ON MODELS AND COMPONENTS OF VARIANCE

by

Basilio A. Rojas
School of Statistics, Department of Applied Statistics

Technical Report No. 170

February 1972

University of Minnesota
Minneapolis, Minnesota

SUMMARY

Set concepts are introduced to derive general expressions for an experimental linear model and their related effects, sums of squares, components of variance, etc. Procedures are also described to systematize derivation of the coefficients, called K's, involved in the expectations, variances and covariances of both the uncorrected sums of squares, U_z's, and of linear functions of these U_z's. The use of set ideas proved useful to arrive at simple expression for some K's in particular cases. The random model is only discussed.

1. Introduction.

Components of variance have been discussed profusely in the literature evidencing both the theoretical implications and practical uses to which they lead to. Searle [1971] gives an excellent presentation of our status of knowledge in this matter and his paper's references well exceed a hundred in number.

Elementary algebra of sets are here introduced to represent a linear experimental model in a concise and general manner, which covers the nested, crossed classification and combinations of these two situations in an unbalanced case. Furthermore, the sums of squares and sums of cross products, and other features arising in an analysis of variance attain generalized expressions, from which interesting particular cases can be discovered without much difficulty.

The attainment of algebraic formulas for the general case of the variances of components of variance have been proved practically intractable. Searle [1971] gives reference to those formulas available in the literature for particular experimental situations. Hartley [1967] deviated from previous efforts to search for mathematical formulas and devised his "synthesis" method primarily geared to obtain numerical values of the coefficients found in the expectations, variances and covariances of quadratic forms. Rao [1968] extended the synthesis procedure. However, Hartley's method becomes prohibitive as the number of levels and factors increase.

This paper introduces a method to obtain the expectations, variances and covariances of the uncorrected sums of squares found in the appropriate analysis of variance of a random model in consideration. The writer proceeds to deal with the expectations, variances and covariances of linear functions of the uncorrected sums of squares. The method proposed involves the construction of matrices called $N(w,s|z)$ and the Hadamard products between pairs of those matrices. This procedure can be generalized to the mixed model without difficulty, but estimation of the components of variance is not as simply attained as in the random model case.

2. Notation and Concepts.

The purpose of this Section is to describe briefly the mathematical concepts and notation used here which are not common in the statistical literature on linear models.

2.1 Sets.

A set is a collection of objects having some common characteristic. Those objects are called elements. Sets are generally designated by capital letters and their elements by small letters. If the element x belongs to the set A we write $x \in A$. When x is not an element of A then $x \notin A$. The elements of a set are written within braces, for example, $A = \{-1, 0, 2\}$ means that the set A is composed of the elements

-1, 0 and 2. An empty or null set is represented by ϕ, and it is a set with no elements. Two sets A and B are equal, A = B, if and only if every element of A is an element of B and every element of B is an element of A. The set A is a subset of B, written $A \subset B$, if every element of A is also an element of B.

The union of sets A and B, denoted $A \cup B$ is a new set whose elements x are such that x is in A or x in B or in both. This statement is written as follows:

$$A \cup B = \{x | x \in A \text{ and/or } x \in B\}.$$

The intersection of sets A and B, written $A \cap B$ is the new set of elements common to both A and B:

$$A \cap B = \{x | x \in A \text{ and } x \in B\}.$$

A and B are called disjoint if they do not have an element in common, that is $A \cap B = \phi$.

The difference of sets A and B or relative complement of B in A is also written B' and is defined as follows:

$$A - B = B' = \{x | x \in A \text{ but } x \notin B\} .$$

From definition $B \cap B' = \phi$ and $B \cup B' = A$.

If A is a finite set of m elements then the new set whose elements are all the subsets of A is known as the power set of A, and denoted by P(A). The set P(A) has 2^m elements. For example, if A = {a, b, c} then P(A) has $2^3 = 8$ elements, and P(A) = {ϕ, {a}, {b}, {c}, {a,b}, {a,c}, {b,c}, A}.

2.2 Hadamard product.

If $A = (a_{ij})$ and $B = (b_{ij})$ are each matrices of the same order $n \times p$, then their Hadamard product, written $A*B$, is a new matrix $C = (c_{ij})$ of the same order $n \times p$ and such that $c_{ij} = a_{ij} \times b_{ij}$. In other words the Hadamard product of two matrices A and B is only defined when both matrices are of the same order and the new product elements are obtained by simple elementwise multiplication. Several properties of this product arise from its definition:

i) Multiplication is commutative:

$$A*B = B*A.$$

ii) If a and b are scalars

$$(aA)*(bB) = ab(A*B).$$

iii) If A is square, I the identity, then

$$A*I = D(a_{ii}),$$

where $D(a_{ii})$ denotes a diagonal matrix whose elements are a_{ii}.

iv) The distributive property holds

$$(A+B)*C = A*C + B*C .$$

v) More generally if there are two linear functions of matrices

$$L_1 = \Sigma a_i A_i \quad \text{and} \quad L_2 = \Sigma b_j B_j$$

where all the A_i's and B_j's are matrices of the same order $n \times p$, and the a_i's and b_j's are scalars, then L_1 and L_2 are also matrices both of order $n \times p$, and

$$L_1 * L_2 = \sum_{i,j} a_i b_j A_i * B_j . \qquad (1)$$

Uses of the Hadamard product in statistics is reviewed by Styan [1968]. He also discusses its application in multivariate analysis. He does not refer to property (v) above and deals mostly with properties of rank and positive definiteness of Hadamard products.

We also denote by $J_{n \times p}$ a matrix of order $n \times p$ whose elements are all ones. 1_p represents a row vector of p elements, every one equal to one. If $A = (a_{ij})$ is a matrix $n \times p$ the summation of all its elements will be written ΣA, that is

$$\Sigma A = \sum_{i,j} a_{ij} = tr(AJ_{p \times n}) . \qquad (2)$$

3. <u>The Model</u>.

Let $F = \{\alpha, \beta, \ldots, \theta\}$ be the factor set with the m elements $\alpha, \beta, \ldots, \theta$. Let $I = \{i, j, \ldots, q\}$ be the index set of m subscripts $i, j, \ldots, q$. In order to have a unique correspondence between the elements of F and those of I a one-to-one function $f: F \to I$ is defined such that each element of F has a distinct image in I, as follows,

$$f(\alpha) = i,\ f(\beta) = j, \ldots,\ f(\theta) = q.$$

Let us now define the power set of F as follows,

$$P(F) = \{F(w) | w = 1,2,\ldots, 2^m\}.$$

Therefore $F(w)$ is a subset of F and these subsets are numbered in a certain fashion. Consider the same function f defined above but extended as to be an associated set function, that is, $f[F(w)] = \{f(x) | x \in F(w)\} = I(w)$. This means that to each subset of factors $F(w)$ there corresponds a unique subset of subscripts $I(w)$. There will be 2^m subsets $F(w)$ and 2^m subsets $I(w)$, among which the null subset ϕ and the complete sets F and I are included in $P(F)$ and $P(I)$ respectively.

We assume that factor α is tested at $L(\alpha)$ levels, that is, $i = 1,2,\ldots, L(\alpha)$; that factor β is tested at $L(\beta)$ levels, or $j = 1,2,\ldots, L(\beta)$.

In general, the subscript subset $I(w)$ will go from $1,2,\ldots$ up to $L(w)$. This means that $L(w)$ represents the summation of subscripts contained in $I(w)$. L then represents the summation over all subscripts in the set I. L is then equal to the total number of factor combinations.

$Y_{ij\ldots q,r}$ is the observation corresponding to the combination of the m factors $\alpha, \beta, \ldots, \theta$ at levels $i, j, \ldots, q$ respectively on experimental unit $ij\ldots qr$. There are $n_{ij\ldots q}$ replications for the testing of combination of the m said factors at levels $i, j, \ldots, q$. We may write $Y_{ij\ldots q,r} = Y_{I,r}$ and $n_{ij\ldots q} = n_I$ where I is the index set referred to above. However, to identify more clearly particular combinations of levels of the factors we will write $Y_{(I,\lambda),r}$ as the observation when the subscripts in I attain the combination λ of levels. In similar manner, $n_{(I,\lambda)}$ is the number of observations for the combination λ of levels in I.

With the above preambles we can write an experimental linear model as follows:

$$Y_{(I,\lambda),r} = \sum_{w=1}^{f} F(w)_{I(w,\lambda)} + \epsilon_{(I,\lambda),r} \; . \qquad (3)$$

$F(w)_{I(w,\lambda)}$ is the effect, main effect or interaction, of factor(s) involved in subset $F(w)$ at levels λ of subscripts in $I(w)$. The summation sign extends over all or some of the possible 2^m subsets $F(w)$, depending upon the assumptions made by the experimenter. In general, $w = 1,2,\ldots, f$, where $f \leq 2^m$, for some effects might be considered null. We also use the convention that for $w = 1$ the subsets $F(1) = I(1) = \phi$, the null subset, and $F(1)_{I(1,\lambda)} = \mu$, the overall mean in model (3). Finally, $\epsilon_{(I,\lambda),r}$ is the error term, whose characteristics will be discussed below.

4. Nested, Crossed and Combined Classifications.

Model (3) is a general expression that includes the nested and crossed classifications as well as combinations of those two types as particular cases. To deal with a given model it is only necessary to establish proper equivalence between the model effects and the required F(w)'s. Examples will be given to clarify this assertion.

4.1 Nested classification.

For the nested or hierarchical model, say with 3 stages:

$$Y_{ijkr} = \mu + A_i + B_{ij} + C_{ijk} + \epsilon_{ijkr}$$

we require a set F with 3 factors, $F = \{\alpha, \beta, \gamma\}$ and $I = \{i, j, k\}$. The sets F(w), their respective I(w) and their equivalent correspondence to the above model are shown as follows:

w	1	2	3	4
F(w)	ϕ	α	α,β	α,β,γ
I(w)	ϕ	i	i,j	i,j,k
Equivalence	μ	A_i	B_{ij}	C_{ijk}

4.2 Crossed classification.

Let us suppose the model

$$Y_{ijk} = \mu + A_i + B_j + C_k + (AB)_{ij} + (AC)_{ik} + \epsilon_{ijk} \cdot$$

The sets F(w), I(w) and their equivalence will be

w	1	2	3	4	5	6
F(w)	ϕ	α	β	γ	α,β	α,γ
I(w)	ϕ	i	j	k	i,j	i,k
Equivalence	μ	A_i	B_j	C_k	AB_{ij}	AC_{ik}

The subsets (β, γ) and (α, β, γ) are not required in this case.

4.3 Nested and crossed classification.

Let us consider the following model

$$Y_{ijkr} = \mu + A_i + B_{ij} + C_k + (AC)_{ik} + (BC)_{ijk} + \epsilon_{ijkr} .$$

Here we require an $F = (\alpha, \beta, \gamma)$ and $I = \{i, j, k\}$. The subsets $F(w)$, $I(w)$ and their equivalence to the above model are as follows:

w	1	2	3	4	5	6
F(w)	ϕ	α	α,β	γ	α,γ	α,β,γ
I(w)	ϕ	i	i,j	k	i,k	i,j,k
Equivalence	μ	A_i	B_{ij}	C_k	AC_{ik}	BC_{ijk}

It can be seen that we have eliminated the effect β_j, used the interaction $(\alpha\beta)_{ij}$ as the nested effect B_{ij}, and the interaction $(\alpha\beta\gamma)_{ijk}$ as the effect $(BC)_{ijk}$.

5. Sums of Squares.

In order to obtain a general expression for a sum of squares in an analysis of variance it is convenient to designate $I'(w)$ as the subset in I which is the complement to the subset $I(w)$. We can then write

$$I(w) \cap I'(w) = \phi \qquad\qquad I(w) \cup I'(w) = I .$$

The subscript set I at levels λ will be written

$$(I, \lambda) = I(w, \lambda_1), I'(w,\lambda_2)$$

which means that combination λ of levels in I is equivalent to combination λ_1 of levels in $I(w)$ and levels λ_2 in $I'(w)$.

Similarly we define $L'(w)$ as the summation of all subscripts contained in $I'(w)$, or

$$L'(w) = \sum_{\lambda_2} I'(w, \lambda_2) \quad \text{and} \quad L(w) + L'(w) = L .$$

A general expression for U_z which denotes the uncorrected sum of squares corresponding to the source of variation of component $F(z)$, where z is any of the possible w values, $z = 1,2,\ldots, f$ will be now obtained. We can write

$$Y_{(I,\lambda),r} = Y_{I(z,\lambda_1),I'(z,\lambda_2),r} \quad \text{and} \quad n_{(I,\lambda)} = n_{I(z,\lambda_1),I'(z,\lambda_2)}$$

and we denote, following general nomenclature,

$$Y_{(I,\lambda)\bullet} = \sum_r Y_{(I,\lambda),r} = Y_{I(z,\lambda_1),I'(z,\lambda_2)\bullet}$$

$$Y_{I(z,\lambda_1)\bullet\bullet} = \sum_{\lambda_2 \epsilon I'(z)} Y_{I(z,\lambda_2),I'(z,\lambda_2)\bullet} \; ,$$

$$n_{I(z,\lambda_1)\bullet} = \sum_{\lambda_2 \epsilon I'(z)} n_{I(z,\lambda_1),I'(z,\lambda_1)}$$

where the last two summations extends over all $L'(z)$ combinations λ_2 of levels in subscripts contained in $I'(z)$. Also

$$Y_{\bullet\bullet\bullet} = \sum_{\lambda_1 \epsilon I(z)} Y_{I(z,\lambda_1)\bullet\bullet} \; , \; n = \sum_{\lambda_1 \epsilon I(z)} n_{I(z,\lambda_1)\bullet}$$

so $Y_{\bullet\bullet\bullet}$ is the grand total over all observations and n is the total number of observations.

The uncorrected sums of squares for $F(z)$, denoted by U_z, is

$$U_z = \sum_{\lambda_1 \epsilon I(z)} \frac{Y^2_{I(z,\lambda_1)\bullet\bullet}}{n_{I(z,\lambda_1)\bullet}} \tag{4}$$

valid expressions for $z = 1,2,\ldots, f$. In particular, for $z = 1$, $I(z) = \phi$ and $I'(z) = I$, U_1 corresponds to the correction factor:

$$U_1 = Y^2_{\bullet\bullet\bullet} / n \quad .$$

We also call U_t the uncorrected sum of squares for total:

$$U_t = \sum_{\lambda, r} Y^2_{(I,\lambda),r} . \tag{5}$$

The corrected sum of squares for all sources $F(z)$, $z = 2,3,\ldots, f$, total and error are linear functions, of the (f+1) expressions (4) and (5), which will depend on model (3).

6. Expectations of Sums of Squares.

6.1 The random model.

Model (1) is random when all effects $F(w)_{I(w,\lambda)}$, for $w = 2,3,\ldots, f$, and $\epsilon_{(I,\lambda),r}$ are random variables. Furthermore, we assume that these f random variables are uncorrelated within themselves and that any of the $\binom{f}{2}$ pairs of different random variables are uncorrelated for any coordination of subscripts. These random variables have means equal to zero and their variances are

$$\left.\begin{array}{l} V[F(w)_{I(w,\lambda)}] = \sigma^2_{F(w)} , \quad \text{for all } \lambda \\ V[\epsilon_{(I,\lambda),r}] = \sigma^2 , \text{ for all } \lambda \text{ and } r \end{array}\right\} \tag{6}$$

It is convenient to adopt the convention

$$V[F(1)_{I(1)}] = \sigma^2_{F(1)} = \sigma^2_{\phi} = \mu^2 . \tag{7}$$

We will now discuss the estimation of the components of variance (6) first and then the variances and covariances of such estimates.

6.2 Expectations of uncorrected sums of squares, U_z's.

Define $n(w, \mathfrak{L}|z, \lambda)$ as the number of $F(w)_{I(w,\mathfrak{L})}$ contained in $Y_{I(z,\lambda)\cdot\cdot}$. By $F(w)_{I(w,\mathfrak{L})}$ we mean the effect $F(w)$ when its respective subscript index $I(w)$ takes up the particular combination of levels $\mathfrak{L}$. The possible values of $\mathfrak{L}$ are $1,2,\ldots, L(w)$. For a given $I(z, \lambda)$ the number $n(w,\mathcal{L}|z,\lambda)$ can be zero for some $\mathcal{L}$. It is not difficult to see that

$$n(w,\mathcal{L}|z,\lambda) = n_{I(z,\lambda)\cup I(w,\mathcal{L})\bullet} \tag{8}$$

where the subscript $I(z,\lambda)\cup I(w,\mathcal{L})$ means the union of the subsets $I(z)$ and $I(w)$ when the respective subscripts take the values λ and $\mathcal{L}$ respectively. The dot means summation over the subset complement to $I(z,\lambda)\cup I(w,\mathcal{L})$.

$n(w,\mathcal{L}|z,\lambda)$ is also equivalent to the coefficient of $F(w)_{I(w,\mathcal{L})}$ in the expectation of $Y_{I(z,\lambda)\bullet\bullet}$ if the effects $F(w)_{I(w,\mathcal{L})}$ are, for only such purpose, assumed as fixed. This property can be helpful to obtain the $n(w,\mathcal{L}|z,\lambda)$.

If $K[\sigma^2_w|E(U_z)]$ stands for the coefficient of the component of variance $\sigma^2_{F(w)}$ in the expectation of U_z, it can be seen that

$$K[\sigma^2_w|E(U_z)] = \sum_{\lambda\epsilon I(z)} \frac{\sum_{\mathcal{L}} n^2(w,\mathcal{L}|z,\lambda)}{n_{I(z,\lambda)\bullet}} \tag{9}$$

expression valid for all z and w for the random model (3). The inside summation is over the levels $\mathcal{L} = 1,2,\ldots, L(w)$ of the subscripts in $I(w)$ for each of the levels λ in $I(z)$.

The coefficients $K(\sigma^2_w|E(U_z)]$ attain simpler expressions in particular circumstances. These are given in Table 1 and are easily obtained from the general expression (9). With the knowledge of all those coefficients the expectations of the uncorrected sums of squares for all sources can be written as follows:

$$\left.\begin{aligned} E(U_z) &= L(z)\sigma^2 + \sum_{w=1}^{f} K[\sigma^2_w|E(U_z)]\sigma^2_{F(w)} \\ &\qquad \text{for } z = 1,2,\ldots, f \\ E(U_t) &= n[\sigma^2 + \sum_{w=1}^{f} \sigma^2_{F(w)}] \end{aligned}\right\} \tag{10}$$

6.3 Estimation of components of variance.

The usual procedure is to calculate the corrected sums of squares for each of the f sources of variation $F(z)$, $z = 2,3,\ldots, f$ and error, and equalize these sums of squares with their respective expectations. This procedure furnishes a system of f simultaneous equations with the f components of variance as unknowns. The solution of that system of equations is simple for the balanced case. However, when there is unequality in the numbers n_I it can be worthwhile to proceed directly from the following system of $f + 1$ simultaneous equations, obtained from (10),

$$\left.\begin{array}{l} L(z)\hat{\sigma}^2 + \sum_{w=1}^{f} K(\sigma_w^2|E(U_z)]\hat{\sigma}^2_{F(w)} = U_z \\ \qquad\qquad z = 1,2,\ldots, f \\ n(\hat{\sigma}^2 + \sum_{w=1}^{f} \hat{\sigma}^2_{F(w)}) = U_t \end{array}\right\} \qquad (11)$$

The solution from the system with corrected sums of squares and the solution from system (10) will be identical.

The corrected sums of squares are linear functions of the U_z's and U_T and their expectations can be obtained from the same linear functions of expressions (10). Another procedure will be shown in Section 8.3.

7. Components of Covariance.

Let us consider the same experimental situation described in Section 3 with the factor set $F = \{\alpha,\beta,\ldots, \theta\}$ and the index set $I = \{i,j,\ldots, q\}$ both with m elements. A one-to-one correspondence is established between the elements of I and those of F and the relevant subsets $F(w)$'s and their respective $I(w)$'s are similarly defined. However, there are now two characteristics measured on each experimental unit. Denote by

$Y_{1,(I,\lambda),r}$ and $Y_{2,(I,\lambda),r}$ the measurements for characteristics 1 and 2, respectively, both measured in the same experimental or observation unit specified by the subscript $(I,\lambda),r$. This subscript indicates that the respective measurement corresponds to the r-th replicate of combination λ of levels of the m factors $\alpha,\beta,\ldots,\theta$. Model (3) can be extended as follows:

$$Y_{1,(I,\lambda),r} = \sum_{w=1} F_1(w)_{I(w,\lambda)} + \epsilon_{1,(I,\lambda),r}$$

$$Y_{2,(I,\lambda),r} = \sum_{w=1} F_2(w)_{I(w,\lambda)} + \epsilon_{2,(I,\lambda),r} \quad .$$

$F_v(w)_{I(w,\lambda)}$, (where v equals 1 or 2), is the effect, main effect or interaction,of factor(s) involved in set $F(w)$ for characteristic v, at levels of subscripts in $I(w)$. The summation sign in the above expression extends over the relevant subsets $F(w)$ according to the experimenter's assumptions. In general $w = 1,2,\ldots, f \leq 2^m$. When $w = 1$ then we follow the rule $F(1) = I(1) = \phi$, the null set, and

$$F_1(1)_{I(1)} = \mu_1 \quad \text{and} \quad F_2(1)_{I(1)} = \mu_2$$

where μ_1 and μ_2 are the overall true means of measurements of characteristics 1 and 2, respectively.

There are $n_{(I,\lambda)}$ units for each combination (I,λ) of the m factors. We assume that the 2 models written above are random, that is, all of the terms on the right side, except μ_1 and μ_2 are random variables with means zero and variances

$$V[F_v(w)_{I(w,\lambda)}] = \sigma^2_{F_v(w)} \; ; \; V[\epsilon_{v,(I,\lambda),r}] = \sigma^2_v$$

where $v = 1$ or 2. We also assume that the only covariances, between pairs of those random variables, not zero, are the following,

$$Cov[F_1(w)_{I(w,\lambda)}, F_2(w)_{I(w,\lambda)}] = \sigma_{F_1(w)F_2(w)}$$

for $w = 2,3,\ldots, f$; and

$$Cov[\epsilon_{1,(I,\lambda),r}, \epsilon_{2,(I,\lambda),r}] = \sigma_{\epsilon_1\epsilon_2} \quad .$$

It is to be noticed that the covariances written above are only defined for exactly the same combination of subscripts corresponding to $F(w)$'s, and ϵ's.

The uncorrected sum of cross products for source of variation $F(z)$, called UC_z, can be written

$$UC_z = \sum_{\lambda} \frac{(Y_{1,I(z,\lambda)\cdot})(Y_{2,I(z,\lambda)\cdot})}{n_{I(z,\lambda)\cdot}} \quad .$$

If $KC(w|z)$ denotes the coefficient of the component of covariance $\sigma_{F_1(w)F_2(w)}$ in the expectation of UC_z, it can be shown that

$$KC(w|z) = \sum_{\lambda} [\sum_{\mathfrak{L}} n^2_{I(w,\mathfrak{L}) \cup I(z,\lambda)\cdot}] / n_{I(z,\lambda)\cdot}$$

which is identical to $K(\sigma^2_w|z)$ given in (9). In other words, the coefficient of the component of covariance $\sigma_{F_1(w)F_2(w)}$ in UC_z is the same as of component of variance $\sigma^2_{F(w)}$ in U_z. The expectations of the UC_z's and UC_t (for total) are

$$E(UC_z) = L(z)\sigma_{\epsilon,\epsilon_2} + \sum_{w=1} KC(w|z)\sigma_{F_1(w)F_2(w)}$$

$$E(UC_T) = n[\sigma_{\epsilon,\epsilon_2} + \sum_{w=1} \sigma_{F_1(w)F_2(w)}] \quad .$$

8. Variances and Covariances of Components of Variance.

8.1 Means, variances and covariances of sums of squares.

Here we present a procedure to obtain the expectations, variances and covariances of sums of squares in an analysis of variance for the random model (3). Section 6 dealt only with the expectations of the sums of squares, which will be obtained here also but as a step to arrive at the expressions for the variances and covariances. The method is justified itself for the procedures follow exactly the definition of a summation $Y_{I(z,\lambda)\cdot}/n_{I(z,\lambda)\cdot}$ but taking into account the obvious consequences of these successive operations on each of the effects $F(w)$ contained in model (3). This reason made the writer eliminate unnecessary justification in the assertions that will follow.

i) If $E(Y_{I(z,\lambda)\cdot\cdot}|F(w)_{I(w)})$ stands for the expectation of $Y_{I(z,\lambda)\cdot\cdot}$ when the effects $F(w)_{I(w,\mathcal{L})}$, for all $\mathcal{L} = 1,2,\ldots, L(w)$, are fixed and all others are random variables with mean zero, then

$$E(Y_{I(z,\lambda)\cdot\cdot}|F(w)_{I(w)}) = \sum_{\mathcal{L}} n(w,\mathcal{L}|z,\lambda)F(w)_{I(w,\mathcal{L})} \cdot$$

Call $N'(w|z,\lambda)$ the row vector with $L(w)$ elements, its $\mathcal{L}$-th element being $n(w,\mathcal{L}|z,\lambda)$. The ordering of elements must be kept throughout. In similar manner we define the row vector $N'(s|z,\lambda)$ related to the effects $F(s)_{I(s)}$ in the expectation of the same summation $Y_{I(z,\lambda)\cdot\cdot}$.

ii) Obtain the matrix

$$N(w,s|z,\lambda) = [N(w|z,\lambda)N'(s|z,\lambda)]/n_{I(z,\lambda)\cdot} \cdot$$

The column vector $N(w|z,\lambda)$ with $L(w)$ elements is multiplied by the row vector $N'(s|z,\lambda)$ with $L(s)$ elements giving a matrix of order $L(w) \times L(s)$, which is then divided by the scalar $n_{I(z,\lambda)\cdot}$.

iii) Obtain the matrix

$$N(w,s|z) = \sum_{\lambda} N(w,s|z,\lambda) \tag{13}$$

a matrix of order $L(w) \times L(s)$. Matrix $N(w,s|z)$ contains the numbers of each of the terms $F(w)_{I(w,\mathcal{L})}F(s)_{I(s,m)}$ in U_z, the uncorrected sum of squares for component $F(z)$.

Matrix (13) is defined for all values of $w,s = 2,3,\ldots, f$ for effects; w and s will be also referred to the error term, a total of f values. Matrix $N(w,s|z)$ is also valid for all U_z where $z = 1,2,\ldots, f$ for components $F(z)$ and also valid for $z = t$ or U_t, uncorrected sum of squares for total. There exists then $f(f+1)^2/2$ matrices $N(w,s|z)$. However, many of them attain simple forms easily to write down without the necessity of going through steps (i) to (iii).

iv) It can be seen that the coefficient of $\sigma^2_{F(w)}$ in the expectation of U_z is

$$K(\sigma^2_w|E(U_z)] = tr[N(w,w|z)] \ . \tag{14}$$

v) Consider the following two matrices of the form (13), $N(w,s|z)$ and $N(w,s|v)$. Both referred to the same terms $F(w) \times F(s)$ in U_z and U_v respectively. Obtain the Hadamard product

$$H(w,s|z,v) = N(w,s|z)*N(w,s|v) \ . \tag{15}$$

$H(w,s|z,v)$ is a matrix of order $L(w) \times L(s)$. Expression (15) is valid for $w,s = 2,3,\ldots, f$, and error (f values); $z,v = 1,2,\ldots, f$ and total, (f+1 values), giving a total of $f(f+1)^2(f+2)/4$ different matrices $N(w,s|z,v)$. Fortunately, many of them have simple forms which can be written directly without need of following previous steps. These cases are shown in Table 1.

vi) We make the assumption that all terms, except $\mu = F(1)_{I(1)}$, in model (3) are normally and independently distributed with means zero and variances as stated in (6). The normality assumption is made in order that the fourth moment of the $F(w)_{I(w)}$'s and $\epsilon_{(I,\lambda),r}$ attain the simple form of being equal to 3 times the square of the respective variance, reducing the derivations that follow.

vii) Let $K[\sigma_w^4|\text{cov}(U_z, U_v)]$ and $K[\sigma_w^2\sigma_s^2|\text{cov}(U_z, U_v)]$ be the coefficients of terms $\sigma_{F(w)}^4$ and $\sigma_{F(w)}^2\sigma_{F(s)}^2$ in the covariance between the uncorrected sums of squares U_z and U_v.

viii) We remind here notation introduced in (2) that if $A = (a_{ij})$ is a matrix then ΣA represents the summation of all elements in A, that is, $\Sigma A = \Sigma a_{ij}$.

ix) It can be shown that

$$\left.\begin{aligned} K[\sigma_w^2\sigma_s^2|\text{cov}(U_z, U_v)] &= 4\Sigma\ H(w,s|z,v) \\ \text{and} \qquad & \\ K[\sigma_w^4|\text{cov}(U_z, U_v)] &= 2\Sigma\ H(w,w|z,v) \end{aligned}\right\} \qquad (16)$$

Special cases of (16) are

$$\left.\begin{aligned} K[\sigma_w^2\sigma_s^2|V(U_z)] &= 4\Sigma\ H(w,s|z,z) \\ \text{and} \qquad & \\ K[\sigma_w^4|V(U_z)] &= 2\Sigma\ H(w,w|z,z) \end{aligned}\right\} \qquad (17)$$

where $V(U_z)$ means the variance of U_z.

Formulas (16) and (17) are valid for all $w,s = 2,3,\ldots, f$ and error; also for all $z,v = 1,2,\ldots, f$ and total (U_T).

x) The variance of an uncorrected sum of squares U_z or the covariance between U_z and U_v can be written as follows:

$$V(U_z) = \sum_{w,s} K[\sigma_w^2\sigma_s^2|V(U_z)]\sigma_{F(w)}^2\sigma_{F(s)}^2$$

$$\text{Cov}(U_v, U_z) = \sum_{w,s} K[\sigma_w^2\sigma_s^2|\text{cov}(U_z, U_v)]\sigma_{F(w)}^2\sigma_{F(s)}^2$$

where the summation is over all values $w,s = 2,3,\ldots, f$ and error.

The estimated variance of U_z, $\hat{V}(U_z)$ is obtained by substituting the true unknown components of variance with their respective estimated values. A similar statement applies to the estimated $\text{Cov}(U_z, U_v)$.

8.2 Simple expressions for some K's.

The coefficients of the terms σ^4 and $\sigma^2_{F(w)}\sigma^2$ in $V(U_z)$ and $Cov(U_z, U_v)$ can be obtained using the general method explained in Section 8.1 which leads to formulas (16) and (17). The terms σ^4 and $\sigma^2_{F(w)}\sigma^2$ involve the error terms $\epsilon_{(I,\lambda),r}$ of model (3), and for this reason the matrices required are of order $n \times n$ or $L(w) \times n$. Fortunately and due to the fact that to each observation there corresponds one single error the matrices $N(\epsilon,\epsilon|z,\lambda)$ and $N(w,\epsilon|z,\lambda)$ have elements, which are either one or zero, multiplied, such matrices, by the scalar $1/n_{I(z,\lambda)\cdot}$. This feature simplifies the derivation of the respective K's. It can be shown that

$$K[\sigma^4|cov(U_z,U_v)] = 2 \sum_{\lambda,\mathcal{L}} \frac{(n_{I(z,\lambda)\cup I(v,\mathcal{L})\cdot})^2}{(n_{I(z,\lambda)\cdot})(n_{I(v,\mathcal{L})\cdot})} . \tag{18}$$

For the application of the above formula we construct a two-way table with columns 1,2,..., L(z); and rows 1,2,..., L(v). The cell λ, $\mathcal{L}$ would contain the number $n_{I(z,\lambda)\cup I(v,\mathcal{L})\cdot}$ which is the common number of observations. The marginal table would contain the numbers $n_{I(z,\lambda)\cdot}$ and $n_{I(v,\mathcal{L})\cdot}$ for columns and rows respectively.

It is also not difficult to see that

$$K[\sigma^2_w\sigma^2|cov(U_z,U_v)] \tag{19}$$

$$= 4 \sum_{\lambda,\mathcal{L}} \frac{(n_{I(z,\lambda)\cup I(v,\mathcal{L})\cdot}) \sum_m (n_{I(w,m)\cup I(z,\lambda)\cdot})(n_{I(w,m)\cup I(v,\mathcal{L})\cdot})}{(n_{I(z,\lambda)\cdot})(n_{I(v,\mathcal{L})\cdot})} .$$

The calculation involves a two-way table with L(z) columns and L(v) rows which is equal to the one constructed for the solution of (18). But now in each cell the number indicated by the inside summation in (19) have to be indicated. The operations indicated by (19) can then follow to arrive at the value of the corresponding K.

A simplified version of (19) is obtained when the set $I(v)$ is equal to $I(w)$. The same result is obtained when $I(v) = I(z)$ in (19).

$$K[\sigma_w^2\sigma^2|V(U_z)] = K[\sigma_w^2\sigma^2|\mathrm{cov}(U_w,U_z)] = 4\sum_{\lambda} \frac{\sum_m (n_{I(w,m)\cup I(z,\lambda)\cdot})^2}{n_{I(z,\lambda)\cdot}} . \qquad (20)$$

More special cases of expressions (18) to (20) are given in Table 1. However, it must be emphasized that the procedures indicated in Section 8.1 can be developed in a routine and systematic manner, especially with the help of processing of data devices, which make unnecessary these formulas (18) to (20), and those other special cases shown in Table 1.

8.3 <u>Mean and variance of a linear function of U_z's.</u>

A corrected sum of squares in an analysis of variance as well as the estimator of a component of variance are linear functions of the U_z's, uncorrected sums of squares, defined by (4) and (5). Let us consider the linear function

$$L = \Sigma_{a_g} U_{z_g} . \qquad (21)$$

Our aim is to obtain $E(L)$ and $V(L)$, for which we find coefficients $K[\sigma_w^2|E(L)]$, $K[\sigma_w^4|V(L)]$ and $K[\sigma_w^2\sigma_s^2|V(L)]$, as follows:

i) For each U_{z_g} in (21) we obtain the matrices $N(w,w|z_g)$ and $N(w,s|z_g)$ for all $w,s = 2,\ldots, f$ and error ε.

ii) Construct the matrices

$$N(w,w|L) = \Sigma\, a_g N(w,w|z_g), \text{ and } N(w,s|L) = \Sigma\, a_g N(w,s|z_g). \qquad (22)$$

iii) Calculate the Hadamard products,

$$H(w,w|L,L) = N(w,w|L)*N(w,w|L) \text{ and } H(w,s|L,L) = N(w,s|L)*N(w,s|L). \quad (23)$$

iv) The coefficients K we are searching for are

$$\left.\begin{aligned} K[\sigma_w^2|E(L) &= tr[N(w,w|L)] \\ K[\sigma_w^4|V(L)] &= 2\Sigma H(w,w|L,L) \\ K[\sigma_w^2\sigma_s^2|V(L)] &= 4\Sigma H(w,s|L,L) \end{aligned}\right\} \quad (24)$$

v) Formulas (24) assume that we have actually obtained matrices $N(w,w|L)$, $H(w,w|L,L)$, and $H(w,s|L,L)$ as stated in steps (ii) and (iii) above. There is also the following way which may be easier to apply in certain cases. First, it is easy to see that

$$K[\sigma_w^2|E(L)] = \Sigma\, a_g\; tr[N(w,w|z_g)] \;. \quad (25)$$

Also, using (1) in (22)

$$H(w,s|L,L) = \Sigma\, a_g a_j\; N(w,s|z_g)*N(w,s|z_j) = \Sigma\, a_g a_j\; H(w,s|z_g,z_j) \;.$$

The summation extending over all g's and j's. Substituting this result in (24):

$$\left.\begin{aligned} K[\sigma_w^4|V(L)] &= 2 \sum_{g,j} a_g a_j [\Sigma H(w,w|z_g,z_j)] \\ K[\sigma_w^2\sigma_s^2|V(L) &= 4 \sum_{g,j} a_g a_j [\Sigma H(w,s|z_g,z_j)] \end{aligned}\right\} \quad (26)$$

where $\Sigma H(w,s|z_g,z_j)$, we emphasize it, represents the summation of all elements in matrix $H(w,s|z_g,z_j)$, following notation (2).

Formulas (25) and (26) are to be recommended when we are interested in the expectations and variances of a few linear functions.

We are now able to write down $E(L)$ and $V(L)$,

$$\left.\begin{aligned} E(L) &= n(\Sigma a_g)\mu^2 + \sum_{w=2} K[\sigma_w^2|E(L)]\sigma_{F(w)}^2 \\ V(L) &= \sum_{w,s=2} K[\sigma_w^2\sigma_s^2|V(L)]\sigma_{F(w)}^2\sigma_{F(s)}^2 \end{aligned}\right\} \qquad (27)$$

where the summation is over $w,s = 2,3,\ldots, f$ and error (ε).

8.4 <u>Variances and covariances of estimated components of variance.</u>

Expression (27) provides the variance of the estimator of any component of variance $\hat{\sigma}_{F(w)}^2$ as long as the corresponding linear function (21) is known. There is, however, a more general and compact presentation of the variance-covariance matrix of the vector of estimators of the components of variance.

Let us write down the system of $f + 1$ simultaneous equations (11) in matrix form

$$A\hat{\Omega} = U \qquad (28)$$

where $\hat{\Omega}$ is the column vector of unknowns with $f + 1$ elements: $\hat{\mu}^2$ and f components of variance $\hat{\sigma}_{F(w)}^2$ (including σ^2). A is the matrix of coefficients, square of order $f + 1$. U is the column vector of $f + 1$ U_z's. The solution of (28) is

$$\hat{\Omega} = A^{-1}U . \qquad (29)$$

If $V(\hat{\Omega})$ denotes the variance-covariance matrix of $\hat{\Omega}$ and $V(U)$ is the variance-covariance matrix of the U_z's, then

$$V(\hat{\Omega}) = A^{-1}V(U)(A^{-1})' . \qquad (30)$$

But $V(U)$ can be written as follows:

$$V(U) = \sum_{w,s=2} \sigma^2_{F(w)}\sigma^2_{F(s)} \mathcal{K}(w,s) \tag{31}$$

where $\sigma^2_{F(w)}\sigma^2_{F(s)}$ is a scalar and $\mathcal{K}(w,s)$ is a square matrix of order $f + 1$ whose elements are the coefficients $K[\sigma^2_w\sigma^2_s|\text{cov}(U_z,U_v)]$ and $K[\sigma^2_w\sigma^2_s|V(U_z)]$ for all $z,v = 1,2,\ldots, f$ and total, (U_T). Expression (31) indicates that the $V(U)$ is equal to the summation of $f(f+1)/2$ variance-covariance matrices $\mathcal{K}(w,s)$ for each one of the terms $\sigma^4_{F(w)}$, $w = 2,3,\ldots, f$, and error, and for each $\sigma^2_{F(w)}\sigma^2_{F(s)}$, $w \neq s = 2,3,\ldots, f$ and error. Substituting (31) in (30) we find

$$V(\hat{\Omega}) = \sum_{w,s} \sigma^2_{F(w)}\sigma^2_{F(s)}A^{-1}\mathcal{K}(w,s)(A^{-1})' \;. \tag{32}$$

Elements of matrices $\mathcal{K}(w,s)$ are obtained by the general expressions (16) and (17), by using the expressions (18) to (20) for special cases and by making reference to Table 1 for the simpler cases.

Expression (32) contains the variance of $\hat{\mu}^2$ as well as the covariance of $\hat{\mu}^2$ with all estimated components of variance. If this information about $\hat{\mu}^2$ is of no interest we then partition matrix A^{-1} and take the f rows corresponding to the f components of variance. Call B this partitioned submatrix of A^{-1}. B is rectangular of order $f \times (f+1)$. Denote by $\hat{\Omega}_1$ the partitioned column vector of f elements drawn from $\hat{\Omega}$, by eliminating $\hat{\mu}^2$. We can write

$$\hat{\Omega}_1 = BU, \text{ and } V(\hat{\Omega}_1) = \sum_{w,s} \sigma^2_{F(w)}\sigma^2_{F(s)}B\mathcal{K}(w,s)B' \;. \tag{33}$$

For the estimator of the variance-covariance matrix of $\hat{\Omega}_1$ we use

$$\hat{V}(\hat{\Omega}_1) = \sum \hat{\sigma}^2_{F(w)}\hat{\sigma}^2_{F(s)}B\mathcal{K}(w,s)B' \;. \tag{34}$$

TABLE 1: K COEFFICIENTS

Coefficient of	Coefficient in	Condition for Validity	Formula or Value
$\sigma^2_{F(w)}$	$E(U_z)$	any w and z	(9), (14)
	$E(U_z)$, $E(U_t)$	$I(w) \subset I(z)$	n
	E(C.F.)	any w	k_w
σ^2	$E(U_z)$	any z	L(z)
μ^2	$E(U_z)$	any z	n
$\sigma^4_{F(w)}$	$Cov(U_z,U_v)$	any w,z and v	(16)
	$V(U_z)$	any w and z	(17)
	$V(U_w)$	any w	$2k_w n$
	$Cov(U_w,U_z)$	$I(w) \subset I(z)$	$2k_w n$
	$Cov(U_w,U_t)$	any w	$2k_w n$
	$Cov(U_z,U_v)$	$I(w) \subset I(z) \subset I(v)$	$2k_w n$
$\sigma^2_w\sigma^2_s$	$Cov(U_z,U_v)$	any w, s, z, and v	(16)
	$V(U_z)$	any w,s and z	(17)
	$V(U_w)$ and $V(U_s)$	$I(w) \subset I(s)$	$4k_s n$
	$Cov(U_w,U_s)$	$I(w) \subset I(s)$	$4k_s n$
	$Cov(U_s,U_z)$	$I(w) \subset I(s) \subset I(z)$	$4k_s n$
	$Cov(U_s,U_t)$	$I(w) \subset I(s)$	$4k_s n$
	$Cov(U_w,U_t)$	$I(w) \subset I(s)$	$4k_s n$
$\sigma^2_{F(w)}\sigma^2$	$Cov(U_z,U_v)$	any w,z and v	(19)
	$Cov(U_z$,C.F.)	any w and z	$4k_w$
	$Cov(U_z,U_v)$	$I(w) \subset I(z) \subset I(v)$	4n
	$Cov(U_w,U_z)$	any w and z	(20)
	$Cov(U_w,U_z)$	$I(w) \subset I(z)$	4n
	$Cov(U_w,U_t)$	any w	4n
	$Cov(U_w$,C.F.)	any w	$4k_w$
	$V(U_z)$, $V(U_t)$	any w and z	4n
	V(C.F.)	any w	$4k_w$
σ^4	$Cov(U_z,U_v)$	any z and v	(16), (18)
	$Cov(U_z,U_v)$	$I(z) \subset I(v)$	2L(z)
	$Cov(U_z,U_t)$	any z	2L(z)
	$Cov(U_z$,C.F.)	any z	2
	$V(U_z)$	any z	2L(z)

w,s= 2,3,..., f ; z,v = 1,2,..., f and total (t); U_1 = C.F. , $k_w = (\sum_m n^2_{I(w,m)\cdot})/n$, $k_s = (\sum_m n^2_{I(s,m)\cdot})/n$.

9. Example.

A very simple case of a two-way crossed classification is presented to show the methodology developed in this paper. We assume the model

$$Y_{ijr} = \mu + \alpha_i + \beta_j + \epsilon_{ijr}$$

where the number of levels for i and j are equal to 2. The numbers n_{ij}, $n_{i.}$ and n_{ij} are given in Table 2. The sets $F = \{\alpha, \beta\}$ and $I(i,j)$ give rise to the subsets, relevant to the above model, shown in Table 3, which also includes the values $L(w)$.

TABLE 2: VALUES OF n_{ij}

	α_1	α_2	$n_{.j}$
β_1	2	4	6
β_2	3	5	8
$n_{i.}$	5	9	14 = n

TABLE 3: SET NOTATION

w	1	2	3
F(w)	ϕ	α	β
F(w)	ϕ	i	j
L(w)	1	2	2

Table 3 expresses the fact that F(1), F(2) and F(3) are the mean μ, and the effects α and β respectively. Also U_w, for $w = 1,2,3$ and t will be as follows. U_1 is the correction factor; U_2, U_3 and U_t are the uncorrected sums of squares for columns (α), rows (β), and for total, respectively.

We will show now how to obtain several $N(w,s|z)$ matrices. The matrices for $z = 2$ are $N(2,2|2)$, $N(3,3|2)$, $N(2,3|2)$, $N(2,\epsilon|2)$, $N(3,\epsilon|2)$, and $N(\epsilon,\epsilon|2)$. Matrices $N(2,2|2)$, $N(3,3|2)$ and $N(2,3|2)$, for example are those including forms $\alpha_i\alpha_i'$, $\beta_j\beta_j'$ and $\alpha_i\beta_j$, respectively in U_2 and these are obtained from the expectations of the $Y_{i..}$'s $(i = 1,2)$ assuming that effects α_i and β_j (in this case) are the only ones to be fixed, all others, including μ, have expectation zero:

$$E(Y_{1\cdot\cdot}|\alpha,\beta) = 5\alpha_1 + 0\alpha_2 + 2\beta_1 + 3\beta_2$$

$$E(Y_{2\cdot\cdot}|\alpha,\beta) = 0\alpha_1 + 9\alpha_2 + 4\beta_1 + 5\beta_2 \, .$$

To obtain all the 6 matrices $N(w,s|2)$ indicated above we would use $E(Y_{i\cdot\cdot}|\alpha,\beta,\epsilon)$ for all i's. Now according to expression (13) we can write

$$N(2,2|2) = \frac{(5,0)'(5,0)}{5} + \frac{(0,9)'(0,9)}{9} = \begin{pmatrix} 5 & 0 \\ 0 & 9 \end{pmatrix}$$

$$N(3,3|2) = \frac{(2,3)'(2,3)}{5} + \frac{(4,5)'(4,5)}{9} = \begin{pmatrix} 2.5778 & 3.4222 \\ 3.4222 & 4.5778 \end{pmatrix}$$

where $(5,0)'$ is the transpose of the row vector $(5,0)$. In the same manner

$$N(2,3|2) = \begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix}; \quad N(2,\epsilon|2) = \begin{pmatrix} 1_s & 0 \\ 0 & 1_9 \end{pmatrix}$$

$$N(3,\epsilon|2) = \begin{pmatrix} \frac{2}{5}\,1_5 & \frac{4}{9}\,1_9 \\ \frac{3}{5}\,1_5 & \frac{5}{9}\,1_9 \end{pmatrix}; \quad N(\epsilon,\epsilon|2) = \begin{pmatrix} \frac{1}{5}\,J_{5x5} & 0 \\ 0 & \frac{1}{9}\,J_{9x9} \end{pmatrix}$$

where 1_m is a row vector with all of its m elements being one. J_{mxp} is a matrix of order $m \times p$ with all of its elements being one.

The K coefficients for the expectation of U_2 and for the variance of U_2 can be calculated by simply using formulas (14) and (17). Similar matrices are calculated for each of the U_z's.

Table 4 includes the K coefficients for the expectations of the 4 U_z's. Table 5 gives the K's for the variance of U_2 and for the covariances of U_2 with the other U_z's. These K's are obtained from (14), (16) and (17).

TABLE 4: COEFFICIENTS $K(\sigma_w^2|E(U_z)$

U_z	Component		
	σ^2	σ_2^2	σ_3^2
U_1	1	7.57	7.14
U_2	2	14	7.15
U_3	2	7.58	14
U_t	14	14	14

TABLE 5: COEFFICIENTS $K[\sigma_w^2\sigma_s^2|Cov(U_2U_z)]$

Component	U_1	U_2	U_3	U_t
σ_2^4	122	212	122.17	212
σ_3^4	102.04	102.04	104.18	104.18
σ^4	2	4	2.00	4
$\sigma_2^2\sigma_3^2$	216	216	216	216
$\sigma_2^2\sigma^2$	30.28	56	30.33	56
$\sigma_3^2\sigma^2$	28.57	28.62	28.62	28.62

10. References.

Hartley, H. O. [1967]. Expectation, variances and covariances of ANOVA mean squares by 'synthesis.' Biometrics 23 105-14.

Rao, J.N.K. [1968]. On expectations, variances and covariances of ANOVA mean squares by 'synthesis.' Biometrics 24 963-78.

Searle, S. R. [1971]. Topics in variance component estimation. Biometrics 27 1-76.

Styan, G. P. H. [1968]. Inference in multivariate normal populations with structure. Part I: Inference on variances when correlations are known. University of Minnesota, Department of Statistics Technical Report No. 111, Minneapolis, 81pp.