *then*

$$E(\mathrm{SS}_G) = \sum \{d(G') \textstyle\prod_{j \in G'} (r_j - 1) : G \subset G' \subset N(A)\} \tag{3.50}$$

and

$$E(\mathrm{MS}_G) = \sum \left\{\frac{d(G')[\prod \{1 - 1/r_j : j \in G' - G\}]}{\prod \{r_j : j \in N(G) - G'\}} : G \subset G' \subset N(G)\right\} . \tag{3.51}$$

PROOF. To compute $E(\mathrm{SS}_A)$ for $A \in \mathscr{A} \cup \mathscr{G}$, observe that (3.17) implies that

$$\begin{aligned} E\|P_{\Omega(A)}Y\|_I{}^2 &= E\|P_{\Omega(A)}(Y - \mu)\|_I{}^2 + \|P_{\Omega(A)}\mu\|_I{}^2 \\ &= \sum \{E\|P_{\Delta(A')}(Y - \mu)\|_I{}^2 : A \subset A' \subset N(A)\} + \|P_{\Omega(A)}\mu\|_I{}^2 . \end{aligned} \tag{3.52}$$

By (3.14), $P_{\Delta(A')}(Y - \mu)$ has covariance operator $d(A')P_{\Delta(A')}$. If $A \in \mathscr{A}$, (3.13) and Lemma 2 imply that

$$\begin{aligned} \|P_{\Omega(A)}\mu\|_I{}^2 &= \|\alpha^A \otimes [\otimes \{h^{(j)} : j \in B - N(A)\}]\|_I{}^2 \\ &= [\textstyle\prod \{r_j : j \in B - N(A)\}] \|\alpha^A\|^2_{I(N(A))} , \end{aligned} \tag{3.53}$$

while if $A \in \mathscr{G}$, $A \notin \mathscr{A}$, then

$$\|P_{\Omega(A)}\mu\|_I{}^2 = 0 . \tag{3.54}$$

Since by Lemma 1 and (3.16),

$$\dim \Delta(A') = \textstyle\prod_{j \in A'} (r_j - 1) , \tag{3.55}$$

it follows from (3.52) and Rao (1965, page 180) that (3.47) and (3.50) hold.

Since by Lemma 1, (3.2), and (3.17),

$$\dim \Omega(A) = [\textstyle\prod_{j \in A} (r_j - 1)] \prod \{r_j : j \in N(A) - A\}, \tag{3.56}$$

(3.48) and (3.51) follow. □

COROLLARY 6. *If $A \in \mathscr{A}$ and $d(A') = d(A)$ for $A \subset A' \subset N(A)$, then*

$$\begin{aligned} E(\mathrm{SS}_A) = {} & [\textstyle\prod_{j \in A} (r_j - 1)][\prod \{r_j : j \in N(A) - A\}] d(A) \\ & + [\textstyle\prod \{r_j : j \in B - N(A)\}] \|\alpha^A\|^2_{I(N(A))} \end{aligned} \tag{3.57}$$

and

$$E(\mathrm{MS}_A) = d(A) + [\textstyle\prod \{r_j ; j \in B - N(A)\}] \sigma_A{}^2 . \tag{3.58}$$

If $G \in \mathscr{G}$, $G \notin \mathscr{A}$, and $d(G') = d(G)$ for $G \subset G' \subset N(A)$, then

$$E(\mathrm{SS}_G) = [\textstyle\prod_{j \in G} (r_j - 1)][\prod \{r_j : j \in N(G) - G\}] d(G) \tag{3.59}$$

and

$$E(\mathrm{MS}_G) = d(G) . \tag{3.60}$$

If Y has a multivariate normal distribution, then the sampling properties of SS_A are summarized by the following theorem.

THEOREM 3. *If Y has a multivariate normal distribution, then the sums of squares SS_A, $A \in \mathscr{A} \cup \mathscr{G}$, are independent and SS_A, $A \in \mathscr{A} \cup \mathscr{G}$, is distributed as the sum*

$$\textstyle\sum \{d(A')\chi^2(A') : A \subset A' \subset N(A)\} ,$$

where the $\chi^2(A')$, $A \subset A' \subset N(A)$, are independent noncentral chi-square random variables with $\prod \{r_j - 1 : j \in A'\}$ degrees of freedom and noncentrality parameter $\delta(A')$. If $A \in \mathscr{G}$, $A \notin \mathscr{A}$, then $\delta(A') = 0$. If $A \in \mathscr{A}$, then

$$\delta(A') = [\textstyle\prod \{r_j : j \in B - N(A)\}] \|P_{\Lambda(A,A')}\alpha^A\|^2_{I(N(A))} / d(A') . \tag{3.61}$$

If $d(A') = d(A)$ for $A \subset A' \subset N(A)$, then $\mathrm{SS}_A/d(A)$ has a noncentral chi-square distribution with $[\prod \{r_j : j \in N(A) - A\}] \prod \{r_j - 1 : j \in A\}$ degrees of freedom and noncentrality parameter $\gamma(A)$. If $A \in \mathscr{G}$, $A \notin \mathscr{A}$, then $\gamma(A) = 0$. If $A \in \mathscr{A}$, then

$$\gamma(A) = [\textstyle\prod \{r_j : j \in B - N(A)\}] \|\alpha^A\|^2_{I(N(A))} / d(A) . \tag{3.62}$$

PROOF. The first formula follows from Rao (1965, pages 150–153) and (3.55) since

$$\|P_{\Omega(A)} Y\|_I{}^2 = \textstyle\sum \{\|P_{\Delta(A')} Y\|_I{}^2 : A \subset A' \subset N(A)\} , \tag{3.63}$$

where $P_{\Delta(A')} Y$ has covariance operator $d(A')P_{\Delta(A')}$. If $A \in \mathscr{A}$, $P_{\Delta(A')} Y$ has expected value $P_{\Lambda(A,A')}\alpha^A \otimes [\bigotimes \{h^{(j)} : j \in B - N(A)\}]$, while if $A \in \mathscr{G}$, $A \notin \mathscr{A}$, $P_{\Delta(A')} Y$ has expected value 0. The independence of the SS_A, $A \in \mathscr{A} \cup \mathscr{G}$, follows from Rao (1965, page 152) since $P_{\Omega(A_1)} \sum P_{\Omega(A_2)}$ is 0 whenever $A_1 \neq A_2$. If $d(A') = d(A)$ for $A \subset A' \subset N(A)$, then it follows from Rao (1965, page 147) that the sum $\mathrm{SS}_A/d(A)$ of the $\chi^2(A')$, $A \subset A' \subset N(A)$, has a noncentral chi-square distribution with

$$\begin{aligned} & \textstyle\sum \{\prod_{j \in A'} (r_j - 1) : A \subset A' \subset N(A)\} \\ & \qquad = [\textstyle\prod \{r_j : j \in N(A) - A\}] \prod_{j \in A} (r_j - 1) \end{aligned} \tag{3.64}$$

degrees of freedom and noncentrality parameter

$$\sum \{\delta(A') : A \subset A' \subset N(A)\} = \gamma(A)\,. \tag{3.65}$$ □

EXAMPLE 3. In Example 1, the covariance operators of $\hat{\alpha}^{\varnothing}$, $\hat{\alpha}^{\{1\}}$, $\hat{\alpha}^{\{3\}}$, and $\hat{\alpha}^{\{1,3\}}$ may be estimated by noting that by (3.39) and (3.60),

$$\frac{E(\mathrm{MS}_{\{4\}})}{r_1 r_2 r_3 r_4 r_5} = \frac{d(\{4\})}{r_1 r_2 r_3 r_4 r_5} = \frac{\sigma^2_{\{2\}}}{r_1 r_2 r_4} + \frac{\sigma^2_{\{4\}}}{r_4} + \frac{\sigma^2_{\{5\}}}{r_1 r_2 r_3 r_4 r_5}\,, \tag{3.66}$$

$$\frac{E(\mathrm{MS}_{\{1,4\}})}{r_2 r_3 r_4 r_5} = \frac{d(\{1,4\})}{r_2 r_3 r_4 r_5} = \frac{\sigma^2_{\{2\}}}{r_2 r_4} + \frac{\sigma^2_{\{1,4\}}}{r_4} + \frac{\sigma^2_{\{5\}}}{r_2 r_3 r_4 r_5}\,, \tag{3.67}$$

$$\frac{E(\mathrm{MS}_{\{3,4\}})}{r_1 r_2 r_4 r_5} = \frac{d(\{3,4\})}{r_1 r_2 r_4 r_5} = \frac{\sigma^2_{\{3,4\}}}{r_4} + \frac{\sigma^2_{\{2,3\}}}{r_1 r_2 r_4} + \frac{\sigma^2_{\{5\}}}{r_1 r_2 r_4 r_5}\,, \tag{3.68}$$

and

$$\frac{E(\mathrm{MS}_{\{1,3,4\}})}{r_2 r_4 r_5} = \frac{d(\{1,3,4\})}{r_2 r_4 r_5} = \frac{\sigma^2_{\{2,3\}}}{r_2 r_4} + \frac{\sigma^2_{\{1,3,4\}}}{r_4} + \frac{\sigma^2_{\{5\}}}{r_2 r_4 r_5} \tag{3.69}$$

(see Scheffé (1959, pages 286–287)).

4. Application to the pigeonhole model. In this section, the results of Section 3 are applied to a general version of the pigeonhole model of Cornfield and Tukey (1956). In this model, a table $Z = \{Z_{i^*} : i_j^* \in \bar{r}_j^*, j \in B\}$ is given, where I_j^* contains $r_j^* \geqq r_j \geqq 1$ elements. The observed levels of the factors of the table $Y = \{Y_i : i_j \in \bar{r}_j, j \in B\}$ are chosen by random sampling from the levels of the factors of Z so that

$$Y_i = Z_{\tau(i)}\,, \tag{4.1}$$

where $\tau(i) = \{\tau_j(\{i_{j'} : j' \in \{j\} \cup N_k\}) : j \in D_k,\ k \in \bar{s}\}$. For each $j \in D_k$, $k \in \bar{s}$, and $\{i_j : j \in N_k\}$, $\{\tau_j(\{i'_{j'} : j' \in \{j\} \cup N_k\}) : i_j' \in \bar{r}_j,\ i_{j'} = i_{j'}\ \forall j' \in N_k\}$ is a random sample without replacement from the integers of $\overline{r_j{}^*}$.

If $\mathscr{G} = \{G \subset B : G \neq \varnothing,\ G \cap D_k = \varnothing \text{ or } G \cap N_k = \varnothing,\ 1 < k \leqq s\}$, then Y satisfies

$$Y_i = \alpha^{\varnothing} + \textstyle\sum_{G \in \mathscr{G}} e^G_{\{i_j : j \in N(G)\}}\,, \tag{4.2}$$

where the e^G, $G \in \mathscr{G}$, have expectation 0, are uncorrelated, and satisfy

$$\operatorname{Cov}(e^G) = \sigma_G{}^2 \sum \{\prod \{1 - r_j/r_j{}^* : j \in G \cap [N(G) - G']\} P_{\Lambda(G,G')} : G' \subset N(G)\} \tag{4.3}$$

for some $\sigma_G{}^2$. Thus Y satisfies (3.1).

To show that Y has this form, it is first necessary to show that $E(Y_i)$ is constant for all $i \in I = \prod \{\bar{r}_j : j \in B\}$. Since if $i \in I$ and $i^* \in I^* = \prod \{\overline{r_j^*} : j \in B\}$,

$$P\{\tau(i) = i^*\} = 1/\textstyle\prod_{j \in B} r_j^* , \tag{4.4}$$

$$E(Y_i) = \alpha^{\varnothing} = \textstyle\sum_{i^* \in I^*} Z_{i^*}/\prod_{j \in B} r_j^* . \tag{4.5}$$

To define e^G for $G \in \mathscr{G}$, let $h_*^{(j)} = \{1 : i_j^* \in \overline{r_j^*}\}$,

$$\omega_*(G) = [\otimes_{j \in G} \{h_*^{(j)}\}^{\perp}] \otimes [\otimes \{R^{r_j^*} : j \in N(G) - G\}] , \tag{4.6}$$

$$\Delta_*(G) = [\otimes_{j \in G} \{h_*^{(j)}\}^{\perp}] \otimes [\otimes \{\text{span}\,\{h_*^{(j)}\} : j \in B - G\}] , \tag{4.7}$$

and

$$\begin{aligned} \Omega_*(G) &= \omega_*(G) \otimes [\otimes \{\text{span}\,\{h_*^{(j)}\} : j \in B - N(G)\}] \\ &= \oplus \{\Delta_*(G') : G \subset G' \subset N(G)\} . \end{aligned} \tag{4.8}$$

Observe that if $G' \subset B$, then $G \subset G' \subset N(G)$ and $G \subset \mathscr{G}$ if

$$G = G' - \bigcup \{G' \cap N_k : G' \cap D_k \neq \varnothing, 1 < k \leqq s\} . \tag{4.9}$$

Given Theorem 1, it follows that from Corollaries 1 and 5 that

$$R^{I^*} = \oplus_{G \in \mathscr{G}} \Omega_*(G) . \tag{4.10}$$

As a consequence

$$\begin{aligned} Z &= \textstyle\sum_{G \in \mathscr{G}} P_{\Omega_*(G)} Z \\ &= \textstyle\sum_{G \in \mathscr{G}} e_*^G \otimes [\otimes \{h_*^{(j)} : j \in B - N(G)\}] , \end{aligned} \tag{4.11}$$

where as in (3.25) and (3.26),

$$e_*^G = P_{\omega_*(G)} \bar{Z}^{N(G)} , \tag{4.12}$$

and for $i \in I^*(G) = \prod \{\overline{r_j^*} : j \in G\}$ and $G \subset B$,

$$\bar{Z}_i^G = \textstyle\sum \{Z_{i^*} : i^* \in i' \times I^*(B - G)\}/\prod \{r_j^* : j \in B - G\} . \tag{4.13}$$

Given e_*^G, $G \in \mathscr{G}$, e^G may be defined by the equations

$$e_i^G = e_{*\tau(G,i)}^G , \tag{4.14}$$

where $G \in \mathscr{G}$, $i \in I(N(G))$, and $\tau(G, i) = \{\tau_j(\{i_{j'} : j' \in \{j\} \cup N_k\}) : j \in D_k \cap N(G), k \in \bar{s}\}$. Given (4.1), (4.11), and (4.14), Y satisfies (4.2) for the given choice of e^G, $G \in \mathscr{G}$.

It is now necessary to show that the e^G are uncorrelated, have expectation 0, and satisfy (4.3). To show that $E(e^G) = 0$ for $G \in \mathscr{G}$, note that for any $i^* \in I^*(N(G))$,

$$P\{\tau_j(\{i_{j'} : j' \in \{j\} \cap N_k\}) = i_j^* \quad \forall j \in N(G)\} = 1/\textstyle\prod \{r_j^* : j \in N(G)\} . \tag{4.15}$$

Thus for $i \in I(N(G))$, $G \in \mathscr{G}$,

$$\begin{aligned} E(e_i^G) &= \frac{\sum \{e_{i^*}^G : i^* \in I^*(N(G))\}}{\prod \{r_j^* : j \in N(G)\}} \\ &= \frac{(\otimes \{h_*^{(j)} : j \in N(G)\}, P_{\omega*(G)} \bar{Z}^{N(G)})_{I^*(N(G))}}{\prod \{r_j^* : j \in N(G)\}} . \end{aligned} \tag{4.16}$$

By (4.6) and Lemma 2, $E(e^G) = 0$.

To show that the e^G are uncorrelated, suppose that $G \in \mathscr{G}$, $G' \in \mathscr{G}$, and $G \neq G'$. Suppose that k' is the largest integer such that $G \cap D_{k'} \neq G' \cap D_{k'}$, and choose j'' so that $j'' \in G' \cap D_{k'}$ but $j'' \notin G$. Given $\tau_j(\{i''_{j'} : j' \in \{j\} \cup N_k\})$, $i_j'' \in \bar{r}_j$, $j \in D_k$, $j \neq j''$, $k \in \bar{s}$, the conditional expected value of $e_i^G e_{i'}^{G'}$, $i \in I(N(G))$, $i' \in I(N(G'))$, is

$$\frac{1}{r_{j''}^*} e^G_{*\tau(G,i)} \sum \{e^{G'}_{*i^*} : i_j^* = \tau_j(\{i'_{j'} : j' \in \{j\} \cup N_k\}), j \neq j'', j \in D_k, k \in \bar{s}, i^*_{j''} \in \overline{r_j^*}\},$$

which by Lemma 6, (4.6), and (4.12) must be 0. Thus the unconditional expected value of $e_i^G e_{i'}^{G'}$ is also 0. Since $E(e_i^G) = E(e_{i'}^{G'}) = 0$, $\operatorname{Cov}(e_i^G, e_{i'}^{G'})$ is 0 and e^G and $e^{G'}$ are uncorrelated.

To determine the covariance structure of e^G for $G \in \mathscr{G}$, suppose that $i \in I(N(G))$, $i' \in I(N(G))$, $i_j = i_j'$ for $j \in G'$ and $i_j \neq i_j'$ for $j \in N(G) - G'$. Suppose first that $N(G) - G \not\subset G'$. Then for some $j'' \in D_{k'}$, $k' \in \bar{s}$, $N_{k'} \not\subset G'$, and the conditional expected value of $e_i^G e_{i'}^G$ given $\tau_j(\{i''_{j'} : j' \in \{j\} \cup N_k\})$, $i_j'' \in \bar{r}_j$, $j \in D_k$, $i \neq j''$, $k \in \bar{s}$, is

$$\begin{aligned}
&\frac{1}{r_j^{*2}} \sum \{e^G_{*i^*} e^G_{*i^{**}} : i_j^* = \tau_j(\{i_{j'} : j' \in \{j\} \cup N_k\}),\\
&\qquad i_j^{**} = \tau_j(\{i'_{j'} : j' \in \{j\} \cup N_k\}), j \neq j'', j \in D_k, k \in \bar{s}, i^*_{j''} \in \overline{r_j^*}, i^{**}_{j''} \in \overline{r_j^*}\}\\
&\qquad = \left[\frac{1}{r_j^*} \sum \{e^G_{*i^*} : i_j^* = \tau_j(\{i_{j'} : j' \in \{j\} \cup N_k\}), i^*_{j''} \in \overline{r_j^*}\}\right]\\
&\qquad\quad \times \left[\frac{1}{r_j^*} \sum \{e^G_{*i^{**}} : i_j^{**} = \tau_j(\{i'_{j'} : j' \in \{j\} \cup N_k\}), i^{**}_{j''} \in \overline{r_j^*}\}\right]\\
&\qquad = 0\,.
\end{aligned} \tag{4.17}$$

Thus $E(e_i^G e_{i'}^G) = 0$. If $N(G) - G \subset G'$ and $j'' \in N(G) - G'$, then $j'' \in G$ and the conditional expected value of $e_i^G e_{i'}^G$ given $\tau_j(\{i''_{j'} : j' \in \{j\} \cup N_k\})$, $i_j'' \in \bar{r}_j$, $j \in D_k$, $j \neq j''$, $k \in \bar{s}$, is

$$\begin{aligned}
&\frac{1}{r_j^*} \frac{1}{r_j^* - 1} \sum \{e^G_{*i^*} e^G_{*i^{**}} : i_j^* = \tau_j(\{i_{j'} : j' \in \{j\} \cup N_k\}),\\
&\qquad i_j^{**} = \tau_j(\{i_j' : j' \in \{j\} \cup N_k\}),\\
&\qquad j \in D_k, k \in \bar{s}, i^*_{j''} \in \overline{r_j^*}, i^{**}_{j''} \in \overline{r_j^*}, i^*_{j''} \neq i^{**}_{j''}\}\\
&\qquad\quad = -\frac{1}{r_j^*} \frac{1}{r_j^* - 1} \sum \{e^G_{*i^*} e^G_{*i^{**}} : i_j^* = \tau_j(\{i_j' : j' \in \{j\} \cup N_k\}),\\
&\qquad\qquad i_j^{**} = \tau_j(\{i_j' : j' \in \{j\} \cup N_k\}),\\
&\qquad\qquad j \in D_k, k \in \bar{s}, i^*_{j''} = i^{**}_{j''} \in \overline{r_j^*}\}\,.
\end{aligned} \tag{4.18}$$

Therefore,

$$E(e_i^G e_{i'}^G) = -\frac{1}{r_j^* - 1} E(e_i^G e_{i''}^G)\,, \tag{4.19}$$

where $i_j'' = i_j'$ for $j \in N(G)$, $j \neq j''$, and $i_j'' = i_j$. By induction,

$$E(e_i^G e_{i'}^G) = \left[\prod \left\{-\frac{1}{r_j^* - 1} : j \in N(G) - G'\right\}\right] E(e_i^G)^2\,, \tag{4.20}$$

where

$$E(e_i^G)^2 = \frac{\|P_{\omega_*(G)}\bar{Z}^{N(G)}\|^2_{I^*(N(G))}}{\prod\{r_j^* : j \in N(G)\}}. \tag{4.21}$$

Thus e^G has a permutation-invariant covariance structure. If

$$\sigma_G^2 = \frac{\|P_{\omega_*(G)}\bar{Z}^{N(G)}\|^2_{I^*(N(G))}}{[\prod\{r_j^* : j \in N(G) - G\}]\prod\{r_j^* - 1 : j \in G\}}, \tag{4.22}$$

then (3.11) implies that

$$\mathrm{Cov}\,(e^G) = \sum\{b(G, G')P_{\Lambda(G,G')} : G' \subset N(G)\}, \tag{4.23}$$

where for $G' \subset N(G)$,

$$\begin{aligned} b(G, G') &= \sigma_G^2 \sum \Big\{\Big[\prod_{j \in G}\Big(\frac{r_j^* - 1}{r_j^*}\Big)\Big]\Big[\prod\Big\{-\frac{1}{r_j^* - 1} : j \in N(G) - G''\Big\}\Big] \\ &\qquad \times [\prod\{r_j - 1 : j \in [N(G) - G'] \cap [N(G) - G'']\}] \\ &\qquad s(G'', G') : N(G) - G \subset G'' \subset N(G)\Big\} \\ &= \sigma_G^2\Big[\prod_{j \in G}\Big(\frac{r_j^* - 1}{r_j^*}\Big)\Big]\Big[\prod\Big\{1 - \frac{r_j - 1}{r_j^* - 1} : j \in G \cap [N(G) - G']\Big\}\Big] \\ &\qquad \times \Big[\prod\Big\{1 + \frac{1}{r_j^* - 1} : j \in G \cap G'\Big\}\Big] \\ &= \sigma_G^2 \prod\{1 - r_j/r_j^* : j \in G \cap [N(G) - G']\}. \end{aligned} \tag{4.24}$$

Hence (4.3) holds.

By (3.26), the grand mean $\hat{\alpha}^\varnothing = \bar{Y}^\varnothing$ is the Gauss–Markov estimate of $\alpha^\varnothing$. By (3.15) and (3.28), the variance of $\bar{Y}^\varnothing$ is

$$\begin{aligned} \frac{d(\varnothing)}{\prod\{r_j : j \in B\}} &= \sum\{\sigma_G^2[\prod\{1 - r_j/r_j^* : j \in G\}] \\ &\qquad \times [\prod\{r_j : j \in B - N(G)\}] : G \in \mathscr{G}\}. \end{aligned} \tag{4.25}$$

Special cases of this formula are given in Wilks (1962, pages 228 and 231).

To find expected mean squares MS_G, $G \in \mathscr{G}$, note that since $A \subset N(G)$ is equivalent to $N(A) \subset N(G)$ for $A \in \mathscr{G}$ and since $A \subset N(G)$ implies that $N(A) - A \subset N(G) - G$, it follows that for any A', $A \subset A' \subset N(A)$,

$$\begin{aligned} d(A') &= \sum\{\sigma_G^2[\prod\{1 - r_j/r_j^* : j \in G \cap [N(G) - A]\}] \\ &\qquad \times [\prod\{r_j : j \in B - N(G)\}] : A \subset N(G), G' \in \mathscr{G}\}. \end{aligned} \tag{4.26}$$

By Corollary 6,

$$\begin{aligned} E(\mathrm{MS}_G) &= \sum\{\sigma_{G'}^2[\prod\{1 - r_j/r_j^* : j \in G' \cap [N(G') - G]\}] \\ &\qquad \times [\prod\{r_j : j \in B - N(G')\}] : G \subset N(G'), G \in \mathscr{G}\}. \end{aligned} \tag{4.27}$$

This equation is explicitly derived by Cornfield and Tukey (1956) in the case of the $r \times c$ table with n observations per cell and in the case of the $r \times c \times s$ table with n observations per cell. The general rule provided by (4.27) is stated but not explicitly proven.

REFERENCES

[1] CORNFIELD, J. and TUKEY, J. W. (1956). Average values of mean squares in factorials. *Ann. Math. Statist.* **27** 907–949.

[2] EATON, M. L. (1970). Gauss–Markov estimation for multivariate linear models: a coordinate-free approach. *Ann. Math. Statist.* **41** 528–538.

[3] GOOD, I. J. (1958). The interaction algorithm and practical Fourier analysis. *J. Roy. Statist. Soc. Ser. B* **20** 361–372.

[4] HABERMAN, S. J. (1974). *The Analysis of Frequency Data.* Univ. of Chicago Press.

[5] HALMOS, P. (1958). *Finite-dimensional Vector Spaces.* Van Nostrand, Princeton.

[6] HOFFMAN, K. and KUNZE, R. (1961). *Linear Algebra.* Prentice-Hall, Englewood Cliffs, N. J.

[7] KEMPTHORNE, O. and WILK, M. (1955). Fixed, mixed, and random models. *J. Amer. Statist. Assoc.* **50** 1144–1167.

[8] KEMPTHORNE, O. and WILK, M. (1956). Some aspects of the analysis of factorial experiments in a completely randomized design. *Ann. Math. Statist.* **27** 950–985.

[9] KRUSKAL, W. H. (1961). The coordinate-free approach to Gauss–Markov estimation, and its applications to missing and extra observations. *Fourth Berkeley Symp. Math. Statist. Prob.* **1** 435–451. Univ. of California Press.

[10] KRUSKAL, W. H. (1968). When are least squares and Gauss-Markov estimators identical? A coordinate-free approach. *Ann. Math. Statist.* **39** 70–75.

[11] LOOMIS, L. and STERNBERG, S. (1968). *Advanced Calculus.* Addison-Wesley, Reading, Mass.

[12] RAO, C. R. (1965). *Linear Statistical Inference and Its Applications.* Wiley, New York.

[13] RAO, C. R. and MITRA, S. K. (1971). *Generalized Inverse of Matrices and Its Applications.* Wiley, New York.

[14] SCHEFFÉ, H. (1956a). A "mixed model" for the analysis of variance. *Ann. Math. Statist.* **27** 23–36.

[15] SCHEFFÉ, H. (1956b). Alternative models for the analysis of variance. *Ann. Math. Statist.* **27** 251–271.

[16] SCHEFFÉ, H. (1959). *The Analysis of Variance.* Wiley, New York.

[17] WILKS, S. S. (1962). *Mathematical Statistics.* Wiley, New York.

[18] ZYSKIND, G. (1962). On structure, relation, Σ, and expectation of mean squares. *Sankhyā Ser. A* **24** 115–148.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CHICAGO
118 EAST 58TH STREET
CHICAGO, ILLINOIS