SAMPLING VARIANCES OF ESTIMATES OF COMPONENTS OF VARIANCE

BY S. R. SEARLE

New Zealand Dairy Board, Wellington, N. Z.[1]

1. Outline. In earlier work (4) matrix methods have been developed for obtaining the sampling variances of estimates of components of variance. These rely on the fact that if $y = \mathbf{x}'F\mathbf{x}$ is a function of variables $\mathbf{x}$, having a multinormal distribution with variance-covariance matrix V, then the variance of y is given by

$$\text{var}(y) = 2 \text{ tr } (VF)^2. \tag{1}$$

The use of the method was demonstrated by obtaining for the case of a 1-way classification with unequal numbers in the sub-classes, the sampling variances of the estimates of variance components, as summarized in (1); it was then extended to the sampling variances of estimates of components of covariance.

The present paper makes further use of this matrix technique to obtain the sampling variances of estimates of components of variance from data in a 2-way classification having unequal sub-class numbers. The model assumed is Eisenhart's Model II, [2], and the method of estimating the components is taken to be Henderson's Method 1, [3].

2. Model and analysis of variance. The observations x_{ijk} are taken as having the linear model

$$x_{ijk} = \mu + A_i + B_j + (AB)_{ij} + \epsilon_{ijk},$$

with $k = 1 \cdots n_{ij}$, $i = 1 \cdots a$, and $j = 1 \cdots b$. μ is a general mean, A_i and B_j are main effects, $(AB)_{ij}$ is an interaction and ϵ_{ijk} is residual error. Under the assumptions of the model, all terms (except μ) are taken as being normally distributed, with zero means, and variances σ_a^2, σ_b^2, σ_{ab}^2, and σ_ϵ^2, which we will write as α, β, γ and ϵ respectively.

For a sample of N observations in N' cells of this 2-way classification an analysis of variance can be written as

Term	d-f	Sums of Square
Between A classes	$a - 1$	$T_a - T_f = S_a$
Between B classes	$b - 1$	$T_b - T_f = S_b$
Interaction $A \times B$	$N' - a - b + 1$	$T_{ab} - T_a - T_b + T_f = S_{ab}$
Residual	$N - N'$	$T_0 - T_{ab} = S_w$
Total	$N - 1$	$T_0 - T_f$

where the T's are uncorrected sums of squares. With $n_{i.} = \sum_j n_{ij}$, and $n_{.j}$

Received December 3, 1956; revised September 5, 1957.

[1] The author's present address is Animal Husbandry Dept., Cornell University, Ithaca, New York.

$= \sum_i n_{ij}$, and using customary notation for means,

$$T_a = \sum_i n_{i.} \bar{x}_{i..}^2, \qquad T_b = \sum_j n_{.j} \bar{x}_{.j.}^2,$$

$$T_{ab} = \sum_i \sum_j n_{ij} \bar{x}_{ij.}^2, \quad T_f = N\bar{x}_{...}^2, \qquad \text{and } T_0 = \sum_i \sum_j \sum_k x_{ijk}^2.$$

We may note in passing that not all the expressions in the "sums of squares" column are in fact sums of squares, notably the interaction term. It would be more correct to label this column "quadratic forms" but the terminology "sums of squares" has historical precedence and will be retained.

Henderson's first method [3] for estimating the components of variance is to equate each of the first four lines in the above analysis to its expected value. Denoting the resulting estimates of α, β, γ, and ϵ as $\hat{\alpha}$, $\hat{\beta}$, $\hat{\gamma}$, and $\hat{\epsilon}$, the equations for obtaining them are

$$(2) \quad \left\{\begin{aligned} T_a - T_f &= S_a = (N - k_1)\hat{\alpha} + (k_{12} - k_2)\hat{\beta} \\ &\qquad + (k_{12} - k_3)\hat{\gamma} + (a - 1)\hat{\epsilon} \\ T_b - T_f &= S_b = (k_{21} - k_1)\hat{\alpha} + (N - k_2)\hat{\beta} \\ &\qquad + (k_{21} - k_3)\hat{\gamma} + (b - 1)\hat{\epsilon} \\ T_{ab} - T_a - T_b + T_f &= S_{ab} = (k_1 - k_{21})\hat{\alpha} + (k_2 - k_{12})\hat{\beta} \\ &\qquad + (N - k_{12} - k_{21} + k_3)\hat{\gamma} \\ &\qquad + (N' - a - b + 1)\hat{\epsilon} \end{aligned}\right\}$$

$$(3) \quad T_0 - T_{ab} = S_w = (N - N')\hat{\epsilon}$$

where the k's are functions of the n_{ij}'s, namely

$$k_{12} = \sum_i \frac{\sum_j n_{ij}^2}{n_{i.}} \qquad k_{21} = \sum_j \frac{\sum_i n_{ij}^2}{n_{.j}},$$

$$k_1 = \frac{1}{N} \sum_i n_{i.}^2, \qquad k_2 = \frac{1}{N} \sum n_{.j}^2, \qquad \text{and } k_3 = \frac{1}{N} \sum n_{ij}^2.$$

3. Variances required. In the analysis of variance S_w has a χ^2-distribution with $N - N'$ degrees of freedom. Hence, from Eq. (3) the variance of $\hat{\epsilon}$ is

$$\sigma_{\hat{\epsilon}}^2 = \frac{2\epsilon^2}{N - N'}.$$

Using (3), Eqs. (2) give $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ as linear functions of S_a, S_b, S_{ab}, and $\hat{\epsilon}$. But S_w, and hence $\hat{\epsilon}$, is distributed independently of S_a, S_b, and S_{ab}. Hence the variances and covariances of $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ can be obtained as linear functions of $\sigma_{\hat{\epsilon}}^2$ and the variances and covariances of S_a, S_b, and S_{ab}. By the nature of the S's it is easier to consider the variances and covariances of T_a, T_b, T_{ab}, and T_f.

Writing P for the matrix of coefficients of $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ in Eqs. (2) these equations can be written as

$$\begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ -1 & -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} T_a \\ T_b \\ T_{ab} \\ T_f \end{pmatrix} = P \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \\ \hat{\gamma} \end{pmatrix} + \hat{\epsilon} \begin{pmatrix} a - 1 \\ b - 1 \\ N' - a - b + 1 \end{pmatrix},$$

which we may write as

$$H\mathbf{t} = P\mathbf{v} + \hat{\epsilon}\mathbf{m}.$$

Since $\hat{\epsilon}$ is independent of the terms in $H\mathbf{t}$, the variance-covariance matrix of $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$, var $(\mathbf{vv}')$, can be expressed in terms of the variance-covariance matrix of the T's, var $(\mathbf{tt}')$, as

$$\text{(4)} \qquad \text{var}\,(\mathbf{vv}') = P^{-1}[H \text{ var}\,(\mathbf{tt}')H' + \mathbf{mm}'\sigma_{\hat{\epsilon}}^2]P^{-1'}$$

and

$$\text{cov}\,(\hat{\epsilon}\mathbf{v}) = -P^{-1}\mathbf{m}\sigma_{\hat{\epsilon}}^2 .$$

The unknown term in these expressions is var $(\mathbf{tt}')$ the variance-covariance matrix of T_a, T_b, T_{ab}, and T_f, which we now proceed to obtain, term by term.

4. Matrix definitions and expressions. Let U be a matrix having a one for every element, its order being denoted by subscripts, thus:—

U-matrix (*all elements 1*)	Order
$U_{ij,kl}$	$n_{ij} \times n_{kl}$
$U_{ij,k.}$	$n_{ij} \times n_{k.}$
U_{ij}	$n_{ij} \times n_{ij}$
$U_{i.}$	$n_{i.} \times n_{i.}$
U_N	$N \times N$

Define W-matrices in terms of the U's:

$$W_{ij} = \frac{1}{n_{ij}} U_{ij} ,$$

$$W_{i.} = \frac{1}{n_{i.}} U_{i.} ,$$

$$W_{.j} = \frac{1}{n_{.j}} U_{.j}, \text{ and } W_N = \frac{1}{N} U_N .$$

Then C-matrices are defined, of order $N \times N$, whose only non-zero submatrices are W's along the diagonal:

C_a has $W_{i.}(i = 1 \cdots a)$ in the diagonal,
C_b has $W_{.j}(j = 1 \cdots b)$ in the diagonal,
C_{ab} has $W_{ij}(i = 1 \cdots a, j = 1 \cdots b)$, in the diagonal.

Finally we define D-matrices, the same as C-matrices only having U-matrices instead of W-matrices in their diagonals.

Let $\mathbf{x}'$ be the row vector of the N x_{ijk}'s, arrayed in order, $k = 1 \cdots n_{ij}$, within j-classes, within each i-class; i.e.,

$$\mathbf{x}' = (x_{111} \cdots x_{11n_{11}} \quad x_{121} \cdots x_{12n_{12}} \cdots x_{ab1} \cdots x_{abn_{ab}}).$$

Then if $\mathbf{w}'$ is the vector of the x's arrayed in k-order within i-classes within each j-class, $\mathbf{w}'$ will be a transform of $\mathbf{x}'$, $\mathbf{w}' = \mathbf{x}'R'$, say, where R is an orthogonal elementary operational matrix of order N, of identity matrices I.

The T's can now be expressed in terms of these vectors and matrices:

$$\begin{aligned} T_a &= \mathbf{x}'C_a\mathbf{x}, \\ T_b &= \mathbf{w}'C_b\mathbf{w} = \mathbf{x}'R'C_bR\mathbf{x} = \mathbf{x}'B\mathbf{x}\text{, say.} \\ T_{ab} &= \mathbf{x}'C_{ab}\mathbf{x} = \mathbf{w}'C_{ba}\mathbf{w}, \\ T_f &= \mathbf{x}'U_N\mathbf{x}. \end{aligned}$$

In C_{ab} the W_{ij} in the diagonal are in j-order within i-order; in C_{ba} they are in i-order within j-order.

V, the variance-covariance matrix of the x_{ijk}'s appropriate to $\mathbf{x}'$ can be written as

$$V = J + K,$$

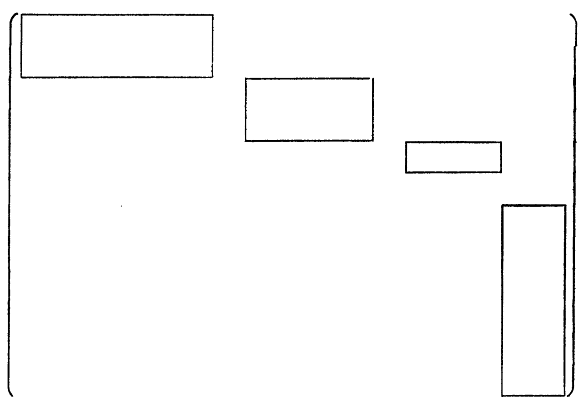
where

$$J = \alpha D_a + (\beta + \gamma)D_{ab} + \epsilon I,$$

and

$$K = \begin{pmatrix} 0 & K_{12} & K_{13} \cdots K_{1a} \\ K_{21} & 0 & \\ \vdots & & \\ K_{a1} & & 0 \end{pmatrix}$$

with $K_{ii'}$, $i \neq i'$, $i, i' = 1 \cdots a$, of order $n_{i.} \times n_{i'.}$ has all elements zero except those in b rectangular matrices $\beta U_{ij,i'j}$, $j = 1 \cdots b$. These b matrices lie "corner to corner" across $K_{ii'}$ thus:



For example the V matrix for a sample of 7 observations with $n_{11} = 1$, $n_{12} = 2$, $n_{21} = 3$ and $n_{22} = 1$ would be

$$\left(\begin{array}{ccc|cccc}
\alpha+\beta+\gamma+\epsilon & \alpha & \alpha & \beta & \beta & \beta & \cdot \\
\alpha & \alpha+\beta+\gamma+\epsilon & \alpha+\beta+\gamma & \cdot & \cdot & \cdot & \beta \\
\alpha & \alpha+\beta+\gamma & \alpha+\beta+\gamma+\epsilon & \cdot & \cdot & \cdot & \beta \\
\hline
\beta & \cdot & \cdot & \alpha+\beta+\gamma+\epsilon & \alpha+\beta+\gamma & \alpha+\beta+\gamma & \alpha \\
\beta & \cdot & \cdot & \alpha+\beta+\gamma & \alpha+\beta+\gamma+\epsilon & \alpha+\beta+\gamma & \alpha \\
\beta & \cdot & \cdot & \alpha+\beta+\gamma & \alpha+\beta+\gamma & \alpha+\beta+\gamma+\epsilon & \alpha \\
\cdot & \beta & \beta & \alpha & \alpha & \alpha & \alpha+\beta+\gamma+\epsilon
\end{array}\right)$$

This is the variance-covariance matrix appropriate to $\mathbf{x}$: that for $\mathbf{w}$ will be RVR'.

5. Variances and covariances. T_a, T_b, T_{ab}, and T_f have now been expressed in the form $\mathbf{x}'F\mathbf{x}$, and the variance-covariance matrix has also been obtained. The sampling variances of the T's will be found from (1), by evaluating 2 trace $(VF)^2$ for each of them.

5.1. $$\text{var}\,(T_a) = 2\,\text{tr}\,(VC_a)^2.$$

VC_a can be expressed as

$$(VC_a) = \begin{pmatrix} P_{11} P_{12} \cdots P_{1a} \\ P_{21} \ddots & & \\ \vdots & & \\ P_{a1} & & P_{aa} \end{pmatrix},$$

where P_{ii} is a column of matrices $x_{ij}U_{ij,i.}$, $(j = 1 \cdots b)$ with

$$x_{ij} = (1/n_{i.})(n_{i.}\alpha + n_{ij}\beta + n_{ij}\gamma + \epsilon).$$

Similarly $P_{ii'}$ is a column of matrices $w_{ij}U_{ij,i'.}$, $(j = 1 \cdots b)$, with

$$w_{i'j} = (n_{i'j}/n_{i'.})\beta.$$

VC_a has here been partitioned into P-matrices, which themselves have been partitioned into sub-matrices of the U-type. Trace $(VC_a)^2$ will therefore depend on two properties of these U matrices, that

(5) $$U_{ij,pq}U_{pq,rs} = n_{pq}U_{ij,rs}$$

and

$$\text{tr}\,(U_{ij}) = n_{ij}.$$

Hence

(6) $$\text{tr}\,(U_{ij,pq}U_{pq,ij}) = n_{ij}n_{pq}.$$

Using these results we have

$$\begin{aligned}\operatorname{tr}(VC_a)^2 &= \sum_i \sum_{i'} \operatorname{tr}(P_{ii'}P_{i'i}) \\ &= \sum_i \operatorname{tr}(P_{ii}^2) + \sum_i \sum_{i'\neq i} \operatorname{tr}(P_{ii'}P_{i'i}) \\ &= \sum_i \sum_j n_{ij}x_{ij} \sum_j n_{ij}x_{ij} + \sum_i \sum_{i'\neq i} \sum_j n_{ij}w_{i'j} \sum_j n_{i'j}w_{ij}.\end{aligned}$$

On substituting for x_{ij} and w_{ij} this gives

$$\tfrac{1}{2}\operatorname{var}(T_a) = \sum_i \Big[\sum_j n_{jj}(n_{i.}\alpha + n_{ij}\beta + n_{ij}\gamma + \epsilon)/n_{i.}\Big]^2 + \sum_i \sum_{i'\neq i} \frac{\big(\sum_j n_{ij}n_{i'j}\big)^2}{n_{i.}n_{i'.}}\beta^2.$$

5.2. $$\operatorname{var}(T_{ab}) = 2\operatorname{tr}(VC_{ab})^2.$$

V and C_{ab} are such that their product can be written as

$$VC_{ab} = L + K,$$

where K is as in V, and

$$L = \alpha D_a + (\beta + \gamma)D_{ab} + \epsilon C_{ab}.$$

Hence,

$$VC_{ab} = V + \epsilon(C_{ab} - I). \tag{7}$$

Since V and C_{ab} are symmetric, VC_{ab} is also, and hence squaring (7) gives

$$(VC_{ab})^2 = V^2 + \epsilon^2(C_{ab}^2 - I).$$

Hence,

$$\tfrac{1}{2}\operatorname{var}(T_{ab}) = \operatorname{tr} V^2 + \epsilon^2(\operatorname{tr} C_{ab}^2 - \operatorname{tr} I)$$

$$= \sum_i \sum_j n_{ij}[(\alpha+\beta+\gamma+\epsilon)^2 + (n_{ij}-1)(\alpha+\beta+\gamma)^2 + (n_{i.} - n_{ij})\alpha^2 + (n_{.j} - n_{ij})\beta^2] + \epsilon^2\left[\sum_i \sum_j n_{ij}n_{ij}\frac{1}{n_{ij}^2} - N\right],$$

which reduces to

$$\tfrac{1}{2}\operatorname{var}(\mathrm{T}_{ab}) = \sum_i \sum_j n_{ij}[n_{ij}(\alpha+\beta+\gamma+\epsilon/n_{ij})^2 + (n_{i.} - n_{ij})\alpha^2 + (n_{.j} - n_{ij})\beta^2].$$

5.3. $$\operatorname{var}(T_f) = 2\operatorname{tr}(VW_N)^2.$$

Similar to the form of the P-matrices in 5.2, VW_N can be expressed as a column of matrices $y_{ij}U_{ij.N}$, $(i = 1 \cdots a, j = 1 \cdots b)$, where

$$y_{ij} = (n_{i.}\alpha + n_{.j}\beta + n_{ij}\gamma + \epsilon)/N.$$

Hence,

$$\text{tr}\,(VW_N)^2 = \sum_i \sum_j n_{ij} y_{ij} \sum_i \sum_j n_{ij} y_{ij},$$

giving

$$\tfrac{1}{2}\,\text{var}\,(T_f) = [\sum_i \sum_j n_{ij}(n_{i.}\alpha + n_{.j}\beta + n_{ij}\gamma + \epsilon(/N]^2.$$

5.4. In general, for any two square matrices of the same order, A and B say, it can be shown that $\text{tr}\,(A + B)^2 = \text{tr}\,A^2 + \text{tr}\,B^2 + 2\,\text{tr}\,AB$. If then, $\hat{a}$ and $\hat{b}$ are two function of the same set of variables such that $\text{var}\,(\hat{a}) = 2\,\text{tr}\,A^2$, and $\text{var}\,(\hat{b}) = 2\,\text{tr}\,B^2$, it follows at once that

$$\text{cov}\,(\hat{a}\hat{b}) = 2\,\text{tr}\,AB = 2\,\text{tr}\,BA. \tag{8}$$

This result will be used for obtaining the covariances among T_a, T_b, T_{ab}, and T_f.

5.5. $$\text{cov}\,(T_a, T_{ab}) = 2\,\text{tr}\,(VC_a)(VC_{ab}).$$

In 5.1 VC_a has been partitioned into P_{ii}'s and $P_{ii'}$'s. If VC_{ab}, expressed as $L + K$ in 5.2 is partitioned in the same manner, into L_{ii}'s and $K_{ii'}$'s, then

$$\tfrac{1}{2}\,\text{cov}\,(T_a, T_{ab}) = \sum_i \sum_{l=1}^{n_{i.}} (\text{inner product of } l\text{'th row of } P_{ii} \text{ and } l \text{ th column of } L_{ii})$$
$$+ \sum_i \sum_{i' \neq i} \sum_{l=1} (\text{inner product of } l\text{'th row of } P_{ii'} \text{ and } l \text{ th column of } K_{i'i})$$

and after substitution this reduces to

$$\tfrac{1}{2}\,\text{cov}\,(T_a, T_{ab}) = \sum_i \sum_j n_{ij}(n_{i.}\alpha + n_{ij}\beta + n_{ij}\gamma + \epsilon)^2/n_{i.}$$
$$+ \beta^2 \sum_i \sum_j n_{ij}^2(n_{.j} - n_{ij})/n_{i.}.$$

5.6. $$\tfrac{1}{2}\,\text{cov}\,(T_a, T_f) = \text{tr}\,(VW_N)(VC_a).$$

Using 5.1 and 5.4, and Eq. (6), this can be expressed as

$$\tfrac{1}{2}\,\text{cov}\,(T_a, T_f) = \sum_i \sum_j n_{ij} y_{ij} \,(\sum_j n_{ij} x_{ij} + \sum_j \sum_{i' \neq i} n_{i'j} w_{ij}),$$

which on substitution for the x's, y's and w's, reduces to

$$\tfrac{1}{2}\,\text{cov}\,(T_a, T_f) = \sum_i \sum_j \frac{n_{ij}}{N}(n_{i.}\alpha + n_{.j}\beta + n_{ij}\gamma + \epsilon)$$
$$\cdot \left(n_{i.}\alpha + \beta \frac{\sum_j n_{.j} n_{ij}}{n_{i.}} + \gamma \frac{\sum_j n_{ij}^2}{n_{i.}} + \epsilon\right).$$

5.7. $$\tfrac{1}{2}\,\text{cov}\,(T_{ab}, T_f) = \text{tr}\,(VW_N)(VC_{ab})$$
$$= \sum_i \sum_j n_{ij} y_{ij}\,[\textstyle\sum \text{terms in } ij\text{'th column of } V + \epsilon(C_{ab} - I)]$$
$$= \sum_i \sum_j n_{ij}(n_{i.}\alpha + n_{.j}\beta + n_{ij}\gamma + \epsilon)^2/N.$$

5.8. In Eq. (8) it is required that $\hat{a}$ and $\hat{b}$ be functions of the same set of variables; therefore, in terms of paragraph 4, the covariance of T_a and T_b must be expressed as

$$\text{cov}\,(T_a\,,\,T_b) = 2\,\text{tr}\,(VC_a)(VB).$$

This covariance is a little more cumbersome to evaluate than previous ones; the method used is essentially a generalization of earlier paragraphs.

B is the same form as V, (4.5) but with matrices $(1/n_{.j})U_{ij}$ in the diagonal $j = 1 \cdots b$, for $i = 1 \cdots a$, and with $K_{ii'}$-matrices having terms

$$(1/n.j)U_{ij,i'j}\,.$$

Now partition V into matrices $(V)_{ij:kl}$ of order $n_{ij} \times n_{kl}$, there being four different forms of this matrix according as k and l are equal or not equal to i and j respectively, namely:

$$\begin{aligned}
(V)_{ij:ij} &= (\alpha + \beta + \gamma)U_{ij} + \epsilon I; & \\
(V)_{ij:il} &= \alpha U_{ij,i} & \text{for } l \neq j; \\
(V)_{ij:kj} &= \beta U_{ij,kj} & \text{for } k \neq l; \\
(V)_{ij:kl} &= 0.U_{ij,kl} \text{ a zero matrix,} & \text{for } k \neq i,\ l \neq j.
\end{aligned}$$

B can be partitioned similarly for $k \neq i$ and $l \neq j$:

$$\begin{aligned}
(B)_{ij:ij} &= \frac{1}{n_{.j}}\,U_{ij}\,, \\
(B)_{il:ij} &= 0.U_{il,ij}\,, \\
(B)_{kj:ij} &= \frac{1}{n_{:j}}\,U_{kj,ij}\,, \\
(B)_{kl:ij} &= 0.U_{kl,ij}\,.
\end{aligned}$$

Consider now the identity

$$(9) \qquad (VB)_{pq:tu} = \sum_f \sum_g V_{pq:fg} B_{fg:tu}\,,$$

whose right-hand side can be expanded as

$$(V)_{pq:tu}(B)_{tu:tu} + \sum_{f \neq t} (V)_{pq:fu}(B)_{fu:tu} + \sum_{g \neq u} (V)_{pq:tg}(B)_{tg:tu} + \sum_{f \neq t} \sum_{g \neq u} (V)_{pq:fg}(B)_{fg:tu}$$

or as

$$(V)_{pq:pq}(B)_{pq:tu} + \sum_{f \neq p} (V)_{pq:fq}(B)_{fq:tu} + \sum_{g \neq q} (V)_{pq:pg}(B)_{pg:tu} + \sum_{f \neq p} \sum_{g \neq q} (V)_{pq:fg}(B)_{fg:tu}\,.$$

These expressions are true for any values of the subscripts.

Applying this identity to the partitioned forms of V and B given above, and using the principle of (5) in 5.1 gives

$$(VB)_{ij:ij} = n_{ij}(\alpha + \beta + \gamma)/n_{.j} U_{ij} + \epsilon/n_{.j} U_{ij} + \beta \sum_{k \neq i} \frac{1}{n_{.j}} U_{ij,kj} U_{kj,ij}$$

$$= (n_{ij}\alpha + n_{.j}\beta + n_{ij}\gamma + \epsilon)/n_{.j} U_{ij} = b_{ij} U_{ij}, \quad \text{say.}$$

Similarly, for $r \neq i$, and $s \neq j$,

$$(VB)_{is:ij} = \alpha \frac{n_{ij}}{n_{.j}} U_{is,ij} = b'_{ij} U_{is,ij}, \quad \text{say,}$$

$$(VB)_{rj:ij} = (n_{ij}\alpha + n_{.j}\beta + n_{ij}\gamma + \epsilon)/n_{.j} U_{rj,ij} = b_{ij} U_{rj,ij},$$

$$(VB)_{rs:ij} = \alpha \frac{n_{rj}}{n_{.j}} U_{rs,ij} = b'_{rj} U_{rs,ij}.$$

Likewise:

$$(VC_a)_{ij:ij} = (n_{i.}\alpha + n_{ij}\beta + n_{ij}\gamma + \epsilon)/n_{i.} U_{ij} = a_{ij} U_{ij}, \quad \text{say,}$$

$$(VC_a)_{ij:is} = (n_{i.}\alpha + n_{ij}\beta + n_{ij}\gamma + \epsilon)/n_{i.} U_{ij,is} = a_{ij} U_{ij,is},$$

$$(VC_a)_{ij:rj} = \beta \frac{n_{rj}}{n_{r.}} U_{ij,rj} = a'_{rj} U_{ij,rj}, \quad \text{say,}$$

$$(VC_a)_{ij:rs} = \beta \frac{n_{rj}}{n_{r.}} U_{ij,rs} = a'_{rj} U_{ij,rs}.$$

Now

$$\tfrac{1}{2} \operatorname{cov} (T_a T_b) = \operatorname{tr} (VC_a)(VB)$$

$$= \sum_i \sum_j \operatorname{tr} (VC_a \cdot VB)_{ij:ij}$$

$$= \sum_i \sum_j [\operatorname{tr} \sum_r \sum_s (VC_a)_{ij:rs}(VB)_{rs:ij}].$$

Applying the identity (9) and results (5) and (6) again, and using the forms of the elements of the sub-matrices of VC_a and VB given above, gives

$$\tfrac{1}{2} \operatorname{cov} (T_a, T_b)$$

$$= \sum_i \sum_j n_{ij}\{n_{ij} a_{ij} b_{ij} + \sum_{s \neq j} n_{is} a_{ij} b'_{ij} + \sum_{r \neq i} n_{rj} a'_{rj} b_{rj} + \sum_{r \neq i} \sum_{s \neq j} n_{rs} a'_{rj} b'_{rj}\}.$$

On substituting for the a's and b's, this reduces to

$$\tfrac{1}{2} \operatorname{cov} (T_a, T_b) = \sum_i \sum_j \frac{n_{ij}^2}{n_{i.} n_{.j}} (n_{i.}\alpha + n_{.j}\beta + n_{ij}\gamma + \epsilon)^2.$$

5.9. We have now found some of the variances and covariances of T_a, T_b, T_{ab} and T_f. These and those which follow from them by symmetry, are summarized in the following table.

VARIANCES AND COVARIANCES OF UNCORRECTED SUMS OF SQUARES

var (T_a)

$$= 2\left\{\sum_i [\sum_j n_{ij}(n_{i.}\alpha + n_{ij}\beta + n_{ij}\gamma + \epsilon)/n_{i.}]^2 + \sum_i \sum_{i' \neq i} \frac{(\sum_j n_{ij}n_{i'j})^2}{n_{i.}n_{i'.}} \beta^2\right\}$$

var (T_b)

$$= 2\left\{\sum_j [\sum_i n_{ij}(n_{ij}\alpha + n_{.j}\beta + n_{ij}\gamma + \epsilon)/n_{.j}]^2 + \sum_j \sum_{j' \neq j} \frac{(\sum_i n_{ij}n_{ij'})^2}{n_{.j}n_{.j'}} \alpha^2\right\}$$

var (T_{ab})

$$= 2 \sum_i \sum_j n_{ij}[n_{ij}(\alpha + \beta + \gamma + \epsilon/n_{ij})^2 + (n_{i.} - n_{ij})\alpha^2 + (n_{.j} - n_{ij})\beta^2]$$

var (T_f)

$$= 2[\sum_i \sum_j n_{ij}(n_{i.}\alpha + n_{.j}\beta + n_{ij}\gamma + \epsilon)/N]^2$$

cov (T_a, T_b)

$$= 2 \sum_i \sum_j \frac{n_{ij}^2}{n_{i.}n_{.j}} (n_{i.}\alpha + n_{.j}\beta + n_{ij}\gamma + \epsilon)^2$$

cov (T_a, T_{ab})

$$= 2\{\sum_i \sum_j n_{ij}(n_{i.}\alpha + n_{ij}\beta + n_{ij}\gamma + \epsilon)^2/n_{i.} + \beta^2 \sum_i \sum_j n_{ij}^2(n_{.j} - n_{ij})/n_{i.}\}$$

cov (T_a, T_f)

$$= 2 \sum_i \sum_j \frac{n_{ij}}{N} (n_{i.}\alpha + n_{.j}\beta + n_{ij}\gamma + \epsilon) \left(n_{i.}\alpha + \beta \frac{\sum_j n_{.j}n_{ij}}{n_{i.}} + \gamma \frac{\sum_j n_{ij}^2}{n_{i.}} + \epsilon\right)$$

cov (T_b, T_{ab})

$$= 2\{\sum_i \sum_j n_{ij}(n_{ij}\alpha + n_{.j}\beta + n_{ij}\gamma + \epsilon)^2/n_{.j} + \alpha^2 \sum_i \sum_j n_{ij}^2(n_{i.} - n_{ij})/n_{.j}\}$$

cov (T_b, T_f)

$$= 2 \sum_i \sum_j \frac{n_{ij}}{N} (n_{i.}\alpha + n_{.j}\beta + n_{ij}\gamma + \epsilon) \left(\alpha \frac{\sum_i n_{i.}n_{ij}}{n_{.j}} + \beta n_{.j} + \gamma \frac{\sum_i n_{ij}^2}{n_{.j}} + \epsilon\right)$$

cov (T_{ab}, T_f)

$$= 2 \sum_i \sum_j n_{ij}(n_{i.}\alpha + n_{.j}\beta + n_{ij}\gamma + \epsilon)^2/N$$

The expressions in the above table are those of the elements of the matrix var $(\mathbf{tt}')$ of Eq. (4). These elements are quadratic functions of the variance components α, β, γ, and ϵ, with coefficients being sums of functions of the n_{ij}'s. The other terms in (4) are not such as would simplify var ($\mathbf{vv}'$) if the elements of var ($\mathbf{tt}'$) as now known were inserted into (4), and therefore, as in any numerical case after calculating the expressions in the table these steps will be quite straightforward, it seems convenient to leave the results in their present form.

6. Balanced Data. It is easily shown that the formulae developed in the last paragraph reduce to the well-known results for balanced data when all the n_{ij} are put equal to n. For example, consider the variance of S_a. From the Analysis of Variance table, the expected value of S_a is given by

$$E(S_a) = (a - 1)(bn\alpha + n\gamma + \epsilon).$$

Then

$$\text{var}(T_a) = 2[a(bn\alpha + n\beta + n\gamma + \epsilon)^2 + a(a - 1)n^2\beta^2],$$

$$\text{var}(T_f) = 2(bn\alpha + an\beta + n\gamma + \epsilon)^2,$$

$$\text{cov}(T_a, T_f) = 2(bn\alpha + an\beta + n\gamma + \epsilon)^2.$$

Hence,

$$\begin{aligned}\text{var}(S_a) &= \text{var}(T_a - T_f)\\ &= 2(a - 1)(bn\alpha + n\gamma + \epsilon)^2\\ &= 2[E(Sa)]^2/(a - 1)\end{aligned}$$

and with $M_a = S_a/(a - 1)$, this gives the familiar result for mean squares

$$\text{var}(M_a) = 2[E(M_a)]^2/(a - 1).$$

Results similar to this can be obtained for M_b and M_{ab}, the mean squares for B-effects and interaction.

7. Conclusion. Matrix methods have been developed for finding the sampling variances of estimates of components of variance. In earlier work (4) these were used for data in a 1-way classification, and this paper has extended them to data for a 2-way classification, with unequal numbers of observations in the subclasses. The estimates of the components of variance for main effects and interaction are expressed as linear functions of the corrected sums of squares and the estimate of the error variance component. By expressing the corrected sums of squares as functions of the uncorrected sums of squares, the variance-covariance matrix of the estimates of the components of variance has been expressed as a function of that for the uncorrected sums of squares, (Eq. 4). Expressions have then been found for the elements of this, the variance-covariance matrix of the uncorrected sums of squares. It has been checked that when the data are assumed balanced, i.e., all n_{ij} equal to n, these expressions reduce to the appropriate forms for variances of mean squares then having independent χ^2-distributions. Estimates with any optimum properties have not been obtained, and it would seem that the only feasible estimation procedure in a practical case would be that of replacing the variance components in these formulae by their estimates.

It is hoped that these methods can next be extended to data in a 3-way classification with unequal subclass numbers, still based on Eisenhart's Model II and using Henderson's Method I for estimation.

REFERENCES

[1] S. L. Crump, "Present status of variance components analysis," *Biometrics*, Vol. 7 (1951), pp. 1–16.

[2] C. Eisenhart, "Assumptions underlying analysis of variance," *Biometrics*, Vol. 3 (1947), pp. 1–21.

[3] C. R. Henderson, "Estimation of variance and covariance components," *Biometrics*, Vol. 9 (1953), pp. 226–252.

[4] S. R. Searle, "Matrix methods in components of variance and covariance analysis," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 737–748.