Statistical Research Report
Institue of Mathematics
University of Oslo

No. 2
August 1966

CONFIDENCE INTERVALS AND TESTS FOR VARIANCE RATIOS IN UNBALANCED VARIANCE COMPONENTS MODELS.*

By

Emil Spjøtvoll

*Revised March 1967.

1. INTRODUCTION. In this paper we consider the variance components model in the analysis of variance, which in the notation of Graybill and Hultquist [2] can be described as follows. An n x 1 vector of observations Y is assumed to be a linear sum of $k+2$ quantities,

$$Y = \sum_{i=0}^{k+1} X_i \beta_i \tag{1.1}$$

where β_0 is a fixed unknown constant, β_i $(i = 1,2,\dots,k+1)$ is a $p_i \times 1$ vector of multinormally distributed random variables with expectation 0 and covariance matrix $\sigma_i^2 I$. I denotes the identity matrix and 0 the null matrix. The vectors β_1, $\beta_2,\dots,$ β_{k+1} are mutually independent. X_0 is an $n \times 1$ vector of 1's, X_i $(i = 1,2,\dots,k)$ is an $n \times p_i$ matrix of known constants and $X_{k+1} = I$.

Some general theorems concerning this model have been dericed by Graybill and Hultquist [2] under one or both of the assumptions

(I) A_i and A_j commute, where $A_i = X_i X_i'$ $(i,j = 1,2,\dots,k)$.

(II) The matrix X_i is such that $j'_n X_i = r_i j'_{p_i}$ and $X_i j_{p_i} = j_n$, where r_i is a positive integer and j_p generally denotes a $p \times 1$ vector of 1's$(i = 1,2,\dots,n)$.

It has been observed in [2] that the usual balanced experimental design models satisfy these assumptions.

Wald [7] considers the model (1,1) where some of the β_i's are fixed and not random effects, and derives confidence intervals for the ratios $\sigma_i^2/\sigma_{k+1}^2$ without the restrictions (I) and (II). In Section 2 it is shown, contrary to what formerly seems to be accepted ([1] ,p.11 and p.15, [5] p.206) , that Wald's method for obtaining confidence intervals does not in general apply to experimental design models.

In Section 3 it is proved that the usual method of transforming the elements of Y into independent variables by a linear transformation is not possible when assumption (I) does not hold, and Theorem 2 gives an equivalent condition to (I) which is easier to apply and which shows that practically only balanced models satisfy condition (I).

In Section 4 a method for obtaining confidence intervals and testing hypotheses concerning the $\sigma_i'^2$'s for unbalanced experimental design models is given. The method is illustrated on the two-way classification model.

2. WALD'S METHOD. Wald obtains a confidence for $\sigma_i'^2/\sigma_{k+1}^2$ as follows. Derive the least square estimate $\hat{\beta}_i$ of β_i as in a regression model. Then conditionally given β_i the difference $\hat{\beta}_i - \beta_i$ is a random multinormally distributed vector e_i with $E(e_i) = 0$ and $E(e_i e_i') = C\,\sigma_{k+1}^2$ where C may be computed by means of the X_i's . Since the conditional distribution of e_i given β_i does not depend upon β_i, the two vectors are independent. It follows that $\hat{\beta}_i = \beta_i + e_i$ has a multinormal distribution with $E(\hat{\beta}_i) = 0$ and $E(\hat{\beta}_i \hat{\beta}_i') = (C + (\sigma_i'^2/\sigma_{k+1}'^2)\,I)\,\sigma_{k+1}^2$.

Let Q_a be independent of $\hat{\beta}_i'(C+(\sigma_i^2/\sigma_{k+1}^2)\,I)^{-1}\hat{\beta}_i$, such that Q_a/σ_{k+1}^2 is χ^2 - distributed.

Then

$$(2.1) \qquad \frac{\hat{\beta}_i'(C+(\sigma_i^2/\sigma_{k+1}^2)I)^{-1}\hat{\beta}_i}{Q_a}$$

multiplied with a suitable constant has an F-distribution. It can be shown that (2.1) decreases with $\sigma_i^2/\sigma_{k+1}^2$, and hence a confidence intefval is obtained in the usual way.

Unfortunately this method cannot generally be applied to experimental design models. In these models there are side conditions on the β_i's, which the realizations of the random vectors β_i with probability 1 do not satisfy. Without the side conditions the different elements of the β_i's are not in general estimable, and the above method cannot be applied.

This is not easily remedied as shown by the following example. Consider the two-way classification model with random effects.

$$(2.2) \qquad y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk} \qquad \begin{array}{l} i = 1,2,\ldots,r, \\ j = 1,2,\ldots,s, \\ k = 1,2,\ldots,m_{ij}, \end{array}$$

where μ is an unknown constant, α_i, β_j, γ_{ij} and e_{ijk} are independent normal random variables with expectations 0 and variances $\sigma_A^2, \sigma_B^2, \sigma_{AB}^2$ and σ^2 respectively.

We shall show that there does not exist a linear function T of the y_{ijk}'s

$$T = \sum_{i,j,k} a_{ijk}\, y_{ijk}$$

such that the variance of T depends only upon σ_A^2 and σ^2. Note that the elements of $\hat{\beta}_i$ given above satisfy this condition with $\sigma_i^2 = \sigma_A^2$ and $\sigma_{k+1}^2 = \sigma^2$.

Define $a_{i..} = \sum_{j,k} a_{ijk}$, $a_{.j.} = \sum_{i,k} a_{ijk}$ and $a_{ij.} = \sum_k a_{ijk}$. Then $\operatorname{Var} T = \sigma_A^2 \sum_i a_{i..}^2 + \sigma_B^2 \sum_j a_{.j.}^2 + \sigma_{AB}^2 \sum_{i,j} a_{ij.}^2 + \sigma^2 \sum_{i,j,k} a_{ijk}^2$. If $\operatorname{Var} T$ shall not depend upon σ_{AB}^2, we must have $a_{ij.} = 0$ for all i and j. But then $a_{i..} = \sum_j a_{ij.} = 0$ and $a_{.j.} = \sum_i a_{ij.} = 0$. Hence

$$\operatorname{Var} T = \sigma^2 \sum_{i,j,k} a_{ijk}^2 .$$

It follows that we cannot derive a confidence interval for σ_A^2/σ^2 by finding linear combinations of the y_{ijk}'s with covariance matrix depending only upon σ_A^2 and σ^2.

3. ON THE POSSIBILITY OF TRANSFORMING INTO INDEPENDENT VARIABLES.

The covariance matrix of Y is $\sum_{i=1}^{k+1} A_i \sigma_i^2$. Let $Z = PY$ be a linear transformation of Y. Then the covariance matrix of Z is $\sum_{i=1}^{k+1} PA_iP' \sigma_i^2$. If P is such that the elements of Z are independent for all values of σ_i $(i = 1,2,\ldots,k+1)$, then the matrices PA_iP' must be diagonal, and since $A_{k+1} = I$, PP' must be diagonal. Hence if there exists a matrix transforming Y into a vector with independent elements, then there exists an orthogonal matrix with this property.

$A_1, A_2, \ldots, A_k$ are real symmetric matrices. By a theorem from matrix theory ([4] , p.213), there exists an orthogonal matrix P such that PA_iP' $(i = 1,2,\ldots,k)$ is diagonal if and only if $A_1, A_2, \ldots, A_k$ commute in pairs. Hence we have the following theorem.

THEOREM 1. There exists a linear transformation P such that the elements of $Z = PY$ are independent for all values of σ_i $(i = 1,2,\ldots,k+1)$ if and only if the matrices $A_1, A_2, \ldots, A_k$ commute in pairs. The matrix P can be chosen to be orthogonal.

If such a transformation P exists the variances of the elements of PY can be obtained and used to find estimates and confidence intervals, and to test hypotheses concerning $\sigma_1, \sigma_2, \ldots, \sigma_{k+1}$. This is the procedure used e.g. by Herbach [3] and Graybill and Hultquist [2].

We shall find a simpler form of the condition in Theorem 1 when the following assumption is satisfied.

(III) The elements of the matrix X_i are 0's and 1's and

$$X_i j_{p_i} = j_n \quad (i = 1,2,\ldots,k).$$

This holds for the usual experimental design models, balanced or unbalanced.

We shall need the following lemma which is obvious from the definition of A_i as $X_i X_i'$.

LEMMA. A_i and A_j commute if and only if $A_i A_j$ is symmetric. Incidentally this gives a simpler form of assumption (I).

Since the matrices shall commute only pairwise we will in the following be concerned only with the matrices A_1 and A_2 corresponding to what we will call factors 1 and 2. We shall say that factor 1 has P_1 levels $\alpha_1, \alpha_2, \ldots, \alpha_{p_1}$ and factor 2 has p_2 levels $\gamma_1, \gamma_2, \ldots, \gamma_{p_2}$. Let $x_{ij}^{(m)}$ denote the (i,j) th element of the matrix X_m. We shall say that the observation Y_k denoting the k th element of Y, belongs to the i th level of factor m if $x_{ki}^{(m)}$ is 1, and it does belong to this level if it is 0 (m = 1,2). Then $n_{ij} = \sum_{\nu=1}^{n} x_{\nu i}^{(1)} x_{\nu j}^{(2)}$ is the number of observations having simultaneously factor 1 on the i th level and factor 2 on the j th level. Let N denote the $p_1 \times p_2$ matrix with elements n_{ij} .

If assumption (III) is satisfied, then the following theorem holds.

THEOREM 2. A_1 and A_2 commute it and only if the matrix N after suitable relabelling of the levels has for same t the form

$$\begin{bmatrix} N_1 & 0 & 0 & \ldots & 0 \\ 0 & N_2 & 0 & \ldots & 0 \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ 0 & 0 & 0 & \ldots & N_t \end{bmatrix}$$

where the N_i are matrices with all elements equal to an integer n_i .

Proof. Define $B = A_1 A_2$ and let b_{kh}, $a_{kh}^{(1)}$, $a_{kh}^{(2)}$ denote the elements of B, A_1 and A_2 respectively. Then $a_{kh}^{(m)} =$

$\sum_{\nu=1}^{p_m} x_{k\nu}^{(m)} x_{h\nu}^{(m)}$ (m = 1,2), and $b_{kh} = \sum_{\mu=1}^{n} a_{k\mu}^{(1)} a_{\mu h}^{(2)}$. It is seen that

$a_{k\mu}^{(m)} = 1$ if the observations Y_k and Y_μ belong to the same level of factor m, and $a_{k\mu}^{(m)} = 0$ if Y_k and Y_μ belong to different levels of factor m (m= 1,2). It follows that $a_{k\mu}^{(1)} a_{\mu h}^{(2)} = 1$ if Y_k and Y_μ belong to the same level of factor 1 and Y_μ and Y_h belong to the same level of factor 2, and $a_{k\mu}^{(1)} a_{\mu h}^{(2)} = 0$ otherwise. The sum $b_{kl} = \sum_{\mu=1}^{n} a_{k\mu}^{(1)} a_{\mu h}^{(2)}$ is then the number of observations which simultaneously has the same level of factor 1 as Y_k and the same level of factor 2 as Y_h.

Assume first that A_1 and A_2 commute or equivalently that B is symmetric. Let (α_k, γ_h) denote the k th level of factor 1 and the h th level of factor 2. Suppose that $n_{1i} > 0$ and $n_{1j} > 0$. Let Y_k be an observation from (α_1, γ_i) and Y_h an observation from (α_1, β_j). Then $b_{kh} = n_{1j}$ and $b_{hk} = n_{1i}$. If the matrix B is symmetric, then $b_{kh} = b_{hk}$ and hence $n_{1i} = n_{1i}$. It follows that all n_{1j} $(j=1,2,\ldots,p_2)$ not equal to 0 are identical. The same argument apply to all n_{ij} $(j=1,2,\ldots,p_2)$ for fixed i, and to all n_{ij} $(i=1,2,\ldots,p_1)$ for fixed j.

Suppose that $n_{ij} > 0$ and $n_{1j} = 0$. If for some s $n_{si} > 0$, then $n_{sj} = 0$. Suppose on the contrary that for some s we have $n_{si} > 0$ and $n_{sj} > 0$. In that case there exist observations X_k and X_h which are from (α_1, γ_i) and (α_s, γ_j) respectively. We have $b_{kh} = n_{1j}$ and $b_{hk} = n_{si}$. Symmetry implies $n_{si} = n_{1j}$. Hence $n_{si} = 0$ which is a contradiction. It follows that the form of the matrix N must be as stated in the theorem.

Conversely suppose that the matrix N has the given form. Let X_k and X_h be observations from (α_i, γ_j) and (α_r, γ_s) respectively. then $b_{kh} = n_{is}$ and $b_{hk} = n_{rj}$, and $n_{ij} > 0$

and $n_{rs} > 0$. If $n_{is} > 0$ then by the condition on N we have $n_{is} = n_{ij}$ and $n_{is} = n_{rs}$. We now have $n_{is} = n_{ij} = n_{rs} > 0$. By the condition on N we must have $n_{rj} = n_{is} = n_{ij} = n_{rs} > 0$. Hence $b_{kh} = b_{hk}$. If $n_{is} = 0$ then by the condition of N we have $n_{rj} = 0$, and hence $n_{is} = n_{rj} = 0$. Hence $b_{kh} = b_{hk}$.

From Theorem 1 and Theorem 2 it is seen that for what is usually called unbalanced models it is impossible to transform into a vector of independent observations by means of linear transformations. An exception is the one-way classification model, since then k = 1.

4. A MODIFICATION OF WALD'S METHOD.

When it is not possible to transform to independent random variables, we can use a modification of Wald's method which in this section is illustrated on the two-way classification model (2.2).

Define new variables by

$$\left.\begin{aligned} \mu^* &= \mu + \alpha_. + \beta_. + \gamma_{..} \\ \alpha_i^* &= \alpha_i - \alpha_. + \gamma_{i.} - \gamma_{..} \\ \beta_j^* &= \beta_j - \beta_. + \gamma_{.j} - \gamma_{..} \\ \gamma_{ij}^* &= \gamma_{ij} - \gamma_{i.} - \gamma_{.j} + \gamma_{..} \end{aligned}\right\} \quad \begin{aligned} i &= 1,2,\ldots,r, \\ j &= 1,2,\ldots,s, \end{aligned}$$

where $r\alpha_. = \sum_i \alpha_i$, $s\beta_. = \sum_j \beta_j$, $s\gamma_{i.} = \sum_j \gamma_{ij}$, $r\gamma_{..} = \sum_{i,j} \gamma_{ij}$, $r\gamma_{.j} = \sum_i \gamma_{ij}$. Independence of $\mu^*, \{\alpha_i^*\}, \{\beta_j^*\}$ and $\{\gamma_{ij}^*\}$ is either verified directly or seen from the general theory of fixed-effects models see Scheffé [6] (p.111).

Equation (2.2) may now be written

$$(4.1) \qquad Y_{ijk} = \mu^* + \alpha_i^* + \beta_j^* + \gamma_{ij}^* + e_{ijk}$$

where $\sum_i \alpha_i^* = \sum_j \beta_j^* = \sum_i \gamma_{ij}^* = \sum_j \gamma_{ij}^* = 0$. We shall in the following assume that we have at least one observation per cell i. e. $m_{ij} \geq 1$ for all i and j. Let $\hat{\alpha}^*$, $\{\hat{\alpha}_i^*\}$, $\{\hat{\beta}_j^*\}$ and $\{\hat{\gamma}_{ij}^*\}$ be the least square estimates of the corresponding parameters, when the latter are considered as fixed. The estimates are unique and defined by

$$\left.\begin{aligned} \hat{\alpha}_i^* &= y_{i\cdot\cdot} - y_{\cdot\cdot\cdot} \\ \hat{\beta}_j^* &= y_{\cdot j\cdot} - y_{\cdot\cdot\cdot} \\ \hat{\gamma}_{ij}^* &= y_{ij\cdot} - y_{i\cdot\cdot} - y_{\cdot j\cdot} + y_{\cdot\cdot\cdot} \end{aligned}\right\} \quad \begin{aligned} i &= 1,2,\ldots,r, \\ j &= 1,2,\ldots,s, \end{aligned}$$

where $m_{ij}\, y_{ij\cdot} = \sum_k y_{ijk}$, $s\, y_{i\cdot\cdot} = \sum_j y_{ij\cdot}$,

$r\, y_{\cdot j\cdot} = \sum_i y_{ij\cdot}$ and $rs\, y_{\cdot\cdot\cdot} = \sum_{i,j} y_{ij\cdot}$.

From (4.1) we now find

$$(4.2) \qquad \left.\begin{aligned} \hat{\alpha}_i^* &= \alpha_i^* + e_{i\cdot\cdot} - e_{\cdot\cdot\cdot} \\ \hat{\beta}_j^* &= \beta_j^* + e_{\cdot j\cdot} - e_{\cdot\cdot\cdot} \\ \hat{\gamma}_j^* &= \gamma_{ij}^* + e_{ij\cdot} - e_{i\cdot\cdot} - e_{\cdot j\cdot} + e_{\cdot\cdot\cdot} \end{aligned}\right\} \quad \begin{aligned} i &= 1,2,\ldots,r, \\ j &= 1,2,\ldots,s, \end{aligned}$$

with the same notation for the e's as for the y's.

Let $\hat{\gamma}^*$ be a vector of $(r-1)(s-1)$ linear independent elements of the set $\{\hat{\gamma}_{ij}^*\}$.

From (4.1) and (4.2) it is seen that the covariance matrix of $\hat{\gamma}^*$ is of the form $(B_1 \Delta_{AB} + B_2)\sigma^2$ where $\Delta_{AB} = \sigma_{AB}^2/\sigma^2$ and B_1 and B_2 are known matrices. It follows that

$$(4.3) \qquad \frac{\hat{\gamma}^{*\prime}(B_1 \Delta_{AB} + B_2)^{-1} \hat{\gamma}^*}{\sigma^2}$$

has a χ^2-distribution with $(r-1)(s-1)$ degrees of freedom.

It is known from matrix theory that there exsists a matrix P such that $PB_1P' = I$ and $PB_2P' = K$ where K is a diagonal matrix with diagonal elements $k_1, k_2, \ldots, k_{(r-1)(s-1)}$. Introduce $\tilde{\gamma}^* = P\hat{\gamma}^*$. The covariance matrix of $\tilde{\gamma}^*$ is then $(I\Delta_{AB} + K)\sigma^2$ and

$$(4.4) \qquad \begin{aligned} \hat{\gamma}^{*\prime}(B_1 \Delta_{AB} + B_2)^{-1}\hat{\gamma}^* &= \tilde{\gamma}^{*\prime}(I\Delta_{AB} + K)^{-1}\tilde{\gamma}^* \\ &= \sum_i \tilde{\gamma}_i^{*2}/(\Delta_{AB} + k_i) \end{aligned}$$

showing that (4.3) is a decreasing function of Δ_{AB}.

Define $Q = \sum_{i,j,k} (y_{ijk} - y_{ij.})^2$.

Then Q/σ^2 has a χ^2-distribution with $n-rs$ degrees of freedom and is independent of (4.3). It follows that

$$F(\Delta_{AB}) = \frac{n-rs}{(r-1)(s-1)} \frac{\hat{\gamma}^{*\prime}(B_1 \Delta_{AB} + B_2)^{-1}\hat{\gamma}^*}{Q}$$

has an F-distribution. Since (4.3) decreases with Δ_{AB} also $F(\Delta_{AB})$ decreases with Δ_{AB}, hence a confidence interval can be obtained in the usual way.

To test a hypothesis of the form $\Delta_{AB} \leq \Delta_o$ against $\Delta_{AB} > \Delta_o$, we reject when $F(\Delta_o)$ is larger than the upper α- point $f_{1-\alpha}$ of the corresponding F-distribution. From (4.4) and the covariance properties of $\tilde{\gamma}^*$ it is found that the power function of the test is given by

$$\beta(\Delta_{AB}) = P\left(\frac{n - rs}{(r-1)(s-1)} \; \frac{\sum_i \tilde{\sigma}_i^{*2}/(\Delta_o + k_i)}{Q} > f_{1-\alpha}\right)$$

$$\beta(\Delta_{AB}) = P\left(\frac{n-rs}{(r-1)(s-1)} \; \frac{\sum_i \frac{\Delta_{AB}+k_i}{\Delta_o + k_i} R_i^2}{Q} \quad f_{1-\alpha}\right),$$

where $R_1, R_2, \ldots, R_{(r-1)(s-1)}$ are independent χ^2-distributed random variables, with 1 degree of freedom. It is seen that the power increases with Δ_{AB}. Hence in particular the test is unbiased.

The vector $\hat{\alpha}^* = \left[\hat{\alpha}_1^*, \hat{\alpha}_2^*, \ldots, \hat{\alpha}_{r-1}^*\right]'$ can be used to make inference concerning σ_A^2. From (4.1) and (4.2) we find

$$\mathrm{Var}\ \hat{\alpha}_i^* = (s\sigma_A^2 + \sigma_{AB}^2)(r-1)/(rs) + \mathrm{Var}\ (e_{i\cdot\cdot} - e_{\cdot\cdot\cdot})$$

(4.5)

$$\mathrm{Cov}(\hat{\alpha}_i^*, \hat{\alpha}_j^*) = -(s\sigma_A^2 + \sigma_{AB}^2)/(rs) + \mathrm{Cov}(e_{i\cdot\cdot} - e_{\cdot\cdot\cdot}, e_{\cdot j\cdot} - e_{\cdot\cdot\cdot})$$

Hence the covariance matrix of $\hat{\alpha}^*$ may be written in the form $(C_1 \Delta_A + C_2)\sigma^2$ where $\Delta_A = (s\sigma_A^2 + \sigma_{AB}^2)/\sigma^2$.

Using the same method as above it can be proved that

$$G(\Delta_A) = \frac{n-rs}{r-1} \; \frac{\hat{\alpha}^{*\prime}(C_1\Delta_A + C_2)^{-1}\hat{\alpha}^*}{Q}$$

has an F-distribution, is a decreasing function of Δ_A, and hence may be used to find a confidence interval for Δ_A. To test the hypothesis $\Delta_A \leq \Delta_o$ against $\Delta_A > \Delta_o$ we reject when $G(\Delta_o)$ is larger than a constant, and as above it is seen that the power function increases with Δ_A.

Unless $\sigma_{AB}^2 = 0$ we cannot by this method make inferance about σ_A^2/σ^2 exclusively. The situation is of course the same when considering σ_B^2.

It may be noted that the elements of B_1, B_2, C_1 and C_2 are easily obtained. The elements of C_1 is given by (4.5), and

the elements of B_1 are found from the fact that $\operatorname{Var} \gamma_{ij}^{*} =$
$= \sigma_{AB}^{2}(r-1)(s-1)/(rs)$, $\operatorname{Cov}(\gamma_{ij}^{*}, \gamma_{ik}^{*}) = -\sigma_{AB}^{2}(r-1)/(rs)$,
$\operatorname{Cov}(\gamma_{ij}^{*}, \gamma_{hj}^{*}) = -\sigma_{AB}^{2}(s-1)/(rs)$, $\operatorname{Cov}(\gamma_{ij}^{*}, \gamma_{hk}^{*}) = -\sigma_{AB}^{2}/(rs)$,
$(i = 1,2,\ldots,r,\ j = 1,2,\ldots,s,\ j \neq k,\ i \neq h)$. The elements of
B_2 and C_2 are obtained from (4.2) and the definition of these matrices.
(For C_2 see in particular (4.5)).

Finally it should be noted that the tests reduces to the usual ones when the model is balanced.

REFERENCES

[1] CRUMP, S.L. (1951). The present status of variance component analysis. Biometrics 7 1-16.

[2] GRAYBILL, F.A. and HULTQUIST, R.A. (1961). Theorems concerning Eisenhart's model II. Ann.Math.Statist. 32 261-269.

[3] HERBACH, L.H. (1959). Properties of model II - Type analysis of variance tests, A: Optimum nature of the F-test for model II in the balanced case. Ann.Math.Statist. 30 939-959.

[4] PERLIS, S. (1952) . Theory of Matrices. Addison-Wesley, Reading, Massachusetts.

[5] PLACKETT, R.L. (1960). Models in the analysis of variance. J. Roy. Statist.Soc. Ser.B 22 195-209.

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

[7] WALD, A. (1947). A note on regression analysis. Ann.Math.Statist. 18 586-589.

ESp/elr

7/4-1967.