DERIVATION OF EXPECTED MEAN SQUARES IN VARIANCE COMPONENTS

MODELS HAVING FINITE POPULATIONS

S. R. Searle* and R. F. Fawcett

Texas A&M University

Summary

Expected mean squares in variance components models having populations of finite size are shown related to their corresponding values with infinite populations. The relationship applies uniformly to both nested and crossed classifications, and it is the same for both unbalanced and balanced data.

Introduction

In variance components models the random effects are usually assumed to have infinite populations. Finite populations have also been considered: for example, Bennett and Franklin (1954), Cornfield and Tukey (1956) and Wilk and Kempthorne (1956) discuss them for various cases of balanced data. However, for unbalanced data, apart from Tukey's (1956 and 1957) treatment of the 1-way classification, the only discussion of finite population models appears to be that of Gaylor and Hartwell (1969), who deal in detail with only the 3-way nested classification. In estimating the variance components of the model by the familiar method of equating observed mean squares to expected values, somewhat tedious algebra is involved in deriving the expected mean squares even though the ultimate results amount to making only minor adjustments to the expectations under infinite populations. In place of this algebra we develop a rule for converting expectations under infinite models into expectations under finite models. It applies universally to balanced and unbalanced data, and to nested and crossed classifications as well as to mixtures of them. Since expectations under infinite populations are already available in the literature, the rule provides the change from infinite to finite population models.

---

* On leave from Cornell University, 1968-1969.

The model is assumed to have factors A, B, ..., K, with the population of effects in the θ-factor having size N_θ, its effects being θ_i for i = 1,2, ..., N_θ, with θ having the values A, B, ..., K. As is customary [e.g. Gaylor and Hartwell (1969)], it is assumed that the mean of each population is zero, so that

$$\sum_{i=1}^{N_\theta} \theta_i = 0 \; ; \tag{1}$$

its variance is correspondingly defined as

$$\sigma_\theta^2 = \frac{\sum_{i=1}^{N_\theta} \theta_i^2}{N_\theta - 1} \, . \tag{2}$$

A consequence of (1) is that

$$\left(\sum_{i=1}^{N_\theta} \theta_i\right)^2 = \sum_{i=1}^{N_\theta} \theta_i^2 + \sum_{i \neq}^{N_\theta} \sum_{i'}^{N_\theta} \theta_i \theta_{i'} = 0 \tag{3}$$

and so from (2) and (3)

$$\sum_{i \neq}^{N_\theta} \sum_{i'}^{N_\theta} \theta_i \theta_{i'} = -(N_\theta - 1)\sigma_\theta^2 \, . \tag{4}$$

In the sampling that has taken place in obtaining data it is assumed that the finite populations of effects (and errors) in the model have been sampled at random without replacement. If θ_r is a sampled value of the θ-effects then, because of (1),

$$\text{mean } (\theta_r) = E(\theta_r) = \frac{1}{N_\theta} \sum_{i=1}^{N_\theta} \theta_i = 0 \; ; \tag{5}$$

and by (2) and (4)

$$\text{var}\,(\theta_r) = E(\theta_r^2) = \frac{1}{N_\theta}\sum_{i=1}^{N_\theta} \theta_i^2 = (1 - N_\theta^{-1})\sigma_\theta^2 \;; \tag{6}$$

and similarly, for two sample values θ_r and θ_s

$$\text{cov}\,(\theta_r\theta_s) = E(\theta_r\theta_s) = \frac{1}{N_\theta(N_\theta - 1)} \sum_{i}^{N_\theta}\sum_{\neq i'}^{N_\theta} \theta_i\theta_{i'} = \frac{-\sigma_\theta^2}{N_\theta} \,. \tag{7}$$

Since the corresponding values of (6) and (7) for infinite populations are σ_θ^2 and 0, respectively, the expected values of mean squares in finite population models are not the same as with infinite populations. In either case the expected values are linear functions of the variance components; the coefficients of these components are determined for finite population models in accord with (6) and (7).

Suppose $\underline{y}$ represents the vector of observations, having vector of means $\underline{\mu}$ and covariance matrix $\underline{V}$. Then any mean square is a quadratic form in $\underline{y}$, $\underline{y}'\underline{Q}\underline{y}$ say, and its expected value is

$$E(\underline{y}'\underline{Q}\underline{y}) = tr(\underline{Q}\underline{V}) + \underline{\mu}'\underline{Q}\underline{\mu} \,. \tag{8}$$

Now analysis of variance mean squares with balanced data, and their analogues with unbalanced data, are quadratic forms of the observations such that row rums of $\underline{Q}$ are zero; i.e. $\underline{Q}\underline{1} = 0$, where $\underline{1}$ is a vector of 1's. Furthermore, in random models $\underline{\mu} = \mu'\underline{1}$ and so in (8), $\underline{\mu}'\underline{Q}\underline{\mu} = \underline{\mu}'\underline{Q}\underline{1}\mu = 0$. Therefore, in considering only expressions of this nature, denoted by M, (8) can be taken as

$$E(M) = E(\underline{y}'\underline{Q}\underline{y}) = tr(\underline{Q}\underline{V}) \,.$$

This result is independent of whether the model has finite or infinite populations. It therefore holds for both and we write

$$E_\infty(M) = tr(\underline{Q}\underline{V}_\infty) \quad \text{and} \quad E_F(M) = tr(\underline{Q}\underline{V}_F) \tag{9}$$

for infinite and finite population models, respectively. Since $\underline{Q}$ is the same in both cases the only differences between the two expected mean squares are those brought about by using $\underline{V}_F$ in place of $\underline{V}_\infty$. The exact nature of these differences can therefore be ascertained by looking at the manner in which $\underline{V}_\infty$ gets altered to become $\underline{V}_F$ when changing from infinite to finite populations. The alterations depend on results like (6) and (7). Whatever they are, it is immaterial to (9) whether $\underline{Q}$ has come from balanced or unbalanced data. In either case (9) holds, and so the changes to be made to derive $\underline{V}_F$ from $\underline{V}_\infty$ will yield $E_F(M)$ from $E_\infty(M)$ whether the data be balanced or unbalanced. The discussion and results that follow therefore apply equally to both kinds of data.

Nested Models

It is opportune to developing the general result to consider three separate cases: nested models (consisting solely of nested classifications), crossed-classification models (having no nested classifications) and models involving combinations of nested and crossed classifications. To discuss nested models we first consider a 3-way nested model of factors A, B within A, and C within B. The derivation of $\underline{V}_F$ from $\underline{V}_\infty$ is then a matter of considering the effect of (6) and (7) on the various elements of $\underline{V}_\infty$. First, the

$$\text{diagonal elements in } \underline{V}_\infty \text{ are } \sigma_A^2 + \sigma_B^2 + \sigma_C^2 + \sigma_e^2 \tag{10}$$

and by (6) they become

$$\text{diagonal elements in } \underline{V}_F\text{:} \quad (1 - N_A^{-1})\sigma_A^2 + (1 - N_B^{-1})\sigma_B^2 + (1 - N_C^{-1})\sigma_C^2 + (1 - N_e^{-1})\sigma_e^2 , \tag{11}$$

where N_e is the size of the population of error terms. Another term is

$$\sigma_A^2 + \sigma_B^2 + \sigma_C^2 \quad \text{in} \quad \underline{V}_\infty ; \tag{12}$$

For finite population models the σ^2's of (12) will be changed just as they were in (10); but since in $\underline{V}_\infty$, (12) is the covariance between two observations that are in the same sub-cell of the data but have different error terms, those error terms in the finite population case have a covariance $-\sigma_e^2/N_e$, in accord with (7). Hence (12) becomes

$$\sigma_A^2(1 - N_A^{-1}) + \sigma_B^2(1 - N_B^{-1}) + \sigma_C^2(1 - N_C^{-1}) - \sigma_e^2/N_e \text{ in } \underline{V}_F \,. \tag{13}$$

Similarly

$$\sigma_A^2 + \sigma_B^2 \text{ in } \underline{V}_\infty \tag{14}$$

is the covariance between two observations that are in the same levels of A and B but, within that level of B, they are in different levels of C. Therefore, in the finite population case this covariance involves $-N_C^{-1}\sigma_C^2$, as in (7). Hence (14) becomes

$$(1 - N_A^{-1})\sigma_A^2 + (1 - N_B^{-1})\sigma_B^2 - N_C^{-1}\sigma_C^2 \text{ in } \underline{V}_F \,. \tag{15}$$

Likewise

$$\sigma_A^2 \text{ in } \underline{V}_\infty \tag{16}$$

becomes

$$(1 - N_A)^{-1}\sigma_A^2 - N_B^{-1}\sigma_B^2 \text{ in } \underline{V}_F \tag{17}$$

Also, for the same reason,

$$\text{zero elements in } \underline{V}_\infty \tag{18}$$

become

$$-N_A^{-1}\sigma_A \text{ in } \underline{V}_F \,. \tag{19}$$

From (11), (13), (15), (17) and (19) it is seen that $\underline{V}_F$ has $-N_A^{-1}\sigma_A^2$ in every element. Therefore $\underline{V}_F$ can be expressed as

$$\underline{V}_F = (-\sigma_A^2/N_A)\underline{1}\underline{1}' + \underline{V}_F^* . \tag{20}$$

Now, as already discussed, $\underline{Q}\underline{1} = \underline{0}$ and so

$$\underline{Q}\underline{V}_F = \underline{Q}\underline{V}_F^* .$$

Therefore in (9) ,

$$E_F(M) = \mathrm{tr}(\underline{Q}\underline{V}_F^*) \tag{21}$$

Now in (20), $\underline{V}_F^*$ is obtained by subtracting $(-\sigma_A^2/N_A)$ from every element in $\underline{V}_F$, and from (13), (15), (17) and (19) this gives elements of $\underline{V}_F^*$ as

$$(\sigma_A^2 - \sigma_B^2/N_B) + (\sigma_B^2 - \sigma_C^2/N_C) + (\sigma_C^2 - \sigma_e^2/N_e) + \sigma_e^2 , \tag{22}$$

$$(\sigma_A^2 - \sigma_B^2/N_B) + (\sigma_B^2 - \sigma_C^2/N_C) + (\sigma_C^2 - \sigma_e^2/N_e) , \tag{23}$$

$$(\sigma_A^2 - \sigma_B^2/N_B) + (\sigma_B^2 - \sigma_C^2/N_C) , \tag{24}$$

$$(\sigma_A^2 - \sigma_B^2/N_B) , \tag{25}$$

and zero. (25a)

These are elements of $\underline{V}_F^*$. The corresponding elements of $\underline{V}_\infty$ are given in (10), (12), (16) and (18); comparing the two sets of elements indicates that $\underline{V}_F^*$ can be derived from $\underline{V}_\infty$ if

$$\left.\begin{matrix} \sigma_A^2 \\ \sigma_B^2 \\ \sigma_C^2 \\ \sigma_e^2 \end{matrix}\right) \text{ of } \underline{V}_\infty \text{ is replaced by } \left(\begin{matrix} \sigma_A^2 - \sigma_B^2/N_B \\ \sigma_B^2 - \sigma_C^2/N_C \\ \sigma_C^2 - \sigma_e^2/N_e \\ \sigma_e^2 \end{matrix}\right. \tag{26}$$

Therefore, because $E_\infty(M) = tr(\underline{\underline{Q}}\underline{V}_\infty)$ as in (9) and $E_F(M) = tr(\underline{\underline{Q}}\underline{V}_F^*)$ as in (21), making the replacements of (26) in $E_\infty(M)$ yields $E_F(M)$; i.e. making the replacements of (26) in the expected mean squares of an infinite population model yields the corresponding expected mean squares of a finite population model.

The above discussion is in terms of the 3-way nested classification, but it clearly extends to any nested model. In general, every diagonal element in $\underline{V}_\infty$ is the sum of all variance components in the model, $\sum_{\theta=A}^{K} \sigma_\theta^2$, and so by (6) it becomes $\sum_{\theta=A}^{K} (1 - N_\theta^{-1})\sigma_\theta^2$ in $\underline{V}_F$; (10) and (11) illustrate this. Similarly, every non-zero off-diagonal element in $\underline{V}_\infty$ is a covariance between two observations and is a sum of certain σ_θ^2's of the model, which by (6) become $(1 - N_\theta^{-1})\sigma_\theta^2$ in $\underline{V}_F$. Furthermore, this covariance also contains a zero covariance which by (7) is non-zero in $\underline{V}_F$, of the form $-\sigma_\varphi^2/N_\varphi$, namely the covariance between two levels of the factor nested within the submost factor of those whose variances are in the element of $\underline{V}_\infty$. Examples of this are seen in equations (12) through (17). And finally, non-zero elements in $\underline{V}_\infty$ are covariances between observations in different levels of the A-factor and so in $\underline{V}_F$ these covariances are $-\sigma_A^2/N_A$, as exemplified in (18) and (19). Of course, observations in different levels of the A effect are also in different levels of all other effects, but these populations are defined only within the factors within which they are nested (and in particular within A), and so only $-\sigma_A^2/N_A$ enters into this covariance. This is also why only one term of the form $-\sigma_\varphi^2/N_\varphi$ arises in each of the other covariance elements of $\underline{V}_F$. The replacements indicated in (26) for the 3-way nested classification therefore extend readily to any k-way nested classification. Furthermore, in (26) we see that in replacing σ_A^2 by $\sigma_A^2 - \sigma_B^2/N_B$, B is the factor nested within A; similarly in replacing σ_B^2 by $\sigma_B^2 - \sigma_C^2/N_C$, C is the factor nested within B; likewise, in $\sigma_C^2 - \sigma_e^2/N_e$ where σ_e^2 is the error variance, the error terms can be thought of as being a 'factor' nested within the submost subclassification of the model. In this way the replacements of (26) can be generalized.

Let $\varphi:\theta$ denote the factor φ nested within θ, and let $N_{\varphi:\theta}$ (or N_φ) be the size of the φ-population within each level of the θ-factor -- for each level, the same sized population. Then we have the following rule.

Rule: Expected mean squares in nested classification models with finite populations are obtained from their values with infinite populations by

$$\text{replacing } \sigma_\theta^2 \text{ by } \sigma_\theta^2 - \sigma_\varphi^2/N_{\varphi:\theta} , \tag{27}$$

where σ_φ^2 is the variance component of the factor $\varphi:\theta$ nested immediately within θ.

Application of this rule means that if expectation of a mean square in the infinite population model is

$$E_\infty(M) = \sum_{\theta=A}^{K} \lambda_\theta \sigma_\theta^2 \tag{28}$$

then expectation in the finite population model is

$$E_F(M) = \sum_{\theta=A}^{K} \lambda_\theta(\sigma_\theta^2 - \sigma_\varphi^2/N_{\varphi:\theta}) . \tag{29}$$

When the θ-factor has no factor θ nested within it then σ_φ^2 in (29) is taken to be zero; for example, when θ represents the error "factor".

Because the σ_φ^2-terms in (29) are just the variance components of the model, just as are the σ_θ^2's, (29) can also be rewritten as

$$E_F(M) = \sum_{\theta=A}^{K} (\lambda_\theta - \lambda_\varphi/N_{\theta:\varphi})\sigma_\theta^2 . \tag{30}$$

In this case, when the θ-factor is nested within no other factor φ does not exist and so $\lambda_\varphi = 0$, e.g. when θ is the main classification of a nested classification it is nested within no other factor. Examples of (29) and (30) are shown below.

(30) provides immediate calculation of the coefficients of the σ^2's in a finite population model.

General use of the rule requires careful interpretation of what is meant by "the factor nested immediately within θ". In the 3-way nested classification, it is clear that C is the factor nested within B, and in (27) σ_B^2 is replaced by $\sigma_B^2 - \sigma_C^2/N_C$. But, although C is nested within A (because C is nested within B within A), σ_A^2 is not replaced by something involving C and B but only by $\sigma_A^2 - \sigma_B^2/N_B$, because B is "the factor nested immediately within A". Furthermore, error terms, already described as being a "factor" are, for purposes of the rule, nested within the sub-most sub-classification; e.g., in the above example σ_C^2 is replaced by $\sigma_C^2 - \sigma_e^2/N_e$.

Several examples of rule (27) are shown in Table 1. To illustrate their use, suppose in the 3-way nested

(Show Table 1)

classification, the data contain b levels of the B factor in each level of A, c levels of the C-factor in each level of the B-factor, and n observations in each cell. Then in the infinite population model the mean square for the A-factor is

$$E_\infty(MSA) = bcn\sigma_A^2 + cn\sigma_B^2 + n\sigma_c^2 + \sigma_e^2 .$$

With finite populations, on making the changes indicated in Table 1, the expected value is

$$E_F(MSA) = bcn(\sigma_A^2 - \sigma_B^2/N_B) + cn(\sigma_B^2 - \sigma_C^2/N_C) + n(\sigma_C^2 - \sigma_e^2/N_e) + \sigma_e^2 \qquad (31)$$

$$= bcn\,\sigma_A^2 + cn(1 - b/N_B)\sigma_B^2 + n(1 - c/N_C)\sigma_C^2 + (1 - n/N_e)\sigma_e^2 . \qquad (32)$$

These results are examples of (29) and (30); and (32) is, of course, the familiar result for finite population given for example, in Bennett and Franklin (1954).

The importance of (27) [and hence of (28) and (29)] is, as has been indicated in its derivation, that it applies to unbalanced data as well as to balanced data. For example, consider the 3-way nested classification with n_{ijk} observations in the k'th level of C in the j'th level of B in the i'th level of A, with a levels of A, in the i'th one of which there are b_i levels of B whose j'th level therein has c_{ij} levels of C. With infinite populations the expected mean square for B within A is given by Anderson and Bancroft (1952, p. 327) as

$$E_\infty(M) = \lambda_B\sigma_B^2 + \lambda_C\sigma_C^2 + \sigma_e^2 \tag{33}$$

where

$$\lambda_B = \sum_{i=1}^{a}\sum_{j=1}^{b_i} n_{ij.}^2 f_{ij} \quad \text{and} \quad \lambda_C = \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{c_{ij}} n_{ijk}^2 f_{ij}$$

with $f_{ij} = \left(n_{ij.}^{-1} - n_{i..}^{-1}\right)\left(\sum_{i=1}^{a} b_i - a\right)^{-1}$. On making the replacements suggested in Table 1, the expected value for finite populations is

$$E_F(M) = \lambda_B(\sigma_B^2 - \sigma_C^2/N_C) + \lambda_C(\sigma_C^2 - \sigma_e^2/N_e) + \sigma_e^2 \tag{34}$$

$$= \lambda_B\sigma_B^2 + \sigma_C^2(\lambda_C - \lambda_B/N_C) + \sigma_e^2(1 - \lambda_C/N_e) . \tag{35}$$

Thus, given (33) for infinite populations, the results (34) and (35) for finite populations follow at once; they are examples of (29) and (30). We note that when N_e is taken as infinite

$$E_F(M) = \lambda_B\sigma_B^2 + \sigma_C^2(\lambda_C - \lambda_B/N_C) + \sigma_e^2$$

the result given by Gaylor and Hartwell (1969). Their other results, and all finite population results, follow similarly.

Another illustration from Table 1 is the 1-way classification. With infinite populations the mean squares MSB (between groups) and MSW (within groups) have expectations

$$E(MSB) = n\sigma_A^2 + \sigma_e^2 \quad \text{and} \quad E(MSE) = \sigma_e^2 \tag{36}$$

With finite populations these become the familiar results

$$E(MSB) = n(\sigma_A^2 - \sigma_e^2/N_e) + \sigma_e^2 = n\sigma_A^2 + (1 - n/N_e)\sigma_e^2 \tag{37}$$

and

$$E(MSE) = \sigma_e^2 ,$$

given, for example, by McHugh and Miekle (1968). They suggest that when (36) leads to the estimate

$$\hat{\sigma}_A^2 = (MSB - MSE)/n \tag{38}$$

being negative, then assuming a finite population of error terms and using (37) might not give a negative estimate; i.e.

$$\hat{\sigma}_A^2 = [MSB - (1 - n/N_e)MSE]/n \tag{39}$$

might not be negative. This is an interesting suggestion. However, on writing

$$F = MSB/MSE$$

it is easily seen that (39) is positive only when

$$N_e < \frac{n}{1-F} . \tag{40}$$

Therefore, to use (39) and have it yield a positive estimate we must conjecture (or have) a value for N_e that satisfies (40). Now in using (39) when (38) is negative, i.e. when $F < 1$, it is clear that (40) provides a positive upper limit to N_e. But this limit will only exceed n appreciably when F (<1) is close to unity. Conversely, when F (<1) is close to zero the limit will be close to n, in which case the concept of a finite population of N_e (only just greater than n) error terms may not be very palatable. Thus the idea of using finite population models as alternatives to infinite populations in order to re-calculate variance components that would otherwise be negative may have appeal only when F-values are just less than unity and not when they are a great deal less.

General application of rule (27) involves two points: (i) The error variance, σ_e^2, is always replaced by itself. (ii) Models in which only some populations are finite have infinite N's for the other populations. For example, in most cases an infinite population would be assumed for the error terms and so N_e would be taken as infinite. With these two provisos (27) applies to any nested classification model, with balanced or unbalanced data.

The 2-way crossed classification

In the two-way crossed classification with factors A and B, the sampled values of the populations of the two main effects, the A_i and the B_j, will both have properties similar to (6) and (7). Furthermore, defining the population of interaction effects in such a way that

$$\sum_{i=1}^{N_A} (AB)_{ij} = 0 \text{ for all } j, \text{ and } \sum_{j=1}^{N_B} (AB)_{ij} = 0 \text{ for all } i \tag{41}$$

leads to defining

$$\sigma_{AB}^2 = \frac{\sum_{p=1}^{A} \sum_{q=1}^{B} (AB)_{pq}}{(N_A - 1)(N_B - 1)} . \tag{42}$$

This in turn, through an extension of the procedure by which (6) and (7) were derived leads to

$$E[(AB)_{ij}(AB)_{i'j'}] = \begin{cases} \dfrac{(N_A - 1)(N_B - 1)}{N_A N_B}\,\sigma^2_{AB} & \text{for } i = i' \text{ and } j = j' ; \\[2ex] \dfrac{-(N_A - 1)}{N_A N_B}\,\sigma^2_{AB} & \text{for } i = i' \text{ and } j \neq j' \\[2ex] \dfrac{-(N_B - 1)}{N_A N_B}\,\sigma^2_{AB} & \text{for } i \neq i' \text{ and } j = j' \\[2ex] \dfrac{\sigma^2_{AB}}{N_A N_B} & \text{for } i \neq i' \text{ and } j \neq j' . \end{cases} \tag{43}$$

where $(AB)_{ij}$ and $(AB)_{i'j'}$ are two sample values of the interaction effects.

To derive $\underline{V}_F$ from $\underline{V}_\infty$ in this case, consider first the elements of $\underline{V}_\infty$. They will be of five different kinds:

$$\begin{aligned} v_1 &= \sigma^2_A + \sigma^2_B + \sigma^2_{AB} + \sigma^2_e &&= \text{var}(y_{ijk}) ; \\ v_2 &= \sigma^2_A + \sigma^2_B + \sigma^2_{AB} &&= \text{cov}(y_{ijk}, y_{ijk'}) \text{ for } k \neq k' ; \\ v_3 &= \sigma^2_A &&= \text{cov}(y_{ijk}, y_{ij'k}) \text{ for } j \neq j' ; \\ v_4 &= \sigma^2_B &&= \text{cov}(y_{ijk}, y_{i'jk}) \text{ for } i \neq i' ; \\ \text{and} \quad v_5 &= 0 &&= \text{cov}(y_{ijk}, y_{i'j'k}) \text{ for } i \neq i', j \neq j' . \end{aligned} \tag{44}$$

Now the model is

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

with $E(y_{ijk}) = \mu$, with A_i, B_j and e_{ijk} being samples from finite populations with properties similar to (6) and (7); and the $(AB)_{ij}$ have the properties indicated in (43). Under these conditions it is not difficult to show that the elements of $\underline{V}_F$ corresponding to those of $\underline{V}_\infty$ are

$$
\begin{aligned}
f_1 &= \sigma_A^2(1 - N_A^{-1}) + \sigma_B^2(1 - N_B^{-1}) + \sigma_{AB}^2(1 - N_A^{-1})(1 - N_B^{-1}) + \sigma_e^2(1 - N_e^{-1}) , \\
f_2 &= \sigma_A^2(1 - N_A^{-1}) + \sigma_B^2(1 - N_B^{-1}) + \sigma_{AB}^2(1 - N_A^{-1})(1 - N_B^{-1}) - \sigma_e^2 N_e^{-1} , \\
f_3 &= \sigma_A^2(1 - N_A^{-1}) - \sigma_B^2 N_B^{-1} \qquad - \sigma_{AB}^2(1 - N_A^{-1})N_B^{-1} , \\
f_4 &= -\sigma_A^2 N_A^{-1} \qquad + \sigma_B^2(1 - N_B^{-1}) - \sigma_{AB}^2 N_A^{-1}(1 - N_B^{-1}) , \\
f_5 &= -\sigma_A^2 N_A^{-1} \qquad - \sigma_B^2 N_B^{-1} \qquad + \sigma_{AB}^2 N_A^{-1} N_B^{-1} .
\end{aligned}
\tag{45}
$$

With a little re-arranging of terms these can be written as

$$
\begin{aligned}
f_1 &= f_5 + f_1^*, \text{ with } f_1^* = (\sigma_A^2 - \sigma_{AB}^2/N_B) + (\sigma_B^2 - \sigma_{AB}^2/N_A) + (\sigma_{AB}^2 - \sigma_e^2/N_e) + \sigma_e^2 , \\
f_2 &= f_5 + f_2^* \text{ with } f_2^* = (\sigma_A^2 - \sigma_{AB}^2/N_B) + (\sigma_B^2 - \sigma_{AB}^2/N_A) + (\sigma_{AB}^2 - \sigma_e^2/N_e) , \\
f_3 &= f_5 + f_3^* \text{ with } f_3^* = (\sigma_A^2 - \sigma_{AB}^2/N_B) , \\
f_4 &= f_5 + f_4^* \text{ with } f_4^* = \qquad (\sigma_B^2 - \sigma_{AB}^2/N_A) , \\
f_5 &= f_5 + f_5^* \text{ with } f_5^* = 0 .
\end{aligned}
\tag{46}
$$

Since the f's of (45) are elements of $\underset{-}{V}_F$, it is clear from (46) that f_5 is part of every element of $\underset{-}{V}_F$ and so $\underset{-}{V}_F = f_5\underline{1}\underline{1}' + \underset{-}{V}_F^*$ where the f^*'s are elements of $\underset{-}{V}_F^*$. Therefore, because $\underline{Q}\underline{1} = \underline{0}$, we have $E_F(M) = tr(\underline{Q}\underline{V}_F) = tr(\underline{Q}\underline{V}_F^*)$ just as in nested models. Furthermore, on comparing (46) with (44) it is evident that the f^*'s are the v's with

$$
\begin{aligned}
&\sigma_A^2 \text{ replaced by } \sigma_A^2 - \sigma_{AB}^2/N_B \\
&\sigma_B^2 \text{ replaced by } \sigma_B^2 - \sigma_{AB}^2/N_A \\
&\sigma_{AB}^2 \text{ replaced by } \sigma_{AB}^2 - \sigma_e^2/N_e \\
&\sigma_e^2 \text{ replaced by } \sigma_e^2 ,
\end{aligned}
\tag{47}
$$

and so $E_F(M)$ is $E_\infty(M)$ after making these same replacements. These are the results shown in Table 2.

The general crossed classification

The 3-way cross-classification model is handled by a natural extension of the methods used above in equation (43)-(47) for the 2-way case. The results are shown in Table 1. In obtaining them, all three of the two-factor interactions of the model behave in the manner of (38), and the behavior of the three-factor interaction is similar, only it has eight different terms in its counterpart of (38), and not four as shown there. The algebraic complexity of the detailed steps is, of course, greater than that for the 2-way case shown above. Furthermore, as Cornfield and Tukey (1956) so rightly say, in carrying out these detailed steps the "systematic algebra can take us deep into the forest of notation. But the detailed manipulation will, sooner or later blot out any understanding we may have started with". With this we wholeheartedly agree! In addition, having accomplished this we would then, so far as developing a general result is concerned, be no more than "ready for another step of induction and so on" as Cornfield and Tukey aptly put it. We therefore give the general result towards which this induction apparently leads. It is simple: in an r-way crossed classification having main effects A, B, ..., R replace the error component

$$\sigma_e^2 \text{ by } \sigma_e^2 ; \tag{48}$$

and for the highest order interaction component

$$\text{replace } \sigma^2_{AB...R} \text{ by } \sigma^2_{AB...R} - \sigma_e^2/N_e \tag{49}$$

and for any other interaction or main effect components $\sigma^2_{DE...L}$, where D, E, ..., L is a subset of any 1, 2, ..., r-1 letters from A, B, ..., R,

$$\text{replace } \sigma^2_{DE\ldots L} \text{ by } \sigma^2_{DE\ldots L} \prod_{\substack{\theta \in \text{ all main} \\ \text{effects that} \\ \text{cross } DE\ldots L}} (1 - g_\theta) \qquad (50)$$

In this expression g_θ is an operational function: multiplying a σ^2 by g_θ adds θ to the subscript of σ^2 and divides the result by N_θ; e.g.,

$$\sigma^2_A g_B = \sigma^2_{AB}/N_B \quad \text{and} \quad \sigma^2_{AB} g_C = \sigma^2_{ABC}/N_C \; .$$

Furthermore, multiplication of g's is symbolic, so that

$$\sigma^2_A g_B g_C = \sigma^2_{ABC}/N_B N_C \quad \text{and} \quad \sigma^2_A g_B g_C g_D = \sigma^2_{ABCD}/N_B N_C N_D \; .$$

Using g_θ in this manner, examples of (50) are seen in Table 2. Thus in the 3-way crossed classification σ^2_A is replaced by

$$\sigma^2_A \prod_{\theta=B}^{C} (1 - g_\theta) = \sigma^2_A (1 - g_B)(1 - g_C) = \sigma^2_A - \sigma^2_{AB}/N_B - \sigma^2_{AC}/N_C + \sigma^2_{ABC}/N_B N_C$$

as shown. Similarly σ^2_{AB} becomes

$$\sigma^2_{AB}(1 - g_C) = \sigma^2_{AB} - \sigma^2_{ABC}/N_C$$

These results are, of course, in agreement with those of Wilk and Kempthorne (1956), whose subscripted Σ's (pages 977 <u>et seq</u>), ignoring their Q's, are identical to the expressions emanating from (50). Of these expressions we can say, as do Wilk and Kempthorne, that their "general pattern has been obtained by one or both of the present authors for more complex designs and situations than we have studied" in detail in this paper. The results for some of these situations are shown in Table 2.

(Show Table 2)

In addition, as shown in the next section, we have extended this result for crossed classification models to a more general result applicable to models that consist of any combination of crossed and/or nested classifications. Some results of these are shown in Table 4.

On some occasions use is made of models that do not include all the interactions that could be included; e.g. the 2-way classification without interaction. In this case, expectations of mean squares with finite populations are obtained by first finding their values with infinite populations with all interactions included, then making the replacements as indicated above, and then putting the interaction variances identically equal to zero. For example, in the 2-way classification without interaction, the expected mean square for the A-factor is developed as shown in Table 3.

(Show Table 3)

Combinations of nested and crossed classifications

The rule developed in (27) for nested classifications conforms with (50) for crossed classifications. For φ nested within θ let $h_{\varphi:\theta}$ be an operational function, similar in nature to g_θ of (50) but such that $h_{\varphi:\theta}$ multiplies only σ_θ^2 and

$$\sigma_\theta^2 h_{\varphi:\theta} = \sigma_\varphi^2/N_\varphi \,. \tag{51}$$

Then the rule (27) for nested classifications is

$$\text{replace } \sigma_\theta^2 \text{ by } \sigma_\theta^2(1 - h_{\varphi:\theta}) \tag{52}$$

where φ is the factor nested immediately within θ. Operationally, the effect of multiplying σ_θ^2 by $h_{\varphi:\theta}$ is to add $\varphi:\theta$ to the subscript of σ_θ^2 and then "cancel" the θ's because θ occurs on the right of the colon in h as well as in the σ^2; and to divide the result, σ_φ^2, by N_φ, so giving (51) as the product. The value of this description of $h_{\varphi:\theta}$ will be evident in its general application.

It is clear that (52) applies to all factors of a completely nested model. It also applies, for crossed classification models, to (48):

$$\text{replace } \sigma_e^2 \text{ by } \sigma_e^2(1 - h_{\varphi:\theta}) = \sigma_e^2 \tag{53}$$

which is (48), this result being so because $h_{\varphi:e} = 0$ - there are no factors nested within the error factor. Similarly, it applies to (49):

$$\text{replace } \sigma^2_{AB\ldots R} \text{ by } \sigma^2_{AB\ldots R}(1 - h_{e:AB\ldots R}) = \sigma^2_{AB\ldots R} - \sigma_e^2/N_e\ , \tag{54}$$

the term in σ_e^2 coming from the definition (51) in which, of course, θ is not necessarily a main effect but can be, as here, an interaction. Finally, the general result (50) can be amended to take account of nested classifications by multiplying it by expressions $\left(1 - h_{\varphi:\theta_{DE\ldots L}}\right)$ for any factor φ that is the factor nested immediately within $\theta_{DE\ldots L}$, where $\theta_{DE\ldots L}$ is any main effect D, E, ..., or L, or any interaction factor of these main effects. Hence, to cover all combinations of nested and crossed classifications, including the error 'factor', (50) is generalized as follows. In a model having factors A, B, ..., R which are nested and/or crossed,

$$\text{replace } \sigma^2_{DE\ldots L} \text{ by } \sigma^2_{DE\ldots L} \prod_{\substack{\theta\,\epsilon \text{ all main} \\ \text{effects that} \\ \text{cross } DE\ldots L}} (1 - g_\theta) \prod_{\varphi} (1 - h_{\varphi:\theta_{DE\ldots L}})\ , \tag{55}$$

where φ is the factor nested within $\theta_{DE\ldots L}$. By using $\theta_{DE\ldots L}$ in the subscript of h, (55) provides for the error term being a nested factor, as illustrated in (54), and in the same way it also makes provision for situations in which a factor might be nested within an interaction of two (or more) other factors. However, nested factors are usually nested within just single factors and not interactions of them, in which case apart from (49), the general result (55) becomes

$$\text{replace } \sigma^2_{DE\ldots L} \text{ by } \sigma^2_{DE\ldots L} \prod_{\substack{\theta\ \epsilon \text{ all main} \\ \text{effects that} \\ \text{cross } DE\ldots L}} (1 - g_\theta) \prod_{\theta=D}^{L} (1 - h_{\varphi:\theta}) \quad (56)$$

Examples of (56) are shown in Table 4. Illustrations of its use in the first three entries of Table 4 are as follows. In deriving them we use the verbal description of the h-operation given below equation (52).

$$\begin{aligned} \sigma^2_A \text{ is replaced by } \quad & \sigma^2_A(1 - g_B)(1 - h_{P:A}) \\ &= (\sigma^2_A - \sigma^2_{AB}/N_B)(1 - h_{P:A}) \\ &= \sigma^2_A - \sigma^2_{AB}/N_B - \sigma^2_P/N_P + \sigma^2_{PB}/N_PN_B \end{aligned}$$

$$\sigma^2_B \text{ is replaced by } \quad \sigma^2_B(1 - g_A) = \sigma^2_B - \sigma^2_{AB}/N_A$$

$$\sigma^2_{AB} \text{ is replaced by } \quad \sigma^2_{AB}(1 - h_{P:A}) = \sigma^2_{AB} - \sigma^2_{PB}/N_P \ .$$

Derivation of the other results in Table 4 follows in similar manner. Clearly, when there are no nested factors, h-terms are zero and (55) and (56) reduce to (50) for crossed classification models; and when there are no crossed factors g-terms are zero and (55) and (56) reduce to (52) and hence to (27) for nested models. In this way (55) - and its slightly simpler form (56) - applies for all models.

Mixed Models

The above discussion concerns random effects models, and relates to expectations of mean squares in analyses of variance of balanced data and of "mean squares" used in the analogous analysis of variance of unbalanced data [Henderson's (1953) Method 1], discussed in Searle (1968). In these cases there are no terms in μ in the mean squares. Similarly, the procedures also apply with mixed models, for mean squares that contain no terms in the fixed effects. With balanced data, enough of these always occur in the analysis of variance to estimate the variance components, and to these mean squares the rule for changing from infinite to finite populations applies. With unbalanced data, the same situation arises in the fitting countants method [Henderson's (1953) Method 3]. In this method, as shown in Searle (1968), there are always differences between certain residuals that contain no fixed effects and so can be used to estimate the variances unbiasedly. To these differences (47) also applies. The general procedure (47) therefore has widespread application, to balanced and unbalanced data in either random or mixed models.

Acknowledgment

This paper has been supported in part by Research Grant GM 13225-04 from the National Institute of General Medical Sciences, Bethesda, Maryland.

References

Anderson, R. L., and Bancroft, T. A. (1952). "Statistical Theory in Research". McGraw-Hill, New York.

Bennett, C. A. and Franklin, N. L. (1954). "Statistical Analysis in Chemistry and the Chemical Industry". Wiley, New York.

Cornfield, J. and Tukey, J. W. (1956). Average values of mean squares in factorials. Ann. Math. Stat., 907-949.

Gaylor, D. W. and Hartwell, T. D. (1969). Expected mean squares for nested classifications. (in press).

Henderson, C. R. (1953). Estimation of variance and covariance components. Biometrics, 9, 226-252.

McHugh, Richard B. and Miekle, Paul W., Jr. (1968). Negative variance estimates and statistical dependence in nested sampling. J. Am. Stat. Assoc. 63, 1000-1.

Searle, S. R. (1968). Another look at Henderson's methods of estimating variance components. Biometrics, 24, 749-788.

Tukey, J. W. (1956). Variances of variance components: I balanced designs. Ann. Math. Stat. 27, 722-736.

Tukey, J. W. (1957). Variances of variance components: II The unbalanced single classification. Ann. Math. Stat. 28, 43-56.

Wilk, M. B. and Kempthorne O. (1956). Some aspects of the analysis of factorial experiments in a completely randomized design. Ann. Math. Stat., 950-985.

Table 1: Nested Classifications

Expectations of mean squares under finite populations are obtained from those under infinite populations by making the replacements shown.

1-way	2-way	3-way
A	B within A	B within A, C within B
Replace	Replace	Replace
σ_A^2 by $\sigma_A^2 - \sigma_e^2/N_e$	σ_A^2 by $\sigma_A^2 - \sigma_B^2/N_B$	σ_A^2 by $\sigma_A^2 - \sigma_B^2/N_B$
	σ_B^2 by $\sigma_B^2 - \sigma_e^2/N_e$	σ_B^2 by $\sigma_B^2 - \sigma_C^2/N_C$
		σ_C^2 by $\sigma_C^2 - \sigma_e^2/N_e$

Table 2: Crossed Classifications

Expectations of mean squares under finite populations are obtained from those under infinite populations by making the replacements shown.

1-way	2-way	3-way
A	A and B	A, B and C
Replace	Replace	Replace
σ_A^2 by $\sigma_A^2 - \sigma_e^2/N_e$	σ_A^2 by $\sigma_A^2 - \sigma_{AB}^2/N_B$	σ_A^2 by $\sigma_A^2-\sigma_{AB}^2/N_B-\sigma_{AC}^2/N_C+\sigma_{ABC}^2/N_BN_C$
	σ_B^2 by $\sigma_B^2 - \sigma_{AB}^2/N_A$	σ_B^2 by $\sigma_B^2-\sigma_{AB}^2/N_A-\sigma_{BC}^2/N_C+\sigma_{ABC}^2/N_AN_C$
	σ_{AB}^2 by $\sigma_{AB}^2 - \sigma_e^2/N_e$	σ_C^2 by $\sigma_C^2-\sigma_{AC}^2/N_A-\sigma_{BC}^2/N_B+\sigma_{ABC}^2/N_AN_B$
		σ_{AB}^2 by $\sigma_{AB}^2-\sigma_{ABC}^2/N_C$
		σ_{AC}^2 by $\sigma_{AC}^2-\sigma_{ABC}^2/N_B$
		σ_{BC}^2 by $\sigma_{BC}^2 - \sigma_{ABC}^2/N_A$
		σ_{ABC}^2 by $\sigma_{ABC}^2-\sigma_e^2/N_e$

Table 3

Expected mean square for the A-factor in a 2-way crossed classification (b levels of factor B in the data, and n observations per cell).

Model	Expectation of Mean Square for A
Infinite populations, no interaction	$bn\sigma_A^2 + \sigma_e^2$
Infinite populations, interaction	$bn\sigma_A^2 + n\sigma_{AB}^2 + \sigma_e^2$
Finite populations, interaction	$bn(\sigma_A^2 - \sigma_{AB}^2/N_B) + n(\sigma_{AB}^2 - \sigma_e^2/N_e) + \sigma_e^2$
Finite populations, no interaction	$bn\sigma_A^2 - n\sigma_e^2/N_e + \sigma_e^2$

Table 4: Combinations of Crossed and Nested Classifications

Expectations of mean squares under finite populations are obtained from those under infinite populations by making the replacements shown.

2 crossed and 1 nested

A and B crossed, and P nested within A

σ^2_A by $\sigma^2_A - \sigma^2_{AB}/N_B - \sigma^2_P/N_P + \sigma^2_{PB}/N_P N_B$

σ^2_B by $\sigma^2_B - \sigma^2_{AB}/N_A$

σ^2_{AB} by $\sigma^2_{AB} - \sigma^2_{PB}/N_P$

σ^2_P by $\sigma^2_P - \sigma^2_{PB}/N_B$

σ^2_{PB} by $\sigma^2_{PB} - \sigma^2_e/N_e$

2 crossed and 2 nested

A and B crossed, P within A and Q within B

σ^2_A by $\sigma^2_A - \sigma^2_{AB}/N_B - \sigma^2_P/N_P + \sigma^2_{PB}/N_P N_B$

σ^2_B by $\sigma^2_B - \sigma^2_{AB}/N_A - \sigma^2_Q/N_Q + \sigma^2_{AQ}/N_A N_Q$

σ^2_{AB} by $\sigma^2_{AB} - \sigma^2_{AQ}/N_Q - \sigma^2_{PB}/N_P + \sigma^2_{PQ}/N_P N_Q$

σ^2_P by $\sigma^2_P - \sigma^2_{PB}/N_B$

σ^2_{PB} by $\sigma^2_{PB} - \sigma^2_{PQ}/N_Q$

σ^2_Q by $\sigma^2_Q - \sigma^2_{AQ}/N_A$

σ^2_{AQ} by $\sigma^2_{AQ} - \sigma^2_{PQ}/N_P$

σ^2_{PQ} by $\sigma^2_{PQ} - \sigma^2_e/N_e$

A and B crossed, P within A and Q within P

σ^2_A by $\sigma^2_A - \sigma^2_{AB}/N_B - \sigma^2_P/N_P + \sigma^2_{PB}/N_P N_B$

σ^2_B by $\sigma^2_B - \sigma^2_{AB}/N_A$

σ^2_{AB} by $\sigma^2_{AB} - \sigma^2_{PB}/N_P$

σ^2_P by $\sigma^2_P - \sigma^2_{PB}/N_B - \sigma^2_Q/N_Q + \sigma^2_{QB}/N_Q N_B$

σ^2_{PB} by $\sigma^2_{PB} - \sigma^2_{QB}/N_Q$

σ^2_A by $\sigma^2_Q - \sigma^2_{QB}/N_B$

σ^2_{QB} by $\sigma^2_{QB} - \sigma^2_e/N_e$