ON THE DISTRIBUTION OF VARIANCE COMPONENT ESTIMATES IN THE UNBALANCED ONE-WAY CLASSIFICATION

S. R. Searle

Biometrics Unit, Cornell University, Ithaca, New York

ABSTRACT

Computer simulation studies have been made of the distribution of between-group variance component estimates customarily derived from unbalanced data of a 1-way classification model. Under normality assumptions, the distribution is, in many instances, akin to a χ^2, although in some cases it is exponential in nature. An approximation to the distribution function appears feasible in some situations but not in others.

Paper No. ~~BU-167~~ in the Biometrics Unit Series and No. 560 in the Plant Breeding Series.

ON THE DISTRIBUTION OF VARIANCE COMPONENT ESTIMATES IN THE UNBALANCED ONE-WAY CLASSIFICATION*

S. R. Searle

Biometrics Unit, Cornell University, Ithaca, New York

Introduction

Analysis of data on a between- and within-groups basis arises on many occasions: between and within groups of people, between and within the crosses of two species, between and within litters of pigs, or herds of cows, or flocks of poultry; and between and within replications of technological processes. The statistical model appropriate to many of these situations is the well-known random effects model, viz. Model II of Eisenhart (1947). If data consist of n_i observations in the i'th group, for $i = 1,2,\ldots,c$, the equation of the model for y_{ij}, the j'th observation in the i'th group is

$$y_{ij} = \mu + a_i + e_{ij}$$

where μ is a general mean, a_i is the effect due to the i'th group and e_{ij} is a random error term. In the random model the a_i are assumed to be a random sample of a's from a population having zero mean and variance σ_a^2, being uncorrelated with each other and with the e_{ij}-terms which themselves are assumed to have zero mean and variance σ_e^2, they too being uncorrelated with each other. In this

Paper No. BU-167 in the Biometrics Unit Series and No. 560 in the Plant Breeding Series.

*The bulk of this paper was presented at the 6th International Biometric Conference at Sydney, Australia, August, 1967. Much of it was prepared with support from Grant Number GM 13225 of the National Institutes of Health.

context the matter of interest is to estimate the variance components σ_a^2 and σ_e^2 from observations y_{ij} for $j = 1,2,\ldots,n_i$, and $i = 1,2,\ldots,c$. i.e. a situation where the data come from c groups with n_i observations in the i'th group, a total of $N = \sum_{i=1}^{c} n_i$ observations in all. Having obtained estimates of σ_a^2 and σ_e^2 one is also interested in the sampling distribution (and variance) of the estimates.

No great problems arise when there is the same number of observations in each group, $n_i = n$ say, for all i. Data of this nature are usually referred to (and shall be here) as balanced data, and in this situation the distribution of the customary estimator of σ_a^2, the between groups variance component, can be derived (Robinson, 1966 and Wang, 1967). However, in many instances data are such that the groups do not each have the same number of observations. These data are called unbalanced. Theoretical considerations of the distribution of the estimator of σ_a^2 are then more complex, and light can be shed on the properties of this distribution by means of computer simulation. Leone and Nelson (1966) have recently pursued this approach for a 4-stage balanced nested design. Anderson and Crump (1967) have also used the same approach for some unbalanced designs, giving major consideration to just the sampling variance of estimators. The designs they deal with are unbalanced but in a manner that could be called planned, for they are largely concerned with the allocation of resources in situations pertinent to industrial experiments. Bainbridge (1963) and Bush and Anderson (1963) have also considered other planned, unbalanced designs suited to industrial contexts, where the inequality of the numbers of observations in the groups is, in some sense, more or less under control. But in biology these numbers are often under little or no control at all; animals in experiments die at will and organisms reproduce freely - and with survey data, such as are available in dairy herd breeding and poultry breeding, for example, the statistician is given the data just as they are, with scant possibility for determining group sizes or even the number of groups. Yet, in genetics especially, we frequently derive estimates of variance components from data of this sort, so it is important that studies be made of the sampling distributions of such estimates. Some of the problems involved and results obtained from initial studies are considered below.

Customary estimates

The usual procedure for deriving estimates from the data having n_i observations in the i'th group, vide Henderson (1953), is to calculate the between- and within-group mean squares,

$$MSB = \frac{1}{c-1}\left[\sum_{i=1}^{c}\left(\sum_{j=1}^{n_i} y_{ij}\right)^2 \Big/ n_i - \left(\sum_{ij}\sum y_{ij}\right)^2 \Big/ N\right]$$

--- (1)

and

$$MSW = \frac{1}{N-c}\left[\sum_{i=1}^{c}\sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^{c}\left(\sum_{j=1}^{n_i} y_{ij}\right)^2 \Big/ n_i\right]$$

and equate these values to their expectations. Solving the resulting equations for σ_a^2 and σ_e^2 leads to estimators

$$\hat{\sigma}_e^2 = MSW\ , \quad \text{and}\ \hat{\sigma}_a^2 = \frac{MSB - MSW}{\dfrac{N^2 - \sum_{i=1}^{c} n_i^2}{N(c-1)}}\ . \qquad \text{--- (2)}$$

Distributional properties of these estimators are usually considered only on the basis of normality assumptions, namely that the a_i and e_{ij} of the model have the properties already alluded to and, additionally, are normally distributed.

Under these conditions the σ_e^2 estimator, which is the within-group mean square, has a χ^2-distribution or, more accurately, $(N - c)MSW/\sigma_e^2$ has a χ^2-distribution with $(N - c)$ degrees of freedom. Hence the variance of $\hat{\sigma}_e^2$ is

$$\mathrm{var}(\hat{\sigma}_e^2) = 2\sigma_e^4/(N - c) \ . \qquad \text{--- (3)}$$

Thus the sampling distribution and variance of $\hat{\sigma}_e^2$ are readily established. However, as is well-known, the same is not true of $\hat{\sigma}_a^2$ and indeed, the two cases of balanced and of unbalanced data are initially best distinguished.

Balanced data

When each group has the same number of observations, $n_i = n$ for all i and in (2) the estimate of σ_a^2 becomes

$$\hat{\sigma}_a^2 = (MSB - MSW)/n \ .$$

Furthermore, $(c - 1)MSB/(n\sigma_a^2 + \sigma_e^2)$ has a χ^2-distribution with c - 1 degrees of freedom, independently of $(N - c)MSW/\sigma_e^2$ which also has a χ^2-distribution, with N - c degrees of freedom, with N = nc. Hence $\hat{\sigma}_a^2$ is the weighted difference between two independent χ^2 distributions:

$$\hat{\sigma}_a^2 = \frac{\sigma_a^2 + \sigma_e^2/n}{c - 1}\left[\frac{(c - 1)MSB}{n\sigma_a^2 + \sigma_e^2}\right] - \frac{\sigma_e^2}{n(N - c)}\left[\frac{(N - c)MSW}{\sigma_e^2}\right]$$

$$= \frac{\sigma_a^2 + \sigma_e^2/n}{c - 1}\chi^2_{c-1} - \frac{\sigma_e^2}{nc(n - 1)}\chi^2_{c(n-1)} \ . \qquad \text{--- (4)}$$

The density function of (4), as available from results of Robinson (1965) and Wang (1967), is a confluent hypergeometric function involving the coefficients of the χ^2's in (4) and so involves the parameters we seek to estimate, σ_a^2 and σ_e^2, in no straightforward manner. Wang (1967) considers this distribution in some detail, and we shall treat it solely as a special case of unbalanced data. Its variance is

$$\mathrm{var}(\hat{\sigma}_a^2) = \frac{2(n\sigma_a^2 + \sigma_e^2)^2}{n^2(c - 1)} + \frac{2\sigma_e^4}{n^2c(n - 1)}$$

and the covariance between $\hat{\sigma}_a^2$ and $\hat{\sigma}_e^2$ is

$$\mathrm{cov}(\hat{\sigma}_a^2, \hat{\sigma}_e^2) = -(1/n)\mathrm{var}(\hat{\sigma}_e^2) .$$

Unbalanced data

For data where the number of observations is not the same in every group, the estimator of σ_a^2 is as given in (2). But now MSB is not a χ^2-variable: it is a weighted sum of 1-degree-of-freedom χ^2-variables. The procedure of Robinson (1965), who obtains the density of $\sum_{i=1}^{r} \alpha_i \chi_{p_i}^2 - \sum_{j=1}^{s} \beta_j \chi_{q_j}^2$ for independent χ^2-variables and constants α_i and β_j, could thus be invoked except that the constants involved in $\hat{\sigma}_a^2$ are functions of σ_a^2, σ_e^2 and the n_i, since the weights in MSB are in terms of expansions like $n_i/(n_i\sigma_a^2 + \sigma_e^2)$. Despite this, the cumulants of the distribution are available, simulation of it can be undertaken, and evaluation can be made of an approximation to the distribution derived from an analogue of that for balanced data. To these topics we now turn.

Variance of Between-Groups Estimate

The estimator $\hat{\sigma}_a^2$ given in (2) is not maximum likelihood, as are those of Herbach (1959) and Thompson (1962) nor is it admissible, Robson (1965). It is, however, unbiased no matter what underlying form of distribution is attributed to the a_i's and e_{ij}'s — so long as they have zero means, variances σ_a^2 and σ_e^2 and are uncorrelated, as previously described. Under normality assumptions the variance of $\hat{\sigma}_a^2$ is

$$\mathrm{var}(\hat{\sigma}_a^2) = 2N \left[\frac{N(N-1)(c-1)}{(N-c)} \frac{\sigma_e^4}{(N^2-S_2)^2} + \frac{2\sigma_e^2\sigma_a^2}{N^2-S_2} + \frac{(N^2S_2+S_2^2-2NS_3)}{N(N^2-S_2)^2} \sigma_a^4 \right] \quad \text{- - - (5)}$$

where $S_2 = \Sigma n_i^2$ and $S_3 = \Sigma n_i^3$; and the covariance between $\hat{\sigma}_a^2$ and $\hat{\sigma}_e^2$ is

$$\mathrm{cov}(\hat{\sigma}_a^2, \hat{\sigma}_e^2) = \frac{-N(c-1)\mathrm{var}\,\hat{\sigma}_e^2}{N^2-S_2} .$$

Expressions given earlier for this variance and covariance are, of course, special cases of these formulae with $n_i = n$ for all i.

The variance of $\hat{\sigma}_a^2$ in (5) is a special case of the familiar result that the r'th cumulant of the quadratic form y'Fy (F being symmetric) is $2^{r-1}\,\mathrm{trace}(VF)^r$ where V is the variance-covariance matrix of y, a vector of normally distributed random variables having zero mean.

Derivation of (5) is given in Searle (1956); a typographical correction to the expression shown there is that the term in $\sigma_a^2\sigma_e^2$ should include the coefficient 2 as shown in (5) above. A correction to Crump's (1951) expression for $\mathrm{var}(\hat{\sigma}_a^2)$ might also be noted here. He uses notation akin to n_o for $(N^2 - S_2)/N(c - 1)$ and w_i for $n_i/(1 + n_i\sigma_a^2/\sigma_e^2)$ and in this way $\mathrm{var}(\hat{\sigma}_a^2)$ can be written as

$$\mathrm{var}(\hat{\sigma}_a^2) = \frac{2\sigma_e^4}{n_o^2} \left\{ \frac{1}{(c-1)^2} \left[\left(\frac{1}{N} \Sigma \frac{n_i^2}{w_i} \right)^2 + \Sigma \frac{n_i^2}{w_i^2} - \frac{2}{N} \Sigma \frac{n_i^3}{w_i^2} \right] + \frac{1}{N-c} \right\} .$$

Unfortunately Crump (1951) omits the 1/N from the first term. The equivalence of the above expression to (5) is readily shown. Summation is with respect to i, for i = 1,2,...,c.

Unbalancedness

Data wherein all groups have the same number of observations are called balanced; and those where the groups have different numbers of observations are called unbalanced. The characteristic of unbalancedness is not, however, a dichotomy, for it can be evident in varying degrees. For example, with five groups and 25 observations balanced data have n_i-values of 5, 5, 5, 5, and 5; moderately unbalanced data might be considered as those having n_i values of 1, 1, 3, 10 and 10; and severely unbalanced data would be those with n_i-values of 1, 1, 1, 1 and 21. Sets of n_i-values such as these will be referred to as n-patterns and the effects of unbalancedness will be studied by considering a variety of such patterns. There is, of course, no end to the number of possible n-patterns that could be used for this purpose so, in order to confine the problem, we limit ourselves in this paper almost entirely to n-patterns of 5 groups having a total of 25 observations ($c = 5$, $N = 25$). Some simple variants thereof are also used, as shown in Table 1. The nine n=patterns shown there represent unbalancedness of quite widely differing degrees and provide opportunity for comparisons. Pattern P_1 is the balanced case; patterns P_2 and P_3 are moderately unbalanced and P_4 and

(Show Table 1)

P_5 are seriously unbalanced. P_6 is merely P_5 with 20 observations added to one of the single-observation groups; and P_7, P_8 and P_9 are just P_4, P_5 and P_6 with five times as many observations per group.

The descriptions "moderately" and "seriously" unbalanced used in Table 1 are adopted on empirical grounds: no quantitative measure of the degree of unbalancedness is specified, although one might seek a statistic of unbalancedness based, presumably, on the n_i-values in an n-pattern. An obvious possibility is the variance of the n_i's,

$$v = v(n_i) = \frac{\Sigma n_i^2 - (\Sigma n_i)^2/c}{c - 1} = \frac{S_2 - N^2/c}{c - 1},$$

or, alternately, the ratio of v to its maximum value for given N and c, this maximum being achieved when c - 1 groups have 1 observation each and one has N - c + 1 observations, so giving

$$v_{max} = \frac{c - 1 + (N - c + 1)^2 - N^2/c}{c - 1} = \frac{(N - c)^2}{c}.$$

Then $v/v_{max} = cv/(N - c)^2$ could be suggested as a statistic of unbalancedness. However, as shall be indicated, neither this nor any other statistic based solely on n_i-values is uniformly suitable for considering the effects of unbalancedness on distributional properties (the variance, for example) of $\hat{\sigma}_a^2$. This is so because of the way in which n-patterns of differing degrees of unbalancedness (P_2 and P_5, for example) can affect the distribution of $\hat{\sigma}_a^2$: the differences in their effects on the distribution vary according to the underlying value of σ_a^2. Thus the effect of unbalancedness is a function of both the n-pattern and σ_a^2 and so, in terms of the effects of unbalancedness on the distribution (variance) of $\hat{\sigma}_a^2$, it seems that a statistic for unbalancedness cannot be one based solely on the n-pattern.

In considering the effects of unbalancedness on the variance of $\hat{\sigma}_a^2$ one might look for the n-pattern which, for given N and c, maximizes $var(\hat{\sigma}_a^2)$ shown in (5). Anderson and Crump (1967) consider the problem of minimizing $var(\hat{\sigma}_a^2)$ when it is known that the n_i cannot all be the same (equal n_i minimizes it absolutely); but if, for a given N and c, the maximum of $var(\hat{\sigma}_a^2)$ could be found, then the value of $var(\hat{\sigma}_a^2)$ for the particular n-pattern at hand could be considered relative to the maximum, which would presumably represent the worst case of unbalancedness for that N and c.

It is clear from the form of (5) that its first two terms are maximum, regardless of σ_a^2 and σ_e^2, when S_2 is greatest, and this occurs when the n-pattern has the form 1,1,1,...,1,N - c + 1, examples of which are patterns P_5 and P_8 in Table 1. However, this kind of n-pattern does not maximize the third term in (5). The coefficient of $2\sigma_a^4$ in that term is

$$\frac{N^2S_2 + S_2^2 - 2NS_3}{(N^2 - S_2)^2} = \frac{2\sum\sum_{i<j} n_i^2 n_j^2 + \sum\sum\sum_{i<j<k} n_i n_j n_k (n_i + n_j + n_k)}{(\sum\sum_{i<j} n_i n_j)^2}$$

$$= \frac{2\sum\sum_{i<j} n_i^2 n_j^2 + \sum\sum\sum_{i<j<k} n_i n_j n_k (n_i + n_j + n_k)}{2\sum\sum_{i<j} n_i^2 n_j^2 + 4\sum\sum\sum_{i<j<k} n_i n_j n_k (n_i + n_j + n_k) + 12\sum\sum\sum\sum_{i<j<k<h} n_i n_j n_k n_h},$$

where, as usual, the limit of all summations is c, the number of groups. The value of this expression is clearly less than unity, but its complexity, as a function of the n_i's, appears to preclude ascertaining what n-patterns (if any) maximize it. The intractability is reduced somewhat when c = 3, for then the last term of the denominator does not exist, and the coefficient reduces to

$$1 - \frac{3N/2}{\frac{n_1 n_2}{n_3} + \frac{n_2 n_3}{n_1} + \frac{n_1 n_3}{n_2} + 2N}$$

which is maximum when the n-pattern is 1, (N - 1)/2, (N - 1)/2, provided N > 6. It was this n-pattern, extrapolated to more than 3 groups, which prompted the use of patterns P_4, P_7 and P_9 shown in Table 1. Patterns of this nature maximize, of course, only the third term in (5) and the whole expression will then be a maximum only for sufficiently large values of σ_a^2. Thus it appears that in general one cannot find n-patterns, for given N and c, which maximize $\text{var}(\hat{\sigma}_a^2)$ uniformly for all values of σ_e^2 and σ_a^2. This reaffirms the suggestion that the effects of unbalancedness are relative to the underlying values of σ_a^2 and σ_e^2 and not independent of them.

Although no n-pattern has been found that maximizes $\text{var}(\hat{\sigma}_a^2)$ uniformly for all σ_a^2 and σ_e^2, it is clear from (5) that

$$\text{var}\ \hat{\sigma}_a^2 \doteq \frac{2N^2(N - 1)(c - 1)}{(N - c)(N^2 - S_2)^2}\ \sigma_e^4 \qquad \text{for } \frac{\sigma_a^2}{\sigma_e^2} < 1 .$$

This is maximum when S_2 is, i.e. when $v(n_i)$ is, and so when $\sigma_a^2/\sigma_e^2 < 1$ a suitable statistic for unbalancedness might be $v(n_i)/v_{max} = cv(n_i)/(N - c)^2$, as suggested earlier. At the other end of the scale, when $\sigma_a^2/\sigma_e^2 > 1$, the approximate value of $\text{var}(\hat{\sigma}_a^2)$ is, from (5),

$$\text{var}(\sigma_a^2) \doteq 2U\sigma_a^4 \qquad \text{for } \sigma_a^2/\sigma_e^2 > 1 ,$$

where

$$U = \frac{N^2S_2 + S_2^2 - 2NS_3}{(N^2 - S_2)^2} .$$

Although no n-pattern has been found that maximizes this, we have seen that for $c = 3$ it is maximized, for $N > 6$, when the n-pattern is $[1, \frac{1}{2}(N - 1), \frac{1}{2}(N - 1)]$. Furthermore, it can be shown that increasing N without increasing c can increase U. For example, for the n-pattern P_5, (1, 1, 1, 1, 21), U = .398, for (1, 1, 1, 1, 41), U = .416, and for n-pattern P_6, (1, 1, 1, 21, 21), U = .699. Hence, when $\sigma_a^2/\sigma_e^2 > 1$, it is possible for $\text{var}(\hat{\sigma}_a^2)$ to be increased by the addition of observations to the data; in other words, increasing the amount of data can increase the variance of the estimator, somewhat of a paradoxical situation. It would seem that an implication of this result is that in situations when $\sigma_a^2/\sigma_e^2 > 1$

one should strive for having as many groups as possible in one's data, rather than numerous observations in each group. In this context it is also interesting to note that if each n_i is increasing by the same fraction, λ, say, then U is not altered, and so $\text{var}(\hat{\sigma}_a^2)$ will be affected very little, especially when $\sigma_a^2/\sigma_e^2 > 1$.

Calculated values of $\text{var}(\hat{\sigma}_a^2)$

In light of the above discussion the effects of unbalancedness on $\text{var}(\hat{\sigma}_a^2)$ have been studied by computing $\text{var}(\hat{\sigma}_a^2)$ for the n-patterns shown in Table 1, each with the series of values for σ_a^2 and σ_e^2 shown in the same table. At all times $\sigma_e^2 = 1$ has been used, in combination with each of the eleven values for σ_a^2 seen in Table 1. Each combination has then been used with each n-pattern to calculate $\text{var}(\hat{\sigma}_a^2)$ from (5). The results are shown in Table 2. As would be expected, for

(Show Table 2)

each n-pattern this variance increases as σ_a^2 increases; and, as indicated in footnotes to the table, other points of interest are also evident.

(1) In n-patterns P_1 through P_5, N and c remain constant, and the largest value of $\text{var}(\hat{\sigma}_a^2)$ for given σ_a^2 is either in P_4, (1, 1, 1, 11, 11), or in P_5, (1, 1, 1, 1, 21). This suggests that for given N, c and σ_a^2 the largest value of $\text{var}(\hat{\sigma}_a^2)$ may be when the n-pattern is of the form (1, 1, ..., k) or (1, 1, 1, ..., q, q) where $k = N - c + 1$ and $q = \frac{1}{2}(N - c + 2)$.

(2) If, for given N and c, the n-pattern giving largest $\text{var}(\hat{\sigma}_a^2)$ is of the form (1, 1, ..., 1, q, q) then the pattern (1, 1, ..., 1, k) does not necessarily give the next largest. For example, with $\sigma_a^2 = 10$ pattern P_4 gives $\text{var}(\hat{\sigma}_a^2) = 113$ and P_2, not P_5, gives the next largest value of $\text{var}(\hat{\sigma}_a^2)$, namely 90.

(3) Increasing the total number of observations, N, can increase var($\hat{\sigma}_a^2$), as already discussed. Thus for $\sigma_a^2 \geq 1$, var($\hat{\sigma}_a^2$) is larger with P_6 than P_7; and for $\sigma_a^2 > \frac{1}{4}$, its values are larger with P_9 than with P_8.

(4) Increasing each n_i by the same proportion does not greatly decrease var($\hat{\sigma}_a^2$) when $\sigma_a^2/\sigma_e^2 > 1$. For example, the values in Table 2 for P_7, P_8 and P_9, for $\sigma_a^2 > 1$, are very little less than those for P_4, P_5 and P_6, although they contain five times as many observations.

Frequency Distributions of Simulated Components

As already indicated, the distribution of $\hat{\sigma}_a^2$ is not known explicitly in the case of unbalanced data, $\hat{\sigma}_a^2$ being a linear function of MSW which is a multiple of a χ^2-variable and of MSB which is a weighted sum of other 1-degree-of-freedom χ^2-variables. Empirical investigation of the distribution has therefore been made by means of computer simulation, using the n-patterns and the values of σ_e^2 and σ_a^2 shown in Table 1. Leone and Nelson (1966) report studies of this nature for balanced data in a 5 X 2 X 2 X 2 nested design, but few studies for unbalanced data, of a survey nature, have been made.

With each combination of n-pattern and σ_a^2-value that was used, 2,000 simulations were made of the estimator $\hat{\sigma}_a^2$. On each occasion $\hat{\sigma}_e^2$ was derived from a simulated χ_{N-c}^2 variate using procedures given in U. S. Steel (1962) for $N - c \leq 30$ and in Zelen and Severo (1964) for $N - c > 30$. In this way only group means and not individual observations had to be simulated. The means were derived by pseudo-random sampling of 1,000 abscissae that are medians of 1,000 equi-probable areas of the standardized normal distribution, using a multiplicative congruential generator for the pseudo-random sampling. (The medians of the 1,000 equi-probable areas of the normal distribution are discussed in Searle (1966): their first three even-order moments are .999, 2.965 and 14.266 respectively.) Thus when the randomly generated integer between 1 and 1,000 was r, the r'th median was chosen, m_r say, and for a group having n_i observations the group mean, $\bar{y}_{i.}$, was simulated as $\sigma_a m_r + 1/\sqrt{n_i}$. Then, using $\hat{\sigma}^2_e$ derived from the simulated χ^2_{N-c} variate, the simulated value of $\hat{\sigma}^2_a$ was, in accord with (1) and (2), computed as

$$\hat{\sigma}^2_a = \frac{N(c-1)(MSB - \hat{\sigma}^2_e)}{N^2 - S_2}$$

where

$$MSB = \frac{1}{N-c}\left[\sum_{i=1}^{c} n_i \bar{y}^2_{i.} - \left(\sum_{i=1}^{c} n_i \bar{y}_{i.}\right)^2 \Big/ N\right].$$

Frequency distributions were then made of these simulated $\hat{\sigma}^2_a$'s, grouping them into 53 intervals based on σ^2_a and the standard error of its estimate, namely $SE = \sqrt{var(\hat{\sigma}^2_a)}$ derived from (5). Fifty-one intervals of finite width (0.1)SE were used, with center points at σ^2_a - 2SE through to σ^2_a + 3SE, tail intervals being from $-\infty$ to σ^2_a - 2.05SE and from σ^2_a + 3.05 to $+\infty$. This choice resulted in barely 2% of the simulated $\hat{\sigma}^2_a$-values being in the tail intervals, and it was also convenient for computer generation of frequency polygons and cumulative frequencies. However, before discussing the results of these simulations other comments are in order.

Negative estimates

It is well known that negative estimates of σ^2_a can be derived by the methods being considered here, namely (2). The frequency with which negative estimates occurred among the simulated estimates is therefore of some interest. Indication of the extent of this occurrence is shown in Table 3, wherein is given the percentage of the 2,000 simulated values of $\hat{\sigma}^2_a$ that were negative in various combinations of n-pattern and σ^2_a values. It is clear that for situations in which σ^2_a is close to zero there may be many negative estimates $\hat{\sigma}^2_a$, but even when σ^2_a/σ^2_e

(Show Table 3)

is in the neighborhood of 0.25 to 0.50 there still seems to be an appreciable likelihood of getting a negative estimate. If this is indeed the case it gives credence to results often obtained by geneticists and others for whom the variance ratio is customarily in this range. Wang (1967) also reports the frequency of negative estimates, as do Leone and Nelson (1966) for their balanced design, in which case they were able to derive analytical expressions for the frequency, due to the χ^2 distributional properties of the mean squares. This is not so here, and we must be content with the empirical results given in Table 3.

Monte Carlo methods

The procedure described above for generating frequency distributions is purely one of simulation. It makes no use of an available (conditional) distribution property of $\hat{\sigma}^2_a$. This can be utilized in a method which we call Monte Carlo, distinct from the method already discussed, henceforth called the simulation method.

The between- and within-group sums of squares are, from (1), SSB = (c-1)MSB and SSW = (N-c)MSW respectively, and with these equation (2) can be written as

$$\hat{\sigma}^2_a = \lambda_1 SSB - \lambda_2 (SSW/\sigma^2_e) \quad --- (6)$$

for

$$\lambda_2 = \frac{N(c - 1)\sigma^2_e}{(N^2 - S_2)(N - c)} \quad --- (7)$$

and $\lambda_1 = N(N^2 - S_2)$. From (6) it is seen at once that the conditional variable $(\hat{\sigma}_a^2|SSB)$ has a χ^2-distribution (multiplied by a constant). Therefore, for any interval I_k on the real line, one can simulate SSB and thence calculate the probability $p_k = Pr\{\hat{\sigma}_a^2|SSB \in I_k\}$. On dividing the real line into n intervals I_k, k = 1,2,...,m, p_k can then be found for every interval for each simulated SSB, and averaging each p_k over a series of simulations would give an estimated probability density function of $\hat{\sigma}_a^2$. The apparent advantage of this procedure over the simulation method is that each simulated SSB contributes information to each of the I_k intervals, whereas in the simulation method each simulated $\hat{\sigma}_a^2$ contributes information to only one interval. Hopefully, for equivalent information about the whole curve, this should mean that the Monte Carlo procedure would require less simulations (of SSB) than would the simulation method (of $\hat{\sigma}_a^2$). Unfortunately this advantage does not always occur in practice. The difficulty is that when the interval I_k has length (0.1)SE say, as used in the simulation method, then p_k is the probability that a χ^2-variable lies in an interval of length $SE/10\lambda_2$ with λ_2 as in (7). And this interval can turn out to be so large that the probability content of five, or even fewer, adjacent intervals can be so close to 1.00 as to leave other intervals with near-zero probability. For example, in Table 1 with $\sigma_a^2 = 1$ and n-pattern (1,1,1,11,11), $var(\hat{\sigma}_a^2) = 1.41$, and so (0.1)SE = 0.118. But $\lambda_2 = 25(4)1/380(20) = 1/76$, so that the interval has length 76(0.118) = 8.7; and the 1% and 99% points respectively of the χ^2_{20} distribution are 8.26 and 37.57. Hence four adjacent intervals of length 8.7 include nearly all of the probability so that, in this case, a simulated SSB would be contributing non-zero information not to all the intervals but only to about four of them. Furthermore, computer time for calculating the probabilities p_k exceeds that of calculating additional $\hat{\sigma}_a^2$ values in the simulation method. Thus the apparent advantage of the Monte Carlo method does not materialize.

Wang's density function

Wang (1967) considers the variable

$$Z = \alpha\chi^2_{2n} - \beta\chi^2_{m} \qquad \text{- - - (8)}$$

where α and β are constants and χ^2_{2n} and χ^2_{2m} are two independent χ^2-variables with 2n and 2m degrees of freedom, respectively, n and m being integers. (This n is not the same as the n used earlier in discussing balanced data.) The density function is defined in two parts, for Z negative and for Z positive. With the constant

$$K = \frac{\alpha^{m-1}\,\beta^{n-1}}{2^{n+m}\,\Gamma(n)\,\Gamma(m)(\alpha+\beta)^{n+m-1}} \qquad \text{- - - (9)}$$

the part of the function for negative Z is

$$f_-(z) = Ke^{z/2\beta}\int_0^\infty e^{-\frac{1}{2}t}t^{n-1}\left[t - z\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)\right]^{m-1}dt \qquad \text{- - - (10)}$$

which, when m is integer, expands to

$$f_-(z) = \left(\frac{\beta}{\alpha+\beta}\right)^{n-1}\frac{1}{2(\alpha+\beta)}\,e^{z/2\beta}\sum_{j=0}^{m-1}\left(\frac{-z}{2\beta}\right)^{m-1-j}\left(\frac{\alpha}{\alpha+\beta}\right)^{j}\frac{(n-1-j)!}{j!(m-1-j)!(n-1)} \qquad (11)$$

as given by Wang (1967). And that part of the function for positive Z is

$$f_+(z) = Ke^{-z/2\alpha}\int_0^\infty e^{-\frac{1}{2}t}t^{m-1}\left[t + z\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)\right]^{n-1}dt\ , \qquad \text{- - - (12)}$$

which Wang also writes as a sum, analogous to (11), when n is integer.

This distribution cannot be used directly on $\hat{\sigma}^2_a$, because it is not a weighted difference between two independent χ^2's, as is Z. But let us approximate by

$$\hat{\sigma}_a^2 = \alpha\chi_q^2 - \lambda_2\chi_{N-c}^2 \ , \qquad --- (13)$$

where χ_{N-c}^2 corresponds to SSW/σ_e^2 in (6), λ_2 comes from (7), and where we assume that χ_q^2 is independent of χ_{N-c}^2. Then (13) is the same form as Z and equating the first two moments of both sides of (13) yields values of α and q that can be used, by means of (13) and (8), to derive an approximate distribution of $\hat{\sigma}_a^2$. This procedure is that alluded to in Wang (1967). Equating the moments of (13) gives

$$E(\hat{\sigma}_a^2) = \sigma_a^2 = \alpha q - \lambda_2(N - c)$$

and

$$\text{var}(\hat{\sigma}_a^2) = 2\alpha^2 q + 2\lambda_2^2(N - c)$$

which can be solved for α and q by using (5) for $\text{var}(\hat{\sigma}_a^2)$ and (7) for λ_2. Writing $\tau \equiv \sigma_a^2/\sigma_e^2$ the solutions are

$$\alpha = \frac{\sigma_e^2[N^2(c - 1) + 2N(N^2 - S_2)\tau + (N^2S_2 + S_2^2 - 2NS_3)\tau^2]}{(N^2 - S_2)[(N^2 - S_2)\tau + N(c - 1)]} \qquad -- (14)$$

and

$$q = \frac{(c - 1)[N^2(c - 1) + 2N(N^2 - S_2)\tau + \dfrac{(N^2 - S_2)^2}{c - 1}\tau^2]}{[N^2(c - 1) + 2N(N^2 - S_2)\tau + (N^2S_2 + S_2^2 - 2NS_3)\tau^2]} \ . \qquad -- (15)$$

For balanced data these expressions reduce to $\alpha = (\sigma_a^2 + \sigma_e^2/n)/(c - 1)$ and $q = c - 1$, as one would expect from equation (4). In addition, when $\sigma_a^2 = 0$, the value of q is always $c - 1$, and α is then $N\sigma_e^2/(N^2 - S_2)$. The approximate degrees of freedom q is, from (15), a multiple of $c - 1$, the degrees of freedom in the balanced case. Furthermore, when σ_a^2 is large, relative to σ_e^2, we have

$$\lim_{\tau \to \infty} q = \frac{(N^2 - S_2)^2}{N^2S_2 + S_2^2 - 2NS_3} \ , \qquad --- (16)$$

the inverse of the coefficient of $2\sigma_a^4$ in $\mathrm{var}(\sigma_a^2)$ considered earlier. We return to this expression later.

Computed values of α and q for the n-patterns and σ_a^2 values of Table 1 are shown in Table 4. Apparent trends are that for each n-pattern α increases and q decreases as σ_a^2 increases; and no value of q exceeds c - 1. Because of the limited extent of this table it would be unwise to speculate on the significance of

(Show Table 4)

trends in α and/or q in terms of unbalancedness, although a more extensive tabulation might lead to developingment of an index of unbalancedness. Clearly it will depend on σ_a^2.

Graphs of simulations and Wang-type approximations

Using α from (14), $n = \frac{1}{2}q$ from (15), $\beta = \lambda_2$ from (7) and $m = \frac{1}{2}(N - c)$, equations (11) and (12) now provide the approximate distribution of $\hat{\sigma}_a^2$. For all cases in which frequency polygons of simulated values of $\hat{\sigma}_a^2$ were obtained, so also was this approximation, computed in each case plotting the two curves alongside one another. The procedures used for computing (11) and (12) are outlined in the Appendix, and the results are shown in Figures I - VI.

The Figures show the computer output as obtained. Headings to each figure show: the n-pattern; the value of σ_a^2 used in the simulation, denoted by A (A = 0.25, for example, in Figure 1); the value of $\mathrm{var}(\hat{\sigma}_a^2)$ calculated from (5) and denoted by var(A); and in most cases the sample mean and variance of the 2,000 values of $\hat{\sigma}_a^2$ obtained from the simulation.

To facilitate computer generation, frequencies have been measured on the horizontal axis, rather than the vertical, with the intervals measured on the vertical axis. The values of the center points, from σ_a^2 - 2SE to σ_a^2 + 3SE, of the 51 finite intervals of width 0.1(SE) are shown in the left-hand columns of the graphs, with an additional column indicating the tail intervals and the position of A, A ± 1.00(SE), A ± 2.00(SE) and A + 3.00(SE). Thus in Figure 1, A = .25, var(A) = .11 (shown as .105 in Table 2), SE = $\sqrt{.105}$ = .32 and A - 1.00(SE) = -0.07.

The full width of the horizontal axis represents a frequency of 0.19, this and the zero frequency points being indicated by arrows and the notations FREQ = 0 and FREQ = .19 in the headings. In all figures except III there is also a right-hand column showing the cumulative frequency of the simulated $\hat{\sigma}_a^2$ values. This and the 0.19 frequency indicator are not shown in Figure III. In all cases the frequency polygon of the simulated values is plotted with an X and that for the approximate density function is plotted with *, the latter being used whenever the two values in an interval coincide. For the approximate density function, values were calculated from (11) and (12) and multiplied by (5.1)SE/51 = 0.1(SE) to put the function on the same scale as that of the simulated values.

Figures I - V pertain largely to n-patterns with N = 25 and c = 5, patterns P_1 through P_5 of Table 1. With one exception, other n-patterns of that Table have not been used but instead, Figure VI shows a much more unbalanced n-pattern, where N = 60 and c = 20 with 19 groups having one observation and one having 41. Comments on the individual figures follow.

Figure I. This is the balanced case, (5,5,5,5,5), with $\sigma_a^2 = \frac{1}{4}$, a relatively small value. The two curves (the frequency polygon of simulated values and the approximate density function) appear to be quite similar and not unlike a χ^2 curve.

Figure II. This is the same balanced case, (5,5,5,5,5), as in Figure I but with a much larger value for σ_a^2, namely $\sigma_a^2 = 20$. The curves are quite similar to those of Figure I except for being a little steeper at values less than A - 1.00(SE).

Figure III. Six pairs of curves are shown here for a moderately unbalanced situation, (1,1,7,8,8), over a range of values for σ_a^2, namely $\frac{1}{4}$, 1, 2, 5, 10 and 20. The curves are still somewhat like a χ^2, with the steepness on the negative side increasing for the larger values of σ_a^2 (e.g. section 6 where $\sigma_a^2 = 20$, compared to section 2 where $\sigma_a^2 = 1$). There also appears to be a tendency for the curves of the simulated values to be slightly 'squeezed' compared to those of the approximate theoretical densities (sections 4 and 5, for example), the simulated curves having "higher and steeper peaks" than those of the density curves.

Figure IV. The four graphs here are for a very unbalanced case, (1,1,1,11,11), with $\sigma_a^2 = \frac{1}{4}$, 1, 5 and 20. The increasing steepness as σ_a^2 increases is now quite noticeable, the curves for $\sigma_a^2 = 5$ and 20 being almost exponential in type, corresponding to the approximate degrees of freedom, q, being close to unity, 1.91 and 1.83 respectively, from Table 4.

Figure V. The four cases shown in this figure illustrate effects of increasing unbalancedness when keeping σ_a^2 constant, equal to unity: (5,5,5,5,5), the balanced case; (1,1,3,10,10), moderate unbalancedness; (1,1,1,1,21), very unbalanced, these three all having 5 groups and 25 observations; and in section 4, (1,1,1,21,21), representing both severe unbalancedness and the addition of 20 observations compared to (1,1,1,1,21) shown in 3. The trend for these cases seems clear: as unbalancedness increases there is increasing steepness on the left, with (1,1,1,21,21) being like an exponential, again corresponding to $q < 2$, in this case $q = 1.41$. (Table 4)

Figure VI. This shows the very unbalanced case of (1,1,1,... for 19 groups, 41), for 60 observations in 20 groups, with four values of $\sigma_a^2 = \frac{1}{4}$, 1, 5 and 20. Two points of interest can be noticed: (i) the simulated curves are 'squeezed' considerably compared to the theoretical approximations, to the point of the latter appearing to be quite a poor fit; (ii) the values of σ_a^2 are the same as those used with (1,1,1,11,11) in Figure IV where, for $\sigma_a^2 = 5$ and 20, the curves are exponential in type with a q-values (degrees of freedom) 1.91 and 1.83: but they are not exponential in Figure VI where the corresponding q-values are 8.35 and 5.11. Thus for large values of σ_a^2, $q < 2$ appears to indicate that the distribution of $\hat{\sigma}_a^2$ is exponential in nature, as would be expected from q being the degrees of freedom of a χ^2 variable. In this connection the limiting value of q for infinite σ_a^2, as given by (16), is 1.81 for the n-pattern of Figure IV, corresponding to the exponential-style frequencies seen there, and it is 4.90 for the n-pattern of Figure VI where the curves are χ^2 in character. This limiting value, be it noted, is not affected by any proportional change in the n_i's of an n-pattern; i.e. (16) remains unchanged if every n_i of an n-pattern is multiplied by the same constant.

Two final, minor comments can be made. In almost all cases 95% of the estimates $\hat{\sigma}_a^2$ lay in the interval σ_a^2 - 1.5(SE) to σ_a^2 + 2.0(SE), and in all cases the mode of the distributions was considerably less than the mean. This raises the question of the unbiasedness of the estimator $\hat{\sigma}_a^2$. It is, as in most estimation procedures, mean unbiased, a concept well-suited to estimation in fixed effects models where one customarily thinks of unbiasedness in terms of repeated sampling and averaging of several estimates of the same parameter. But this concept may not necessarily be appropriate in variance component models where, in the estimation procedure, one might seldom envisage repeated sampling, least of all with the same n-pattern. This being so, estimation might be more truly considered as a one-time-only procedure, and so an estimator that is modally unbiased might be more appropriate.

Appendix: Computing Wang's density function

We here outline the procedures used for computing (11) and (12) when α is in (14), $n = \frac{1}{2}q$ from (15), $\beta = \lambda_2$ of (7) and $m = \frac{1}{2}(N - c)$.

First, the part for negative Z, shown in (11): by suitable choice of N and c, $m = \frac{1}{2}(N - c)$ is an integer and so the sum in (11) is finite. Thus with

$$L = \tfrac{1}{2}\alpha^{m-1}\,\beta^{n-1}/(\alpha + \beta)^{m+n-1} \qquad \text{--- (A1)}$$

(11) can, on replacing $m - 1 - j$ by j, be written as

$$f_-(z) = Le^{z/2\beta}\sum_{j=0}^{m-1}\left(\frac{-z}{2\beta}\right)^j\left(\frac{\alpha + \beta}{\alpha}\right)^j \frac{(m + n - 2 - j)!}{j!(m - 1 - j)!(n - 1)!}\,. \qquad \text{--- (A2)}$$

From (10), we find on putting $z = 0$ and using (9) and (A1) that

$$f_-(0) = \frac{L\Gamma(m + n - 1)}{\Gamma(m)\,\Gamma(n)}\,.$$

Writing this as

$$v_o = \frac{L\Gamma(m + n - 1)}{\Gamma(m)\,\Gamma(n)} = \frac{L(m + n - 2)(m + n - 3)\ \ldots\ (n + 1)n}{(m - 1)(m - 2)\ \ldots .\ 2.\ 1.} \qquad \text{--- (A3)}$$

(A2) can be written

$$f_-(z) = e^{z/2\beta}\sum_{j=0}^{m-1} v_j\,, \qquad \text{--- (A4)}$$

the v_j's being a readily computable recurrent series

$$v_j = \left[\frac{-z(\alpha + \beta)}{2\alpha\beta}\right]\frac{(m - j)}{j(m + n - 1 - j)}\,v_{j-1} \text{ for } j \geq 1 \qquad \text{--- (A5)}$$

the initial term being v_o in (A3).

The other part of the density function is for positive Z: here we use (12) and cannot invoke binomial expansion because n is not integer and t takes all values from 0 to ∞. Instead we use the transformation

$$t = z\left(\frac{1}{\alpha} + \frac{1}{\beta}\right)\tan^2\theta$$

and find that $f_+(z)$ reduces to

$$f_+(z) = \frac{(\tfrac{1}{2}z)^{m+n-1}e^{-z/2\alpha}}{\alpha^n\beta^m\Gamma(n)\Gamma(m)}\int_0^{\pi/2}\exp\left[\frac{-z(\alpha+\beta)\tan^2\theta}{2\alpha\beta}\right]\left(\frac{\sin^{2m-1}\theta}{\cos^{2m+2n-1}\theta}\right)d\theta\ . \quad \text{- - - (A6)}$$

Calculation of $f_-(z)$ was thus achieved from (A4) using (A3) and (A5), and of $f_+(z)$ from (A6). No numerical problems arose in (A4) but some did occur with (A6): those encountered and overcome are detailed in Townsend (1967). In general, the integral in (A6) was computed by the trapezoidal rule with 200 intervals between 0 and $\pi/2$, except in some cases where the integrand was effectively zero beyond some value θ, θ' say, appreciably less than $\pi/2$. In such cases the trapezoidal rule with 150 intervals between 0 and θ' was used. For a wider range of n-patterns than those of Table 1 the Romberg integration algorithm of Bauer (1961) would probably be more suitable.

Acknowledgements

Grateful thanks are given to D. A. Evans, D. S. Robson, E. C. Townsend, and Y. Y. Wang for their many and varied contributions to this work.

References

Anderson, R. L. and Crump, P. P. (1967). Comparisons of designs and estimation procedures for estimating parameters in a two-stage nested process. Technometrics 9, 499-516.

Bainbridge, T. R. (1963). Staggered, nested designs for estimating variance components. ASQC Annual Convention Transactions, 93-103.

Bauer, F. L. (1961). Algorithm 60, Romberg Integration. Comm. Assoc. Comput. Mach. 4, 255.

Bush, N. and Anderson, R. L. (1963). A comparison of three different procedures for estimating variance components. Technometrics 5, 421-440.

Eisenhart, C. (1947). The assumptions underlying the analysis of variance. Biometrics 3, 1.

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

Herbach, L. H. (1959). Properties of Model-II type analysis of variance tests. Ann. Math. Stat. 30, 939-959.

Leone, Fred C. and Nelson, Lloyd S. (1966). Sampling distributions of variance components. I. Empirical studies of balanced nested designs. Technometrics 8, 457-468.

Robinson, J. (1965). The distribution of a general quadratic form in normal variables. Austral. J. of Statistics, 7, 110-114.

Robson, D. S. (1965). On the admissability of non-negative maximum likelihood estimators of the "between-groups" variance component. Paper No. BU-179-M in the mimeo series, Biometrics Unit, Cornell University, Ithaca, N.Y.

Searle, S. R. (1956). Matrix methods in variance and covariance components analysis. Ann. Math. Stat. 27, 737.

Searle, S. R. (1966). Properties of certain discrete distributions for generating approximately normal variables. Paper No. BU-228-M, mimeo series, Biometrics Unit, Cornell University, Ithaca, N.Y.

Thompson, W. A. (1962). The problem of negative estimates of variance components. Ann. Math. Stat. 33, 273-289.

Townsend, E. C. (1967). Numerical problems encountered when calculating a frequency polygon for variance component estimates. Paper No. BU-248-M in the mimeo series, Biometrics Unit, Cornell University, Ithaca, N.Y.

U. S. Steel (1962). The generation of random samples from common statistical distributions. Technical Report 25, 17-016, U. S. Steel, Monroeville, Pa.

Wang, Y. Y. (1967). A comparison of several variance component estimators. Biometrika, 54, 301-305.

Zelen, M., and Severo, N. C. (1964). Probability functions. Chapter 26, "Handbook of Mathematical Functions", Nat. Bur. of Standards.

Table 1. n-patterns and population variances used in simulation studies

n-pattern

Pattern Number	Numbers of observations					
P_1	5	5	5	5	5	(balanced)
P_2	1	1	3	10	10	(moderately unbalanced)
P_3	1	1	7	8	8	
P_4	1	1	1	11	11	(seriously unbalanced)
P_5	1	1	1	1	21	
P_6	1	1	1	21	21	(for comparisons)
P_7	5	5	5	55	55	
P_8	5	5	5	5	105	
P_9	5	5	5	105	105	

Variances

σ_e^2 :	1										
σ_a^2 :	0	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	1	2	3	4	5	10	20

Table 2. Values of var($\hat{\sigma}_a^2$) for 9 n-patterns and 11 sets of σ_a^2, with $\sigma_e^2 = 1$.

	n-pattern	Values of σ_a^2: 0	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	1	2	3	4	5	10	20
P_1	5 5 5 5 5	.024	.105	.25	.46	.72	2.4	5.1	9	14	52	204
P_2	1 1 3 10 10	.035	.150	.37	.71	1.15	4.0	8.6	15	23	90†	354
P_3	1 1 7 8 8	.030	.134	.33	.62	1.01	3.5	7.5	13	20	78	308
P_4	1 1 1 11 11	.042	.177	.45	.86	1.41	5.0*	10.8	19	29	113†	448
P_5	1 1 1 1 21	.185	.374	.66	1.05*	1.53*	4.5	9.0	15	23	85	330
P_6	1 1 1 21 21	.014	.141	.44	.92	1.56⊗	5.9	13.1	23	36	141	562‡
P_7	5 5 5 55 55	.001	.084	.30	.66	1.16	4.5	10.1	18	28	112	444
P_8	5 5 5 5 105	.006	.084	.26	.54	.91	3.4	7.5	13	20	81	320
P_9	5 5 5 105 105	.001	.096⊗	.37	.81	1.43	5.7	12.7	22	35	140	560‡

* For given N and σ_a^2, var($\hat{\sigma}_a^2$) is usually maximum for n-pattern (1, 1, 1, 1, k) or (1, 1, 1, k, k).

† If n-pattern (1, 1, 1, k, k) gives maximum var($\hat{\sigma}_a^2$), pattern (1, 1, 1, 1, k) does not necessarily give the next largest value.

⊗ Increasing N can increase var($\hat{\sigma}_a^2$).

‡ Increasing every n_i by the same proportion does not greatly decrease var($\hat{\sigma}_a^2$), for $\sigma_a^2 \geq 1$.

Table 3. Percentage of negative estimates of σ_a^2 in 2,000 simulations.

n-pattern						Value of σ_a^2						
						0	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	1	2	3
									Per Cent			
P_1	5	5	5	5	5	56%	23	14	7	5	4	1
P_2	1	1	3	10	10	56	31	19	13	13	6	2
P_3	1	1	7	8	8	58	29	20	15	12	6	3
P_4	1	1	1	11	11	59	32	23	18	14	7	6
P_5	1	1	1	1	21	59	40	31	29	23	15	8
P_6	1	1	1	21	21	60	29	19	17	15	8	8
P_9	5	5	5	5	105	63	17	10	6	6	3	1

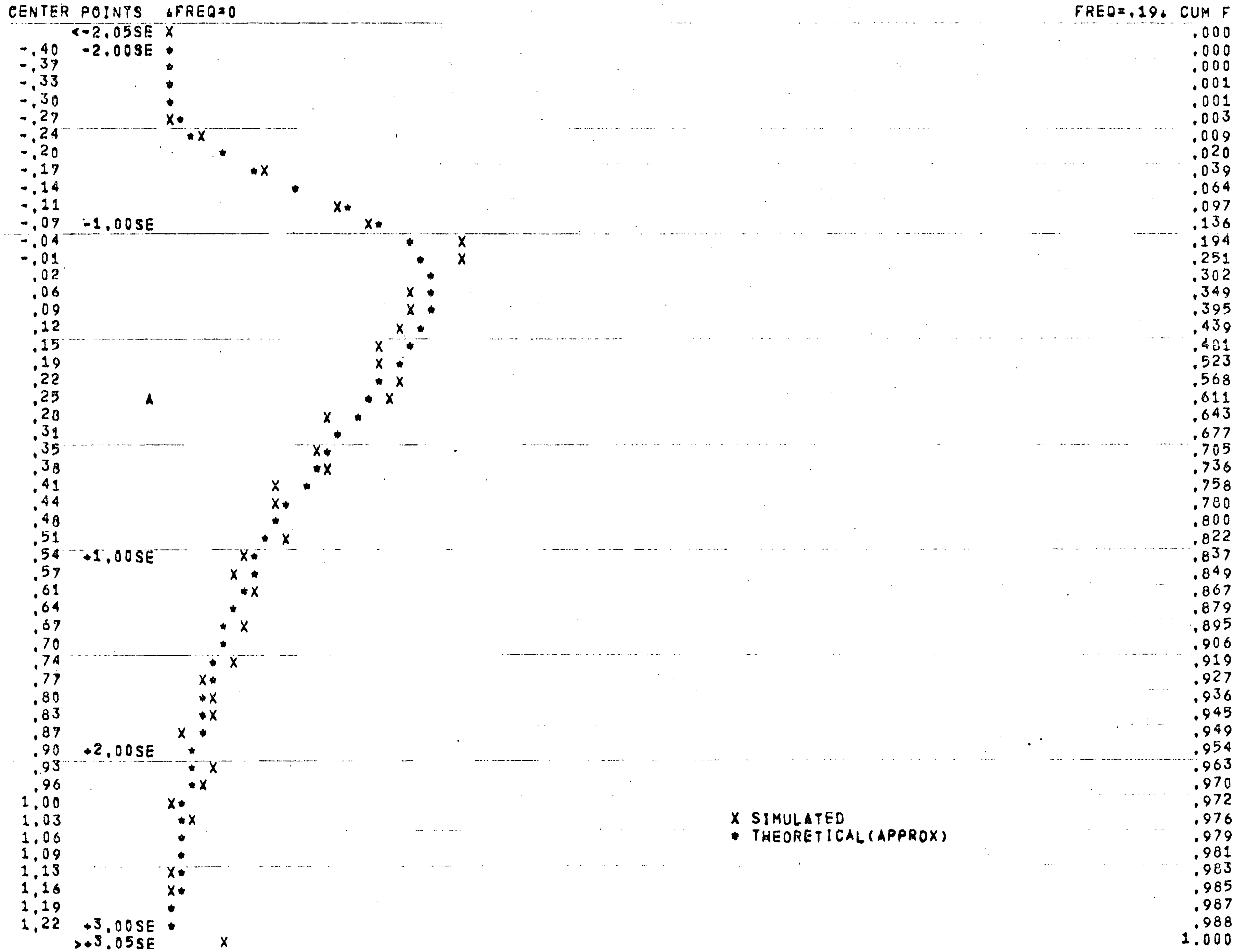
Table 4. Values of α and q in the approximation $\hat{\sigma}_a^2 = \alpha\chi_q^2 - \lambda_2\chi_{N-c}^2$ obtained by fitting the first two moments of $\hat{\sigma}_a^2$ [see equations (12) and (13)].

n-pattern		Values of σ_a^2										
		0	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	1	2	3	4	5	10	20
		The multiplier α:										
P_1	5 5 5 5 5	.05	.11	.18	.24	.30	.55	.80	1.05	1.30	2.55	4.05
P_2	1 1 3 10 10	.06	.15	.25	.35	.46	.89	1.33	1.77	2.20	4.39	8.76
P_3	1 1 7 8 8	.06	.14	.23	.31	.41	.79	1.17	1.55	1.93	3.82	7.62
P_4	1 1 1 11 11	.07	.17	.29	.42	.56	1.10	1.65	2.20	2.76	5.52	11.05
P_5	1 1 1 1 21	.14	.21	.30	.39	.48	.87	1.26	1.66	2.06	4.04	8.02
P_6	1 1 1 21 21	.04	.17	.34	.51	.68	1.37	2.07	2.77	3.46	6.96	13.94
P_7	5 5 5 55 55	.01	.14	.28	.41	.55	1.10	1.66	2.21	2.76	5.53	11.06
P_8	5 5 5 5 105	.03	.12	.21	.31	.41	.81	1.20	1.60	2.00	3.99	7.97
P_9	5 5 5 105 105	.01	.17	.34	.51	.69	1.39	2.09	2.78	3.49	6.98	13.97
		Degrees of freedom q:										
P_1	5 5 5 5 5	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00
P_2	1 1 3 10 10	4.00	3.35	2.98	2.80	2.69	2.50	2.43	2.40	2.38	2.33	2.31
P_3	1 1 7 8 8	4.00	3.50	3.20	3.06	2.97	2.82	2.76	2.73	2.71	2.67	2.66
P_4	1 1 1 11 11	4.00	3.11	2.63	2.40	2.27	2.05	1.97	1.93	1.91	1.85	1.83
P_5	1 1 1 1 21	4.00	3.78	3.53	3.35	3.21	2.93	2.81	2.75	2.70	2.61	2.56
P_6	1 1 1 21 21	4.00	2.38	1.96	1.80	1.71	1.57	1.53	1.50	1.49	1.46	1.45
P_7	5 5 5 55 55	4.00	2.19	2.00	1.94	1.91	1.86	1.84	1.83	1.83	1.82	1.81
P_8	5 5 5 5 105	4.00	3.12	2.86	2.76	2.70	2.61	2.57	2.56	2.55	2.53	2.52
P_9	5 5 5 105 105	4.00	1.65	1.54	1.51	1.49	1.46	1.45	1.45	1.44	1.44	1.43

FREQUENCY DISTRIBUTION OF THE AMONG GROUPS VARIANCE COMPONENT ESTIMATED FROM A 1-WAY ANALYSIS OF VARIANCE

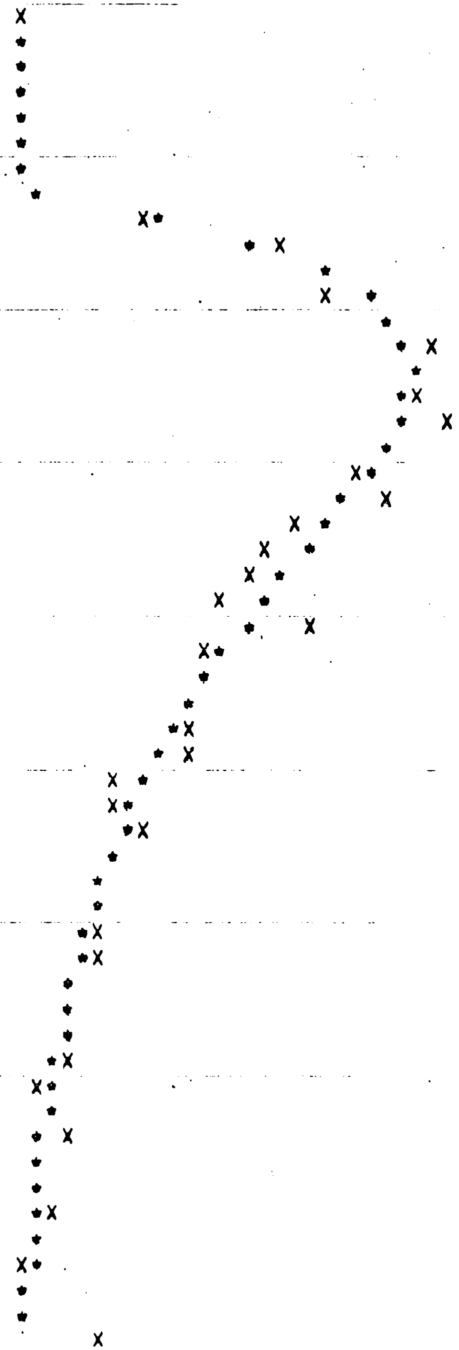
N-PATTERN; 5 5 5 5 5

POP. PARAMETERS; A= .25, VAR(A)= .11. 2000 SIMULATIONS YIELDED MEAN(A)= .251, VAR(A)= .106.



I

FREQUENCY DISTRIBUTION OF THE AMONG GROUPS VARIANCE COMPONENT ESTIMATED FROM A 1-WAY ANALYSIS OF VARIANCE
N-PATTERN: 5 5 5 5 5
POP. PARAMETERS: A= 20.00, VAR(A)= 204.02, 2000 SIMULATIONS YIELDED MEAN(A)= 19.992, VAR(A)= 196.952.



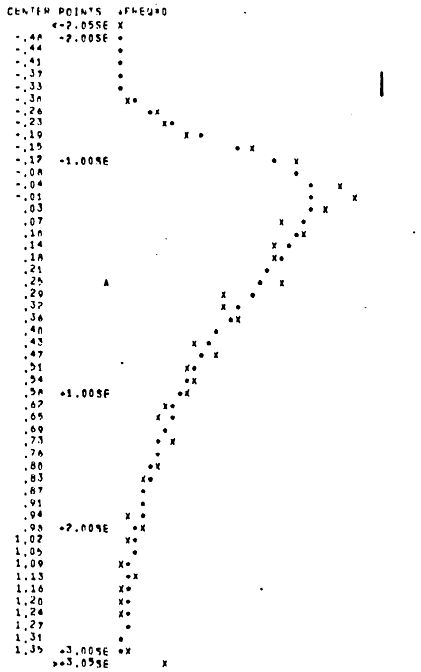
CENTER POINTS		CUM F
	<-2.05SE	0
-8.57	-2.00SE	0
-7.14		0
-5.71		0
-4.28		0
-2.85		0
-1.43		0
.00		.003
1.43		.019
2.86		.053
4.29		.093
5.72	-1.00SE	.134
7.14		.183
8.57		.239
10.00		.292
11.43		.344
12.86		.401
14.29		.450
15.71		.497
17.14		.546
18.57		.583
20.00	A	.616
21.43		.646
22.86		.674
24.29		.713
25.71		.737
27.14		.762
28.57		.784
30.00		.806
31.43		.829
32.86	+1.00SE	.842
34.28		.855
35.71		.872
37.14		.884
38.57		.894
40.00		.904
41.43		.916
42.85		.927
44.28		.934
45.71		.941
47.14		.948
48.57	+2.00SE	.955
50.00		.958
51.42		.963
52.85		.970
54.28		.973
55.71		.976
57.14		.981
58.57		.985
59.99		.986
61.42		.987
62.85	+3.00SE	.989
	>+3.05SE	1.000

↓FREQ=0 — FREQ=.19↓

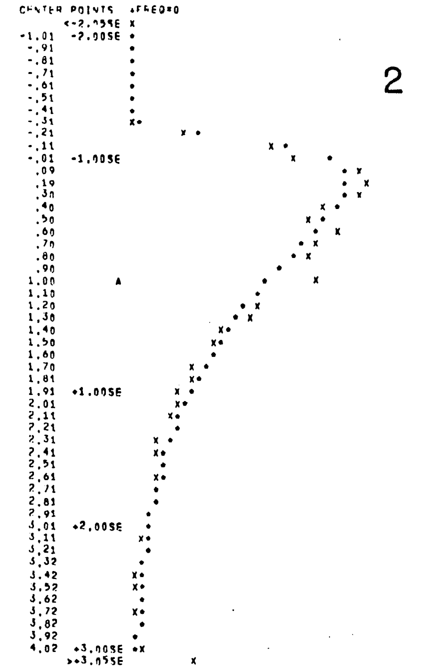
X SIMULATED
* THEORETICAL(APPROX)

III

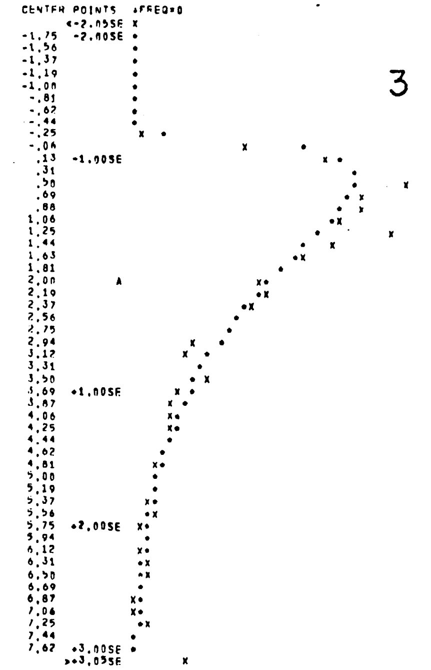
FREQUENCY DISTRIBUTION OF THE AMONG GROUPS VARIANC
N=PATTERN) 1 1 7 8 9
POP. PARAMETERS(A= .25,VAR(A)= .13. 2000 SIMULATIO S



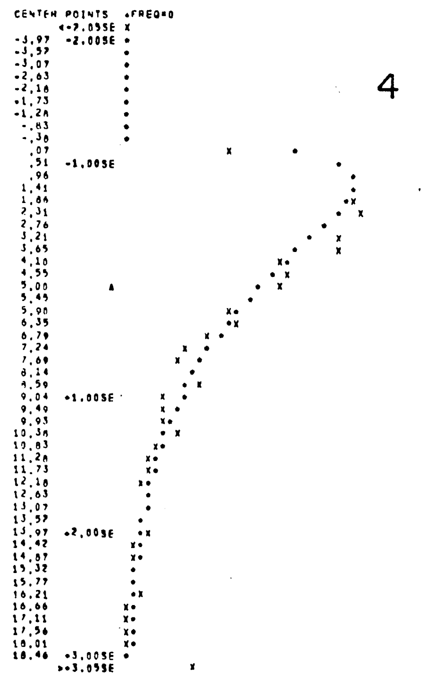
FREQUENCY DISTRIBUTION OF THE AMONG GROUPS VARIANC
N=PATTERN) 1 1 7 8 8
POP. PARAMETERS(A= 1.00,VAR(A)= 1.01. 2000 SIMULATIO S



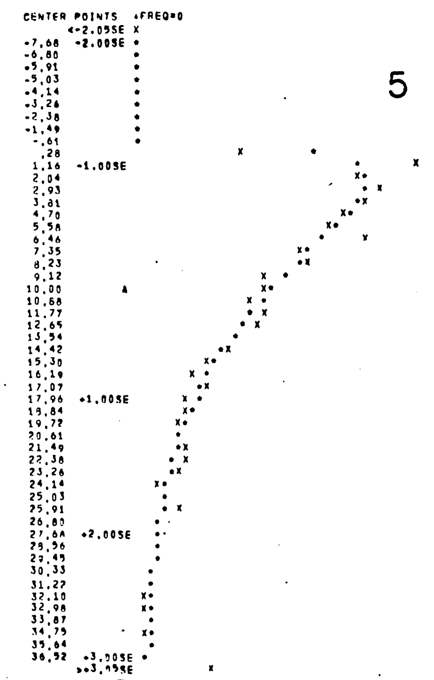
FREQUENCY DISTRIBUTION OF THE AMONG GROUPS VARIANC-
N=PATTERN) 1 1 7 8 8
POP. PARAMETERS(A= 2.00,VAR(A)= 3.91. 2000 SIMULATIO S



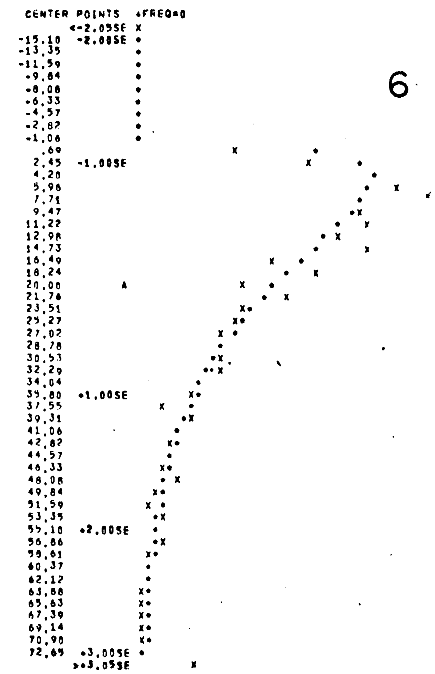
FREQUENCY DISTRIBUTION OF THE AMONG GROUPS VARIANC-
N=PATTERN) 1 1 7 8 8
POP. PARAMETERS(A= 5.00,VAR(A)= 20.12. 2000 SIMULATIO S



FREQUENCY DISTRIBUTION OF THE AMONG GROUPS VARIANC-
N=PATTERN) 1 1 7 8 8
POP. PARAMETERS(A= 10.00,VAR(A)= 78.15. 2000 SIMULATIO S

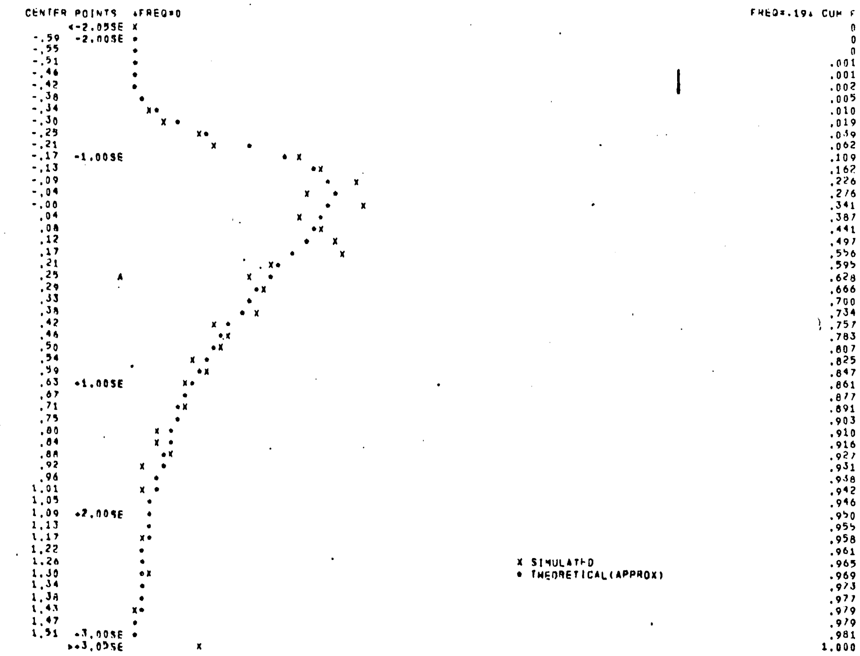


FREQUENCY DISTRIBUTION OF THE AMONG GROUPS VARIANC
N=PATTERN) 1 1 7 8 8
POP. PARAMETERS(A= 20.00,VAR(A)= 308.02. 2000 SIMULATIO S

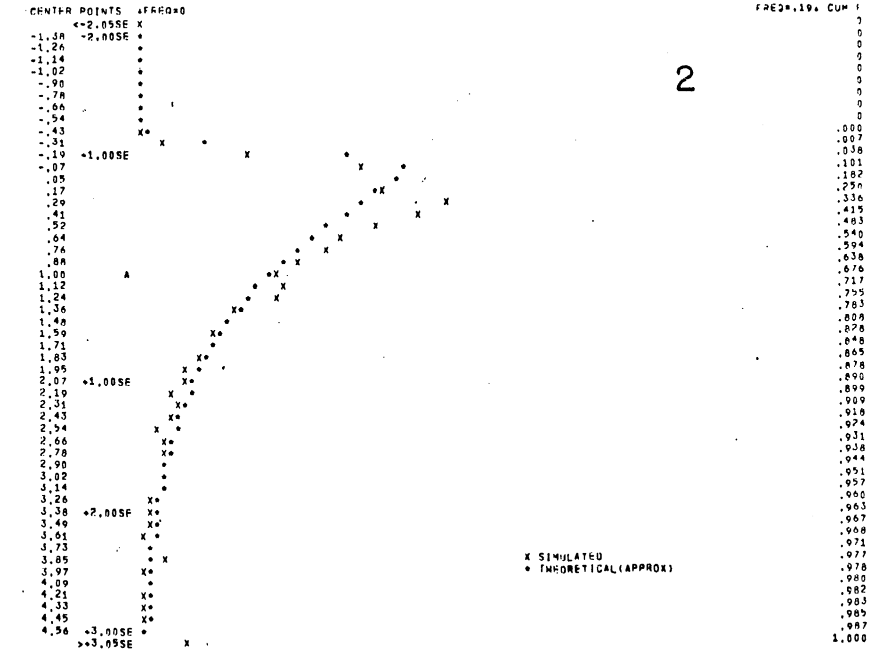


IV

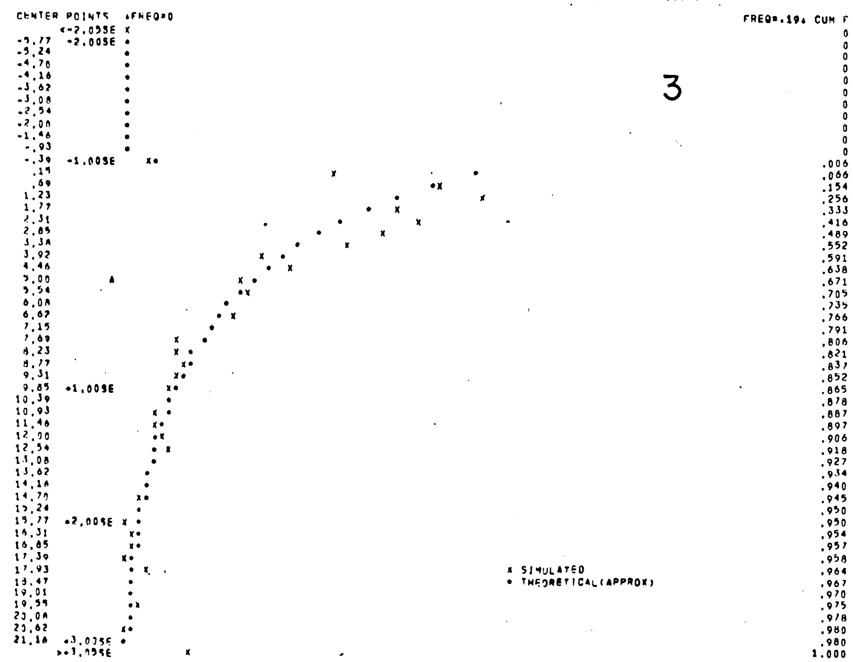
FREQUENCY DISTRIBUTION OF THE AMONG GROUPS VARIANC COMPONENT ESTIMATED FROM A 1-WAY ANALYSIS OF VARIANCE
N-PATTERN: 1 1 1 11 11
POP. PARAMETERS(A= .25,VAR(A)= .18. 2000 SIMULATIO S YIELDED MEAN(A)= .252,VAR(A)= .196.



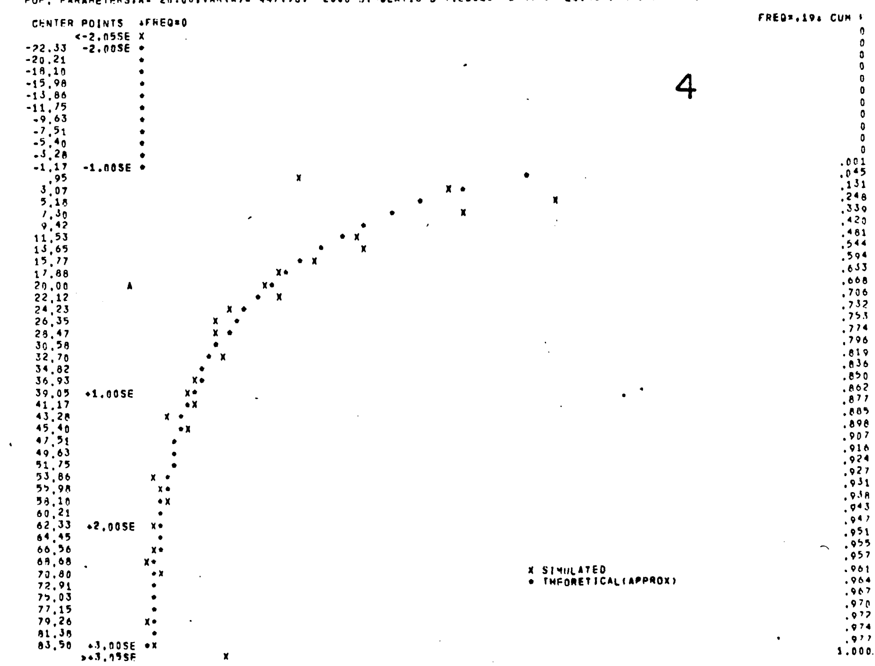
FREQUENCY DISTRIBUTION OF THE AMONG GROUPS VARIANC COMPONENT ESTIMATED FROM A 1-WAY ANALYSIS OF VARIANCE
N-PATTERN: 1 1 1 11 11
POP. PARAMETERS(A= 1.00,VAR(A)= 1.41. 2000 SIMULATIO S YIELDED MEAN(A)= .928,VAR(A)= 1.163.



FREQUENCY DISTRIBUTION OF THE AMONG GROUPS VARIANC COMPONENT ESTIMATED FROM A 1-WAY ANALYSIS OF VARIANCE
N-PATTERN: 1 1 1 11 11
POP. PARAMETERS(A= 5.00,VAR(A)= 29.02. 2000 SIMULATIO S YIELDED MEAN(A)= 5.038,VAR(A)= 30.368.

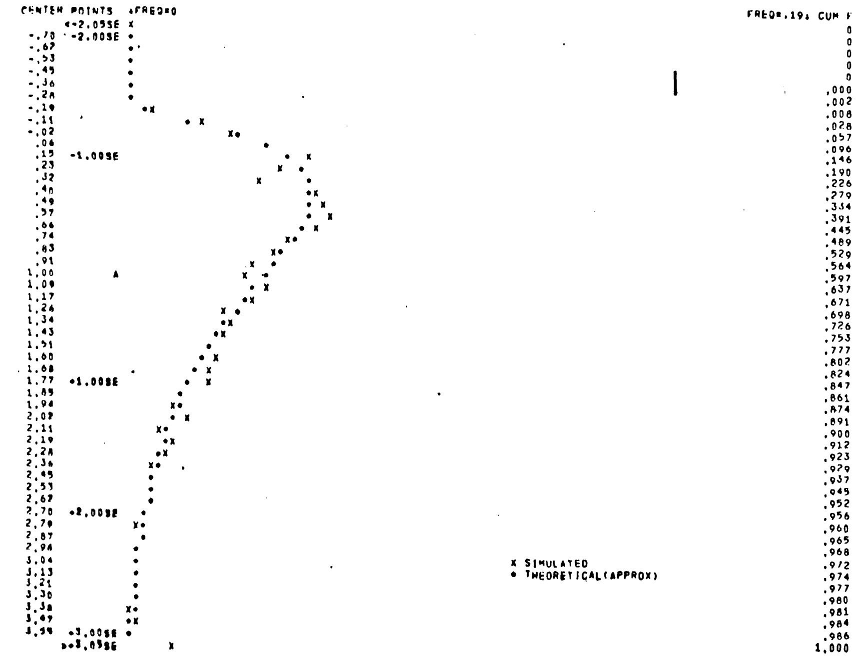


FREQUENCY DISTRIBUTION OF THE AMONG GROUPS VARIANC COMPONENT ESTIMATED FROM A 1-WAY ANALYSIS OF VARIANCE
N-PATTERN: 1 1 1 11 11
POP. PARAMETERS(A= 20.00,VAR(A)= 447.96. 2000 SIMULATIO S YIELDED MEAN(A)= 20.604,VAR(A)= 478.604.

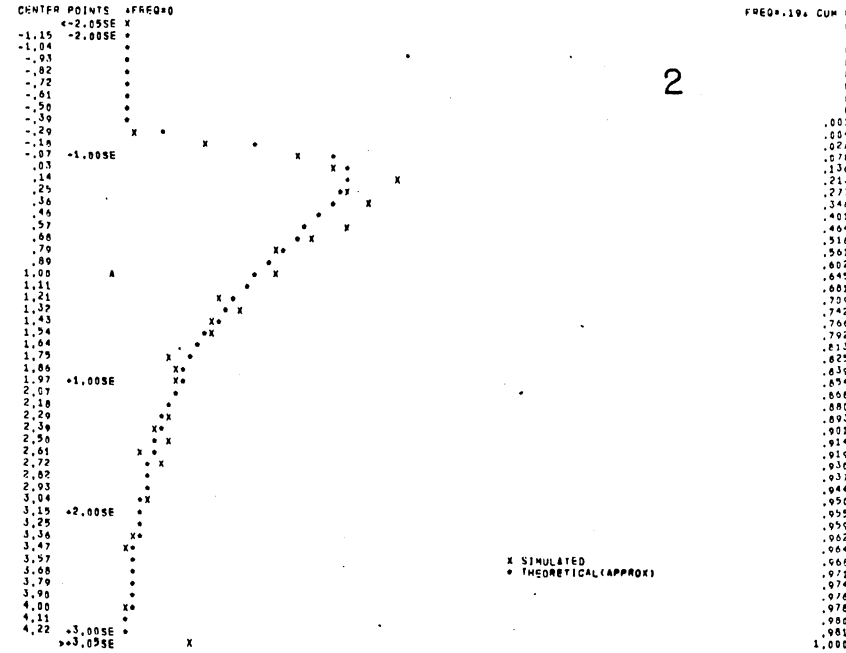


V

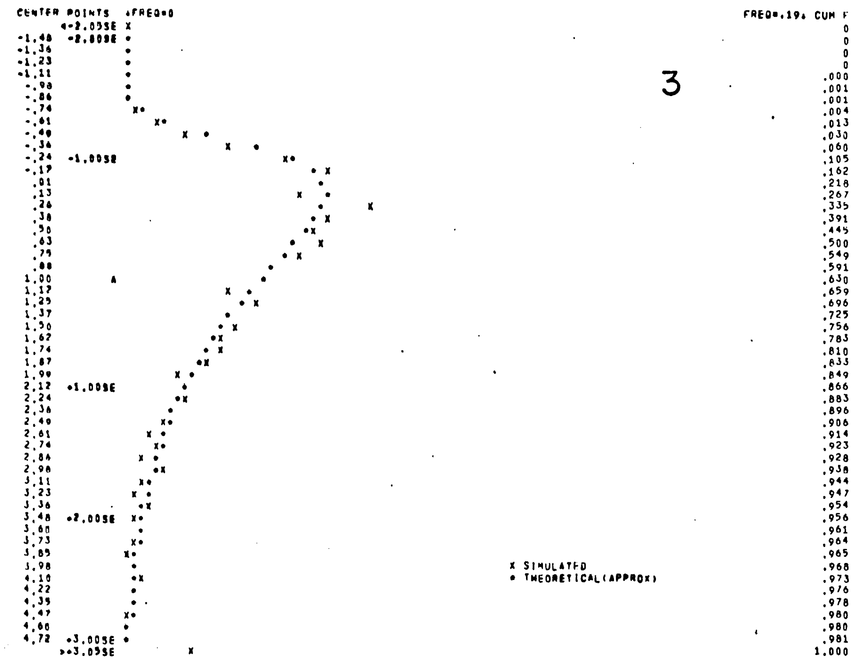
FREQUENCY DISTRIBUTION OF THE AMONG GROUPS VARIANC- COMPONENT ESTIMATED FROM A 1-WAY ANALYSIS OF VARIANCE
N-PATTERN: 5 5 5 5 5
POP. PARAMETERS:A= 1.00,VAR(A)= .72. 2000 SIMULATIO S YIELDED MEAN(A)= 1.012,VAR(A)= .730.



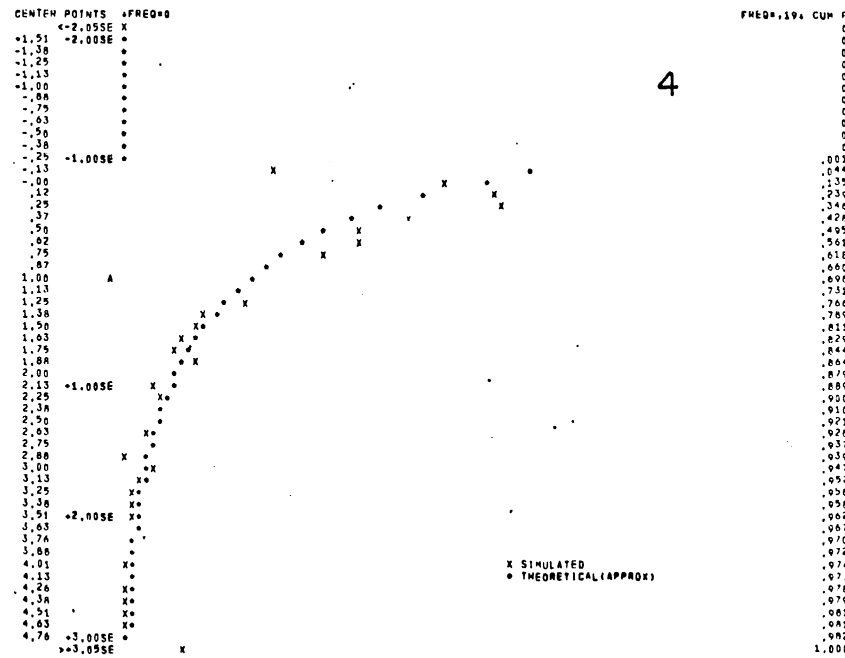
FREQUENCY DISTRIBUTION OF THE AMONG GROUPS VARIANC- COMPONENT ESTIMATED FROM A 1-WAY ANALYSIS OF VARIANCE
N-PATTERN: 1 1 3 10 10
POP. PARAMETERS:A= 1.00,VAR(A)= 1.15. 2000 SIMULATIO S YIELDED MEAN(A)= 1.009,VAR(A)= 1.171.



FREQUENCY DISTRIBUTION OF THE AMONG GROUPS VARIANC- COMPONENT ESTIMATED FROM A 1-WAY ANALYSIS OF VARIANCE
N-PATTERN: 1 1 1 1 21
POP. PARAMETERS:A= 1.00,VAR(A)= 1.54. 2000 SIMULATIO S YIELDED MEAN(A)= .995,VAR(A)= 1.531.



FREQUENCY DISTRIBUTION OF THE AMONG GROUPS VARIANC- COMPONENT ESTIMATED FROM A 1-WAY ANALYSIS OF VARIANCE
N-PATTERN: 1 1 1 21 21
POP. PARAMETERS:A= 1.00,VAR(A)= 1.57. 2000 SIMULATIO S YIELDED MEAN(A)= .935,VAR(A)= 1.336.

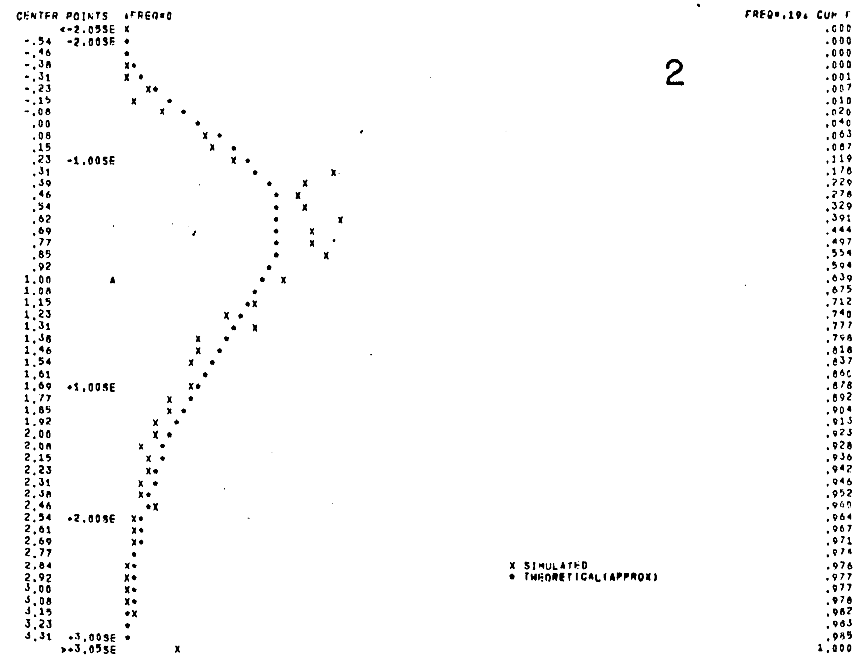


FREQUENCY DISTRIBUTION OF THE AMONG GROUPS VARIANC COMPONENT ESTIMATED FROM A 1-WAY ANALYSIS OF VARIANCE
N-PATTERN: 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 41

POP. PARAMETERS(A= .25,VAR(A)= .11. 2000 SIMULATIO S YIELDED MEAN(A)= .235,VAR(A)= .109.

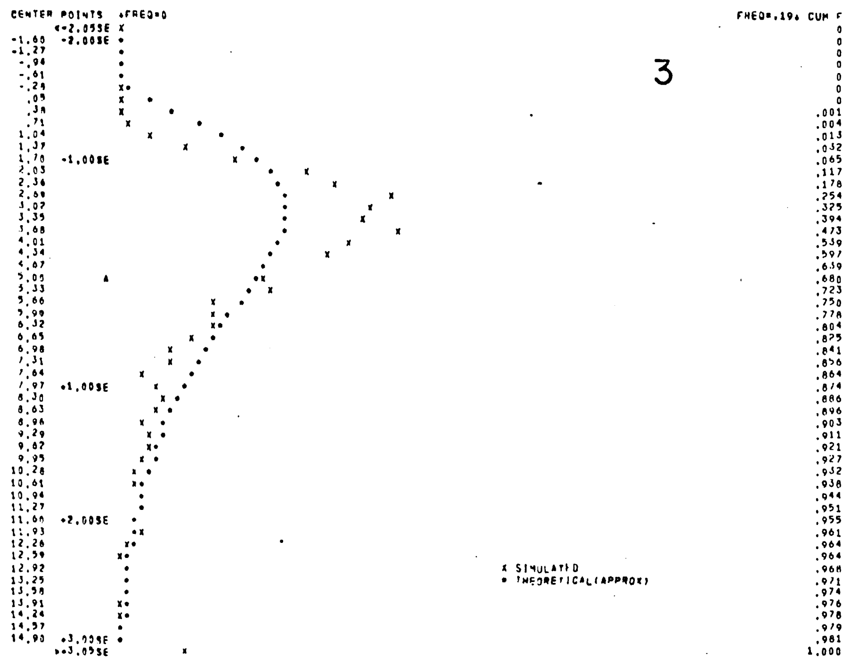
FREQUENCY DISTRIBUTION OF THE AMONG GROUPS VARIANC COMPONENT ESTIMATED FROM A 1-WAY ANALYSIS OF VARIANCE
N-PATTERN: 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 41

POP. PARAMETERS(A= 1.00,VAR(A)= .59. 2000 SIMULATIO S YIELDED MEAN(A)= .962,VAR(A)= .952.



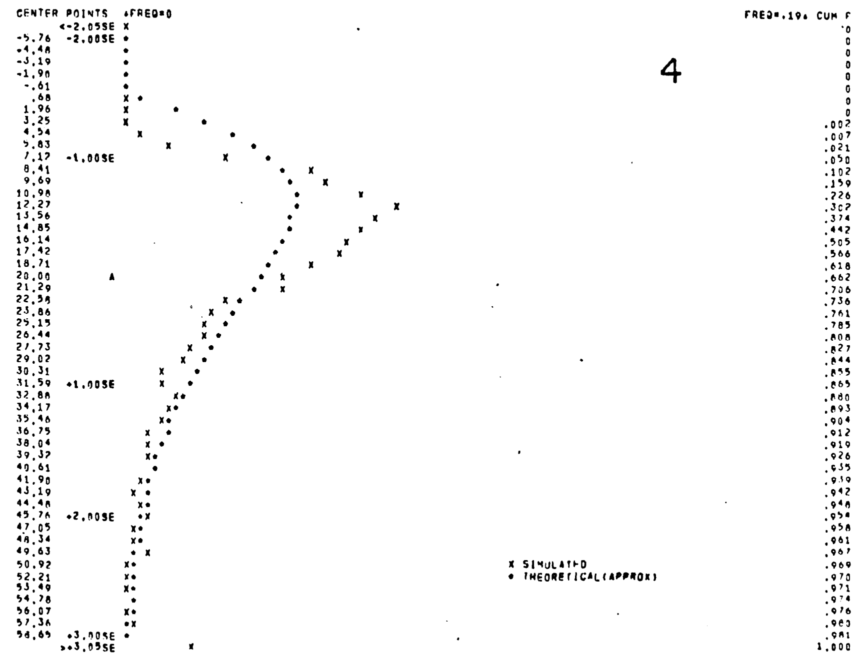
FREQUENCY DISTRIBUTION OF THE AMONG GROUPS VARIANC- COMPONENT ESTIMATED FROM A 1-WAY ANALYSIS OF VARIANCE
N-PATTERN: 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 41

POP. PARAMETERS(A= 5.00,VAR(A)= 10.90. 2000 SIMULATIO S YIELDED MEAN(A)= 4.907,VAR(A)= 10.931.



FREQUENCY DISTRIBUTION OF THE AMONG GROUPS VARIANC COMPONENT ESTIMATED FROM A 1-WAY ANALYSIS OF VARIANCE
N-PATTERN: 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 41

POP. PARAMETERS(A= 20.00,VAR(A)= 165.95. 2000 SIMULATIO S YIELDED MEAN(A)= 28.154,VAR(A)= 163.377.



VI

Appendix

Inverse of V

Since $\underline{V} = \sum_{i=1}^{a+} \underline{V}_i$, as in (19) it is clear from the nature of a direct sum that

$$\underline{V}^{-1} = \sum_{i=1}^{a+} \underline{V}_i^{-1} .$$

For this, $\underline{V}_i^{-1}$ is derived from the following theorem taken from Urquhart (1962).

Theorem Consider a matrix $\underline{A}$ partitioned as

$$\underline{A} = \left\{\underline{A}_{pq} \text{ of order } n_p \times n_q\right\} \text{ for } p,\ q = 1,\ 2,\ \ldots,\ N \qquad \text{(A1)*}$$

such that

$$\underline{A}_{pp} = b_p \underline{I}_{n_p} + g_{pp} \underline{J}_{n_p} \qquad \text{(A2)}$$

and $$\underline{A}_{pq} = g_{pq} \underline{J}_{n_p \times n_q} \quad \text{for } p \neq q \qquad \text{(A3)}$$

with $$\underline{G} = \left\{g_{pq}\right\} . \qquad \text{(A4)}$$

Then the inverse of $\underline{A}$ is

$$\underline{A}^{-1} = \left\{(\underline{A}^{-1})_{pq} \text{ of order } n_p \times n_q\right\} . \qquad \text{(A5)}$$

with $$(\underline{A}^{-1})_{pp} = (1/b_p)\underline{I}_{n_p} + k_{pp} \underline{J}_{n_p} \qquad \text{(A6)}$$

and $$(\underline{A}^{-1})_{pq} = k_{pq} \underline{J}_{n_p \times n_q} \quad \text{for } p \neq q \qquad \text{(A7)}$$

where $$\underline{K} = \left\{k_{pq}\right\}$$

$$= \left[(\underline{G}\underline{D} + \underline{B})^{-1} - \underline{B}^{-1}\right] \underline{D}^{-1} \qquad \text{(A8)}$$

with $$\underline{D} = \text{diag}\left\{n_1,\ \ldots,\ n_N\right\} \text{ and } \underline{B} = \text{diag}\left\{b_1,\ \ldots,\ b_N\right\} . \qquad \text{(A9)}$$

The notation in (A9) indicates that $\underline{D}$ and $\underline{B}$ are diagonal matrices, and in (A5) - (A7) the notation $(\underline{A}^{-1})_{pq}$ does not indicate the inverse of a matrix; it is the pq'th sub-matrix of the inverse of $\underline{A}$.

Comparing the definition of $\underline{V}_i$ in (20) - (22) with that of $\underline{A}$ in (A1) - (A3) indicates that in applying the theorem to find $\underline{V}_i^{-1}$ the N of the theorem is c_i, b_p is e, g_{pp} is $\alpha + \beta$ and g_{pq} is α for $p \neq q$, for p, q = 1, 2, ..., c_i. Hence $\underline{G}$ of (A4), which we now subscript with i to go with $\underline{V}_i$, is

$$\underline{G}_i = \beta \underline{I}_{c_i} + \alpha \underline{J}_{c_i} \quad . \tag{A10}$$

Then from (A5)

$$\underline{V}_i^{-1} = \left\{ (\underline{V}_i^{-1})_{jj'} \text{ of order } n_{ij} \times n_{ij'} \right\} \text{ for } j, j' = 1, 2, \ldots, c_i \tag{A11}$$

with, from (A6) and (A7),

$$(\underline{V}_i^{-1})_{jj} = (1/e)\, \underline{I}_{n_{ij}} + h_{i,jj} \underline{J}_{n_{ij}} \tag{A12}$$

and

$$(\underline{V}_i^{-1})_{jj'} = h_{i,jj'} \underline{J}_{n_{ij} \times n_{ij'}} \quad \text{for } j \neq j' \tag{A13}$$

where, from (A8)

$$\left\{ h_{i,jj'} \right\} \text{ for } j,j' = 1, 2, \ldots, c_i, = \underline{H}_i = \left[(\underline{G}_i \underline{D}_i + \underline{B}_i)^{-1} - \underline{B}_i^{-1} \right] \underline{D}_i^{-1} \tag{A14}$$

with (A9) giving

$$\underline{D}_i = \text{diag} \left\{ n_{i1}, \ldots, n_{ic_i} \right\} \text{ and } \underline{B}_i = e \underline{I}_{c_i} \, .$$

Hence to obtain $\underline{V}_i^{-1}$ we need $\underline{H}_i$ of (A14), first finding the inverse of

$$\underline{G}_i\underline{D}_i + \underline{B}_i = \begin{bmatrix} n_{i1}(\alpha+\beta)+e & n_{i2}\alpha & \cdots & n_{ic_i}\alpha \\ n_{i1}\alpha & n_{i2}(\alpha+\beta)+e & \cdots & n_{ic_1}\alpha \\ \vdots & \vdots & & \\ n_{i1}\alpha & n_{i2}\alpha & \cdots & n_{ic_i}(\alpha+\beta)+e \end{bmatrix}. \quad (A15)$$

For convenience define,

$$m_{ij} = n_{ij}\beta + e, \quad (A16)$$

$$p_i = \prod_{j=1}^{c_i} m_{ij}, \quad (A17)$$

and

$$q_i = 1 + \alpha \sum_{j=1}^{c_i} \frac{n_{ij}}{m_{ij}} , \quad (A18)$$

where m_{ij} and q_i are exactly as in (24) and (25). Then

$$\underline{G}_i\underline{D}_i + \underline{B}_i = \begin{bmatrix} n_{i1}\alpha + m_{i1} & n_{i2}\alpha & \cdots & n_{ic_i}\alpha \\ n_{i1}\alpha & n_{i2}\alpha + m_{i2} & & n_{ic_i}\alpha \\ \vdots & \vdots & & \vdots \\ n_{i1}\alpha & n_{i2}\alpha & \cdots & n_{ic_i} + m_{ic_i} \end{bmatrix}$$

with the determinant being, from diagonal expansion,

$$|\underline{G}_i\underline{D}_i + \underline{B}_i| = \prod_{j=1}^{c_i} m_{ij}\left(1 + \sum_{j=1}^{c_i} \frac{n_{ij}\alpha}{m_{ij}}\right) = p_i q_i \ . \quad (A20)$$

To find the inverse of $\underline{G}_i\underline{D}_i + \underline{B}_i$ we find the cofactors of its elements. That of its j'th diagonal element is, by analogy with (A20)

$$(1/m_{ij}) \prod_{j=1}^{c_i} m_{ij}\left(1 + \sum_{j=1}^{c_i} \frac{n_{ij}\alpha}{m_{ij}} - \frac{n_{ij}\alpha}{m_{ij}}\right) = \frac{p_i}{m_{ij}}\left(q_i - \frac{n_{ij}\alpha}{m_{ij}}\right) ; \quad (A21)$$

and that of its (jj')'th off-diagonal element for $j \neq j'$ is $(-1)^{j+j'}|\underline{M}_{i,jj'}|$ where $\underline{M}_{i,jj'}$ is the corresponding minor. In subtracting the $(j' - 1)$'th row of $|\underline{M}_{i,jj'}|$ - which, for $j < j'$, has come from the j' 'th row of $\underline{G}_i\underline{D}_i + \underline{B}_i$ - from every other row of $|\underline{M}_{i,jj'}|$ we find that for $t \neq j \neq j'$ all elements $n_{it}(\alpha + \beta) + e$ become $n_{it}\beta + e$ and elements $n_{it}\alpha$ become zero; and the only non-zero element in the j'th column is $n_{ij}\alpha$ in the $(j' - 1)$'th position. Expanding $|\underline{M}_{i,jj'}|$ by elements of this column gives

$$|\underline{M}_{i,jj'}| = (-1)^{j'-1+j} n_{ij}\alpha \prod_{t \neq j \neq j'}^{c_i} (n_{ij}\beta + e) = (-1)^{j'-1+j} n_{ij}\alpha p_i / m_{ij} m_{ij'} .$$

This is for $j < j'$. When $j > j'$, the effect is to interchange j and j' in the above result, which merely replaces n_{ij} by $n_{ij'}$. Hence the cofactor of the element in the (jj')'th position, for $j \neq j'$, is

$$(-1)^{j+j'}|\underline{M}_{i,jj'}| = -n_{ij}\alpha p_i / m_{ij} m_{ij'} . \tag{A22}$$

Dividing (A21) and (A22) by (A20) shows that $(\underline{G}_i\underline{D}_i + \underline{B}_i)^{-1}$ has its

$$\text{j'th diagonal element} = \frac{1}{m_{ij}} - \frac{n_{ij}\alpha}{m_{ij}^2 \, q_i}$$

and its

$$\text{(jj')'th element, } j \neq j', = \frac{- n_{ij'}\alpha}{m_{ij} \, m_{ij'} \, q_i} .$$

Therefore, with $\underline{D}_i^{-1} = \text{diag}\left\{1/n_{i1}, \ldots, 1/n_{ic_i}\right\}$ and $\underline{B}_i^{-1} = (1/e)\underline{I}_{c_i}$, the matrix $\underline{H}_i$ of (A14) has diagonal elements

$$h_{i,jj} = \left[\frac{1}{m_{ij}} - \frac{n_{ij}\alpha}{m_{ij}^2 \, q_i} - \frac{1}{e} \right] \frac{1}{n_{ij}} \tag{A23}$$

which, from (A16) reduces to

$$h_{i,jj} = \frac{-\beta}{m_{ij}e} - \frac{\alpha}{m_{ij}^2 q_i} \quad \text{for } j = 1, 2, \ldots, c_i; \tag{A24}$$

and off-diagonal elements

$$h_{i,jj'} = \frac{-\alpha}{m_{ij}\, m_{ij'}\, q_i} \quad \text{for } j \neq j' = 1, 2, \ldots, c_i . \tag{A25}$$

With these terms used in (A12) and (A13). $\underline{V}_i^{-1}$ of (A11) is determined and so $\underline{V}^{-1} = \sum_{i=1}^{a}{}^{+} \underline{V}_i$ is known.

Elements of T

To derive $t_{\alpha\beta} = tr(\underline{V}^{-1}\underline{V}_\alpha\underline{V}^{-1}\underline{V}_\beta)$ for example, we need the differentials of $\underline{V}$ with respect to α and β. This, because $\underline{V} = \sum_{i=1}^{a}{}^{+} \underline{V}_i$, requires the differentials of $\underline{V}_i$, and from the definitions given in (20) - (22) it is readily seen that

$$\underline{V}_{i,\alpha} = \partial\underline{V}_i/\partial\sigma_\alpha^2 = \underline{J}_{n_{i.}} ,$$

$$\underline{V}_{i,\beta} = \partial\underline{V}_i/\partial\sigma_\beta^2 = \sum_{j=1}^{c_i}{}^{+} \underline{J}_{n_{ij}} ,$$

and

$$\underline{V}_{i,e} = \partial\underline{V}_i/\partial\sigma_e^2 = \underline{I}_{n_{i.}} .$$

With these values, and obtaining $\underline{V}_i^{-1}$ of (A11) by using (A23) - (A25) in (A12) and (A13) we now derive the t's. First, $t_{\alpha\alpha}$, utilizing $\underline{V} = \sum_{i=1}^{a}{}^{+} \underline{V}_i$ and $\underline{V}^{-1} = \sum_{i=1}^{a}{}^{+} \underline{V}_i^{-1}$:

$$t_{\alpha\alpha} = tr(\underline{V}^{-1}\underline{V}_\alpha\underline{V}^{-1}\underline{V}_\alpha)$$

$$= \sum_{i=1}^{a} tr(\underline{V}_i^{-1}\underline{V}_{i,\alpha}\underline{V}_i^{-1}\underline{V}_{i,\alpha})$$

$$= \sum_{i=1}^{a} tr(\underline{V}_i\underline{J}_{n_{i.}})^2$$

$$= \sum_{i=1}^{a} \sum_{r=1}^{n_{i.}} \text{i.p.o. (r'th row of } \underline{V}_i^{-1}\underline{J}_{n_{i.}})$$

$$\text{and (r'th column of } \underline{V}_i^{-1}\underline{J}_{n_{i.}})$$

where i.p.o. stands for "inner product of". Now in $\underline{V}_i^{-1}\underline{J}_{n_i}$ every element in a row is the same; and so the columns are equal. Therefore

$$t_{\alpha\alpha} = \sum_{i=1}^{a} \sum_{r=1}^{n_{i.}} (\text{element of r'th row of } \underline{V}_i^{-1}\underline{J}_{n_{i.}})(\text{column sum of } \underline{V}_i^{-1}\underline{J}_{n_{i.}}) \quad \text{(A26)}$$

Now, by the nature of $\underline{V}_i^{-1}$, the $n_{i.}$ rows of $\underline{V}_i^{-1}\underline{J}_{n_{i.}}$ are grouped naturally into c_i sets of n_{ij} rows each, for $j = 1, 2, \ldots, c_i$. Therefore instead of considering the r'th row of $\underline{V}_i^{-1}\underline{J}_{n_{i.}}$ in $t_{\alpha\alpha}$ we consider the k'th row in the j'th set of rows, and replace the summation $\sum_{r=1}^{n_i}$ by $\sum_{j=1}^{c_i}\sum_{k=1}^{n_{ij}}$. Then

an element in the k'th row of the j'th set of rows of $\underline{V}_i^{-1}\underline{J}_{n_{i.}}$

= Σ elements in the k'th row of the j'th set of rows of $\underline{V}_i^{-1}$,

and from the nature of (A12) and (A13) this is

$$= (1/e + n_{ij}\, h_{i,jj}) + \sum_{j'\neq j} n_{ij'}\, h_{i,jj'}$$

which, on substituting from (A23) and (A25) is

$$= \frac{1}{m_{ij}} - \frac{n_{ij}\,\alpha}{m_{ij}^2\, q_i} - \sum_{j'\neq j} \frac{n_{ij'}\,\alpha}{m_{ij}\, m_{ij'}\, q_i}$$

$$= \frac{1}{m_{ij}} - \frac{n_{ij}\,\alpha}{m_{ij}^2\, q_i} - \frac{\alpha}{m_{ij}\, q_i}\left(\sum_{j'=1}^{c_i} \frac{n_{ij'}}{m_{ij'}} - \frac{n_{ij}}{m_{ij}}\right)$$

$$= \frac{1}{m_{ij}} - \frac{\alpha}{m_{ij}\, q_i}\sum_{j=1}^{c_i}\frac{n_{ij}}{m_{ij}} = \frac{1}{m_{ij}} - \frac{1}{m_{ij}\, q_i}(q_i - 1)$$

$$= \frac{1}{m_{ij}\, q_i}, \quad \text{(A27)}$$

on making use of (A16) and (A18) in this reduction. Thus an element in the k'th row of the j'th set of rows of $\underline{V}_i^{-1}\underline{J}_{n_{i.}}$ is $1/m_{ij}q_i$, and summing this over all rows gives

$$\text{column sum of } \underline{V}_i^{-1}\underline{J}_{n_{i.}} = \sum_{j=1}^{c_i}\sum_{k=1}^{n_{ij}} \frac{1}{m_{ij}\,q_i} = \frac{1}{q_i}\sum_{j=1}^{c_i}\frac{n_{ij}}{m_{ij}} . \qquad (A28)$$

Substituting (A27) and (A28) in (A26) gives

$$t_{\alpha\alpha} = \sum_{i=1}^{a}\sum_{j=1}^{c_i}\sum_{k=1}^{n_{ij}} \frac{1}{m_{ij}\,q_i}\left(\frac{1}{q_i}\sum_{j=1}^{c_i}\frac{n_{ij}}{m_{ij}}\right) = \sum_{i=1}^{a}\frac{1}{q_i^2}\left(\sum_{j=1}^{c_i}\frac{n_{ij}}{m_{ij}}\right)^2 .$$

Now define

$$A_{ipq} = \sum_{j=1}^{c_i}\frac{n_{ij}^p}{m_{ij}^q} \qquad \text{for integers p and q,}$$

as in (25), noting from (A18) that

$$q_i = 1 + \alpha A_{i11} = 1 + \sigma_\alpha^2 A_{i11}$$

as in (26). Then $t_{\alpha\alpha}$ gets written as

$$t_{\alpha\alpha} = \sum_{i=1}^{a} A_{i11}^2/q_i^2$$

as shown in (27).

The definitions and procedures introduced in deriving this result are used repeatedly below in obtaining the other t's.

$$t_{\alpha\beta} = tr(\underline{V}^{-1}\underline{V}_\alpha\underline{V}^{-1}\underline{V}_\beta)$$

$$= \sum_{i=1}^{a} tr(\underline{V}_i^{-1}\underline{V}_{i,\alpha}\underline{V}_i^{-1}\underline{V}_{i,\beta})$$

$$= \sum_i tr(\underline{V}_i^{-1}\,\underline{J}_{n_{i.}}\,\underline{V}_i^{-1}\sum_j{}^+\,\underline{J}_{n_{ij}})$$

$$= \sum_i\sum \text{ i.p.o. (r'th row of } \underline{V}_i^{-1}\underline{J}_{n_{i.}}\text{) and (r'th column of } \underline{V}_i^{-1}\sum_j{}^+\,\underline{J}_{n_{ij}}\text{)}$$

$$= \sum_i\sum_{r=1}^{n_{i.}}\text{(element of r'th row of } \underline{V}_i^{-1}\underline{J}_{n_{i.}}\text{)}(\Sigma \text{ elements in r'th column of } \underline{V}_i^{-1}\sum_j{}^+\,\underline{J}_{n_{ij}}\text{)}$$

$$= \sum_{i}^{a} \sum_{j}^{c_i} \sum_{k=1}^{n_{ij}} (1/m_{ij}\ q_i)(\Sigma \text{ elements in k'th column of j'th set of columns of } \underset{-}{V}_i^{-1} \Sigma_j^{+} \underset{-}{J}_{n_{ij}})$$

$$= \sum_i \sum_j \sum_k (1/m_{ij}\ q_i) \left\{\Sigma \text{ elements in the column } [(0 \ldots 0\ \underset{-}{1}_{n_{ij}}\ 0 \ldots 0)\underset{-}{V}_i^{-1}]'\right\}$$

$$= \sum_i \sum_j \sum_k (1/m_{ij}\ q_i)(\Sigma \text{ all elements in the j'th set of columns of } \underset{-}{V}_i^{-1})$$

$$= \sum_i \sum_j \sum_k (1/m_{ij}\ q_i) \left\{\Sigma \text{ all elements in } [(\underset{-}{V}_i^{-1})_{jj} + \sum_{j \neq j} (\underset{-}{V}_i^{-1})_{jj'}]\right\}$$

$$= \sum_i \sum_j \sum_k (1/m_{ij}\ q_i)(n_{ij}/e + n_{ij}^2\ h_{i,jj} + \sum_{j' \neq j} n_{ij}\ n_{ij'}\ h_{i,jj'})$$

$$= \sum_i \sum_j \sum_k (1/m_{ij}\ q_i)(n_{ij}/m_{ij}\ q_i) \text{ from (A27)}$$

$$= \sum_i \frac{1}{q_i^2} \sum_j \frac{n_{ij}^2}{m_{ij}^2}$$

$$= \sum_i A_{i22}/q_i^2 \text{ as shown in (28).}$$

Similarly

$$t_{\alpha e} = \text{tr}(\underset{-}{V}^{-1} \underset{-}{V}_{\alpha} \underset{-}{V}^{-1} \underset{-}{V}_e)$$

$$= \sum_{i=1}^{a} \text{tr}(\underset{-}{V}_i^{-1} \underset{-}{V}_{i,\alpha} \underset{-}{V}_i^{-1} \underset{-}{V}_{i,e})$$

$$= \sum_i \text{tr}(\underset{-}{V}_i^{-1}\ \underset{-}{J}_{n_{i.}}\ \underset{-}{V}_i^{-1})$$

$$= \sum_i \sum_r (\text{element in r'th row of } \underset{-}{V}_i^{-1} \underset{-}{J}_{n_{i.}})(\Sigma \text{ elements in r'th column of } \underset{-}{V}_i^{-1})$$

$$= \sum_i \sum_j \sum_k (1/m_{ij}\ q_i)(1/e + n_{ij}\ h_{i,jj} + \sum_{j' \neq j} n_{ij'}\ h_{i,jj'})$$

$$= \sum_i \sum_j \frac{n_{ij}}{m_{ij}\ q_i} \left(\frac{1}{m_{ij}\ q_i}\right) \text{ from (A26)}$$

$$= \sum_i A_{i12}/q_i^2 \text{ , given in (29).}$$

In deriving the term

$$t_{\beta\beta} = \mathrm{tr}(\underline{V}^{-1}\underline{V}_{\beta}\underline{V}^{-1}\underline{V}_{\beta})$$

$$= \sum_i \mathrm{tr}[(\Sigma_j^+ \underline{J}_{n_{ij}})\underline{V}_i^{-1}]^2$$

we use the partitioning

$$(\Sigma_j^+ \underline{J}_{n_{ij}})\underline{V}_i^{-1} = \left\{\underline{P}_{i,jj'}\right\} \text{ for } j,j' = 1, 2, \ldots, c_i$$

with $\underline{P}_{i,jj} = (\Sigma_j^+ \underline{J}_{n_{ij}})(\underline{V}_i^{-1})_{jj} = (1/e)\underline{J}_{n_{ij}} + n_{ij}\, h_{i,jj}\, \underline{J}_{n_{ij}}$

and $\underline{P}_{i,jj'} = (\Sigma_j^+ \underline{J}_{n_{ij}})(\underline{V}_i^{-1})_{jj} = n_{ij}\, h_{i,jj'}\, \underline{J}_{n_{ij} \times n_{ij'}}$.

This gives

$$t_{\beta\beta} = \sum_i \mathrm{tr}(\left\{\underline{P}_{i,jj'}\right\} j,j' = 1, 2, \ldots, c_i)^2$$

$$= \sum_i \mathrm{tr}(\sum_j \underline{P}^2_{i,jj} + \sum_j \sum_{j'\neq j} \underline{P}_{i,jj'}\underline{P}_{i,j'j})$$

$$= \sum_i \sum_j \mathrm{tr}(\underline{P}^2_{i,jj} + \sum_{j'\neq j} \underline{P}_{i,jj'}\underline{P}_{i,j'j})$$

$$= \sum_i \sum_j \mathrm{tr}[(1/e^2)n_{ij}\underline{J}_{n_{ij}} + n^2_{ij}h^2_{i,jj}n_{ij}\underline{J}_{n_{ij}} + 2(1/e)n_{ij}h_{i,jj}n_{ij}\underline{J}_{n_{ij}}$$

$$+ \sum_{j'\neq j} n_{ij}h_{i,jj'}\underline{J}_{n_{ij} \times n_{ij'}} n_{ij'}h_{i,j'j}\underline{J}_{n_{ij'} \times n_{ij}}]$$

$$= \sum_i \sum_j [n^2_{ij}(1/e + n_{ij}h_{i,jj})^2 + n^2_{ij} \sum_{j'\neq j} n^2_{ij'}h^2_{i,jj'}]$$

$$= \sum_i \sum_j \left[n^2_{ij}\left(\frac{1}{m_{ij}} - \frac{n_{ij}\,\alpha}{m^2_{ij}\,q_i}\right)^2 + n^2_{ij} \sum_{j'\neq j} \frac{n_{ij'}\,\alpha^2}{m^2_{ij}\,m^2_{ij'}\,q^2_i}\right]$$

$$= \sum_i \sum_j \left[\frac{n^2_{ij}}{m^2_{ij}}\left(1 + \frac{n^2_{ij}\,\alpha^2}{m^2_{ij}\,q^2_i} - \frac{2n_{ij}\,\alpha}{m_{ij}\,q_i}\right) + \frac{n^2_{ij}\,\alpha^2}{m^2_{ij}\,q^2_i}\left(\sum_{j'} \frac{n^2_{ij'}}{m^2_{ij'}} - \frac{n^2_{ij}}{m^2_{ij}}\right)\right]$$

$$= \sum_i \sum_j \left(\frac{n_{ij}^2}{m_{ij}^2} - \frac{2\alpha\, n_{ij}^3}{m_{ij}^3\, q_i} + \frac{\alpha^2\, n_{ij}^2}{q_i^2\, m_{ij}^2} \sum_{j'} \frac{n_{ij'}^2}{m_{ij'}^2} \right)$$

$$= \sum_{i=1}^{a} (A_{i22} - 2\sigma_\alpha^2\, A_{i33}/q_i + \sigma_\alpha^4 A_{i22}/q_i^2) \text{ as in (30).}$$

The penultimate term is

$$t_{\beta e} = \mathrm{tr}(\underline{V}^{-1}\underline{V}_\beta \underline{V}^{-1}\underline{V}_e)$$

$$= \sum_i \mathrm{tr}(\underline{V}_i^{-1}\underline{V}_{i,\beta}\underline{V}_i^{-1})$$

$$= \sum_i \mathrm{tr}[(\Sigma_j^+ \underline{J}_{n_{ij}})\underline{V}_i^{-1}\underline{V}_i^{-1}]$$

$$= \sum_i \mathrm{tr}[\Sigma_j \underline{P}_{i,jj}(\underline{V}_i^{-1})_{jj} + \sum_j \sum_{j' \neq j} \underline{P}_{i,jj'}(\underline{V}_i^{-1})_{j'j}]$$

$$= \sum_i \sum_j \mathrm{tr}\Big\{[(1/e)\underline{J}_{n_{ij}} + n_{ij}h_{i,jj}\underline{J}_{n_{ij}}][(1/e)\underline{I}_{n_{ij}} + h_{i,jj}\underline{J}_{n_{ij}}]$$

$$+ \sum_{j' \neq j} n_{ij}h_{i,jj'}\underline{J}_{n_{ij}} \times n_{ij'} h_{i,jj'}\underline{J}_{n_{ij'}} \times n_{ij}\Big\}$$

$$= \sum_i \sum_j \mathrm{tr}[(1/e^2)\underline{J}_{n_{ij}} + 2(1/e)n_{ij}h_{i,jj}\underline{J}_{n_{ij}} + n_{ij}^2 h_{i,jj}^2 \underline{J}_{n_{ij}}$$

$$+ \sum_{j' \neq j} n_{ij}n_{ij'}h_{i,jj'}^2 \underline{J}_{n_{ij}}]$$

$$= \sum_i \sum_j (n_{ij}/e^2 + 2n_{ij}^2 h_{i,jj}/e + n_{ij}^3 h_{i,jj}^2 + n_{ij}^2 \sum_{j' \neq j} n_{ij'}h_{i,jj'}^2)$$

$$= \sum_i \sum_j [n_{ij}(1/e + n_{ij}h_{i,jj})^2 + n_{ij}^2 \sum_{j' \neq j} n_{ij'}h_{i,jj'}^2]$$

$$= \sum_i \sum_j \left[\frac{n_{ij}}{m_{ij}^2} \left(1 + \frac{n_{ij}^2 \alpha^2}{m_{ij}^2 q_i^2} - \frac{2n_{ij}\alpha}{m_{ij}q_i}\right) + \frac{n_{ij}^2 \alpha^2}{m_{ij}^2 q_i^2} \left(\sum_{j'} \frac{n_{ij'}}{m_{ij'}^2} - \frac{n_{ij}}{m_{ij}^2}\right)\right]$$

and, similar to the final reduction of $t_{\beta\beta}$, this becomes, as shown in (31),

$$t_{\beta e} = \sum_{i=1}^{a} (A_{i12} - 2\sigma_\alpha^2 A_{i23}/q_i + \sigma_\alpha^4 A_{i12}A_{i22}/q_i^2).$$

Finally we have

$$t_{ee} = tr(\underset{\sim}{V}^{-1}\underset{\sim}{V}_e\underset{\sim}{V}^{-1}\underset{\sim}{V}_e)$$

$$= \sum_i tr(\underset{\sim}{V}_i^{-1})^2$$

$$= \sum_i \sum_j tr\left\{[(\underset{\sim}{V}_i^{-1})_{jj}]^2 + \sum_{j'\neq j}(\underset{\sim}{V}_i^{-1})_{jj'}(\underset{\sim}{V}_i^{-1})_{j'j}\right\}$$

and on substituting from (A12) and (A13) this becomes

$$t_{ee} = \sum_i \sum_j tr\left\{(1/e^2)\underset{\sim}{I}_{n_{ij}} + 2(1/e)h_{ijj}\underset{\sim}{J}_{n_{ij}} + h^2_{i,jj}n_{ij}\underset{\sim}{J}_{n_{ij}}\right.$$

$$\left.+ \sum_{j'\neq j} h_{i,jj'}\underset{\sim}{J}_{n_{ij}} \times n_{ij'} h_{i,jj'}\underset{\sim}{J}_{n_{ij'}} \times n_{ij}\right\}$$

$$= \sum_i \sum_j (n_{ij}/e^2 + 2n_{ij}h_{i,jj}/e + n^2_{ij}h^2_{i,jj} + n_{ij}\sum_{j'\neq j} n_{ij'}h^2_{ijj'})$$

$$= \sum_i \sum_j [(n_{ij} - 1)/e^2 + (1/e + n_{ij}h_{i,jj})^2 + n_{ij}\sum_{j'\neq j} n_{ij'}h^2_{ijj'}]$$

$$= \frac{n_{..} - c_.}{e^2} + \sum_i \sum_j \left[\frac{1}{m^2_{ij}}\left(1 + \frac{n^2_{ij}\alpha^2}{m^2_{ij}q^2_i} - \frac{2n_{ij}\alpha}{m_{ij}q_i}\right) + \right.$$

$$\left.+ \frac{n_{ij}\ \alpha^2}{m^2_{ij}\ q^2_i}\left(\sum_{j'\neq j}\frac{n_{ij'}}{m^2_{ij'}} - \frac{n_{ij}}{m^2_{ij}}\right)\right]$$

$$= (n_{..} - c_.)/e^2 + \sum_{i=1}^{a}(A_{i02} - 2\sigma^2_\alpha A_{i13}/q_i + \sigma^4_\alpha A^2_{i12}/q^2_i)$$

as shown in (32).

Balanced Data

With $n_{ij} = n$, $c_i = c$, for all i and j,

$$m_{ij} = n\beta + e \qquad A_{ipq} = \frac{cn^p}{(n\beta + e)^q} \quad \text{and} \quad q_i = \frac{cn\alpha + n\beta + e}{n\beta + e}\ .$$

$$x = \frac{a(c - 1)}{(n\beta + e)^2}, \quad y = \frac{a}{(cn\alpha + n\beta + e)^2} \quad \text{and} \quad z = \frac{ac(n - 1)}{e^2}$$

we obtain, from (27) - (32):

$$t_{\alpha\alpha} = a\left(\frac{cn}{n\beta + e}\right)^2 \left(\frac{n\beta + e}{cn\alpha + n\beta + e}\right)^2 = c^2n^2y ,$$

$$t_{\alpha\beta} = \frac{acn^2}{(n\beta + e)^2} \left(\frac{n\beta + e}{cn\alpha + n\beta + e}\right)^2 = cn^2y ,$$

$$t_{\alpha e} = \frac{acn}{(n\beta + e)^2} \left(\frac{n\beta + e}{cn\alpha + n\beta + e}\right)^2 = cny ,$$

and

$$t_{\beta\beta} = \frac{acn^2}{(n\beta + e)^2}\left[1 - \frac{2\alpha n}{cn\alpha + n\beta + e} + \frac{c\alpha^2 n^2}{(cn\alpha + n\beta + e)^2}\right]$$

$$= \frac{an^2}{(n\beta + e)^2}\left[\frac{c(cn\alpha + n\beta + e)^2 - 2\, cn\alpha(cn\alpha + n\beta + e)^2 + c^2n^2\alpha^2}{(cn\alpha + n\beta + e)^2}\right]$$

in which the numerator inside the square brackets can be simplified as

$$(n\beta + e)^2 + (c-1)(n\beta + e)^2 + c^3n^2\alpha^2 + 2c^2n\alpha(n\beta + e) -2cn\alpha(cn\alpha + n\beta + e) + cn^2\alpha^2$$

$$= (n\beta + e)^2 + (c-1)(n\beta + e)^2 + (c-1)c^2n^2\alpha^2 + 2cn\alpha[c(n\beta + e) - (cn\alpha + n\beta + e) + cn\alpha]$$

$$= (n\beta + e)^2 + (c-1)(cn\alpha + n\beta + e)^2$$

and so

$$t_{\beta\beta} = \frac{an^2(c-1)}{(n\beta + e)^2} + \frac{an^2}{(cn\alpha + n\beta + e)^2} = n^2(x + y) .$$

Furthermore, since in (31) the powers of n in the A-terms are one less than those in (30)

$$t_{\beta e} = t_{\beta\beta}/n = n(x + y) .$$

And similarly for the first term of (32), so that

$$t_{ee} = t_{\beta e}/n + (acn - ac)/e^2 = x + y + z.$$

The 1-way classification

With $\sigma_\beta^2 = 0$ and using $w_i = n_i e/(e + n_i \alpha)$, (25) and (26) give

$$A_{ipq} = e^{-q} \sum_j n_{ij}^p$$

and $$q_i = (e + n_{i.} \alpha)/e = n_{i.}/w_i \ .$$

Thus from (27), (29) and (32)

$$t_{\alpha\alpha} = \Sigma\, e^{-2} n_i^2 w_i^2 / n_i^2 = e^{-2} \Sigma\, w_i^2$$

$$t_{\alpha e} = \Sigma\, e^{-2} n_{i.} w_i^2 / n_{i.}^2 = e^{-2} \Sigma\, w_i^2 / n_{i.}$$

and

$$t_{ee} = \Sigma \left[\frac{c_i}{e^2} - \frac{2\alpha\, e}{e + n_{i.}\alpha} \frac{n_{i.}}{e^3} + \frac{e^2 n_{i.}^2}{(e + n_{i.}\alpha)^2 e^4} \right] + \frac{n_{..} - c_.}{e^2}$$

$$= e^{-2}\left[\Sigma(c_i - 1) + \Sigma\left(1 - \frac{\alpha n_{i.}}{e + n_{i.}\alpha}\right)^2 \right] + e^{-2}(n_{..} - c_.)$$

$$= e^{-2}(\Sigma\, w_i^2 / n_{i.}^2 + n_{..} - a),$$

as shown in (33).

The Lagrange identity is

$$\sum_i a_i^2 \sum_i b_i^2 - \left(\sum_i a_i b_i\right)^2 = \tfrac{1}{2} \sum_{i \neq i'} \sum (a_i b_{i'} - a_{i'} b_i)^2 .$$

Hence, with $$w_i = n_{i.} e/(n_{i.}\alpha + e)$$

$$\sum_i w_i^2 \sum_i w_i^2/n_{i.}^2 - (\sum_i w_i^2/n_{i.})^2$$

$$= \tfrac{1}{2}\sum_{i\neq i'}\sum (w_i w_{i'}/n_{i'.} - w_i w_{i'}/n_{i.}) = \tfrac{1}{2}\sum_{i\neq i'}\sum \left[\frac{w_i w_{i'}(n_{i.} - n_{i'.})}{n_{i.}n_{i'.}}\right]^2$$

$$= \tfrac{1}{2}\sum_{i\neq i'}\sum \left[\frac{e^2(n_{i.} - n_{i'.})}{(n_{i.}\alpha + e)(n_{i'.}\alpha + e)}\right]^2 = \tfrac{1}{2}\sum_{i\neq i'}\sum \left[\frac{en_{i.}(n_{i'.}\alpha + e) - en_{i'.}(n_{i.}\alpha + e)}{(n_{i.}\alpha + e)(n_{i'.}\alpha + e)}\right]^2$$

$$= \tfrac{1}{2}\sum_{i\neq i'}\sum (w_i - w_{i'})^2 \qquad = (a - 1)\sum_i w_i^2 - 2\sum_{i>i'}\sum w_i w_{i'}$$

$$= a\,\sum w_i^2 - (\sum w_i)^2 .$$

Hence in (34)

$$|T| = e^{-4}\left[a\,\sum w_i^2 - (\sum w_i)^2 + (n_{..} - a)\,\sum w_i^2\right] = e^{-4}\left[n_{..}\sum w_i^2 - (\sum w_i)^2\right] = e^{-4}D$$

as in (35).

Estimation of μ

$$\underline{1}'\underline{V}^{-1}\underline{y} = \sum_i(\Sigma \text{ elements in } \underline{y}_i'\underline{V}_i^{-1})$$

$$= \sum_i\sum_j\sum_k y_{ijk}(\Sigma \text{ elements in k'th row of j'th set of rows of } \underline{V}_i^{-1})$$

$$= \sum_i\sum_j\sum_k y_{ijk}(1/e + n_{ij}h_{i,jj} + \sum_{j'\neq j} n_{ij'}h_{i,jj'}), \text{ from (A12) and (A13),}$$

$$= \sum_i\sum_j\sum_k y_{ijk}/m_{ij}q_i \quad \text{from (A27)}$$

$$= \sum_i\sum_j n_{ij}\,\bar{y}_{ij.}/m_{ij}q_i$$

$$= \sum_i \frac{1}{q_i}\sum_j \frac{n_{ij}}{m_{ij}}\bar{y}_{ij.} \quad \text{as in (37).}$$

$$\underset{\sim}{1}'\underset{\sim}{V}^{-1}\underset{\sim}{1} = \sum_i (\Sigma \text{ all elements in } \underset{\sim}{V}_i^{-1})$$

$$= \sum_i \sum_j \sum_k (1/e + n_{ij}h_{i,jj} + \sum_{j' \neq j} n_{ij'}h_{i,jj'}) \text{ from above,}$$

$$= \sum_i \sum_j \sum_k \left[\frac{1}{m_{ij}} - \frac{n_{ij}\,\alpha}{m_{ij}^2\, q_i} - \frac{\alpha}{m_{ij}\, q_i} \left(\sum_{j'} \frac{n_{ij'}}{m_{ij'}} - \frac{n_{ij}}{m_{ij}} \right) \right], \text{ from (A23 and (A25)}$$

$$= \sum_i \left(\sum_j \frac{n_{ij}}{m_{ij}} \right) \left(1 - \frac{\alpha}{q_i} \sum_j \frac{n_{ij}}{m_{ij}} \right)$$

and using the definition of q_i this becomes

$$\underset{\sim}{1}'\underset{\sim}{V}^{-1}\underset{\sim}{1} = \sum_i \left(\sum_j \frac{n_{ij}}{m_{ij}} \right) \frac{1}{q_i} \quad \text{as in (38).}$$