The Annals of Mathematical Statistics
1969, Vol. 40, No. 6, 2189–2194

EXPRESSION OF VARIANCE-COMPONENT ESTIMATORS AS LINEAR COMBINATIONS OF INDEPENDENT NONCENTRAL CHI-SQUARE VARIATES

By David A. Harville

Aerospace Research Laboratories

1. Introduction and summary. It is well known that any quadratic form in random variables whose joint distribution is nondegenerate multivariate normal is distributed as a linear combination of independent noncentral chi-square variables. Thus, when normality holds, quadratic estimators of the components of variance associated with Eisenhart's [3] Model II are so distributed, even when the design is unbalanced. The problem considered here is that of determining the appropriate linear combinations for given estimators and designs.

The expressing of quadratic variance-component estimators as linear combinations of independent noncentral chi-squares is useful, for several reasons, in studying the distributions of the estimators: (i) it permits application of results like those of Press [5] on the distributions of such linear combinations; (ii) it leads to Monte-Carlo techniques for approximating the distributions of these estimators which are more efficient than those heretofore used; and (iii) it may give some insight into the ways in which various types of imbalance affect the distributions of the estimators.

We now define the transformation to be used in expressing quadratic forms as linear combinations of independent noncentral chi-squares. Let $\boldsymbol{\beta}$ represent an $n \times n$ real symmetric matrix. Take $\mathbf{y}$ to be an $n \times 1$ random vector having the multivariate normal distribution with mean $\boldsymbol{\mu}$ and symmetric, nonsingular variance-covariance matrix $\mathbf{V}$. Clearly, the $n \times 1$ random vector $\mathbf{z}$ defined by the linear transformation

$$\mathbf{z} = \mathbf{W}'\mathbf{S}^{-1}\mathbf{C}'\mathbf{y};$$

where $\mathbf{C}$ is an $n \times n$ orthogonal matrix whose columns are eigenvectors of $\mathbf{V}$, $\mathbf{S}$ is an $n \times n$ diagonal matrix whose ith diagonal element is the square root of the eigenvalue corresponding to the eigenvector represented by the ith column of $\mathbf{C}$, and $\mathbf{W}$ is an $n \times n$ orthogonal matrix whose columns are eigenvectors of the necessarily symmetric matrix $\mathbf{P} = \mathbf{S}\mathbf{C}'\boldsymbol{\beta}\mathbf{C}\mathbf{S}$; has the multivariate normal distribution with

$$E\{\mathbf{z}\} = \mathbf{W}'\mathbf{S}^{-1}\mathbf{C}'\boldsymbol{\mu} \tag{1}$$

and variance-covariance matrix $\mathbf{I}$ (the identity matrix). Then the distribution of the quadratic form $\mathbf{y}'\boldsymbol{\beta}\mathbf{y}$ is the same as that of $\mathbf{z}'\mathbf{D}\mathbf{z}$, where $\mathbf{D} = \mathbf{W}'\mathbf{P}\mathbf{W}$ is an $n \times n$ diagonal matrix whose ith diagonal element is the eigenvalue of $\mathbf{P}$ corresponding to the eigenvector represented by the ith column of $\mathbf{W}$.

Received 12 August 1968; revised 17 February 1969.

It is clear from the above that, for any specified $\mathbf{V}$ and $\mathfrak{B}$, the expression of $\mathbf{y}'\mathfrak{B}\mathbf{y}$ as a linear combination of independent noncentral chi-squares can be accomplished by determining $\mathbf{C}$, $\mathbf{S}$, and $\mathbf{W}$. The linear combination so obtained can be readily converted into a linear combination of independent noncentral chi-squares having distinct coefficients. It follows from the "only if" part of Theorem 1 of Baldessari [1] that this latter expression is unique; i.e., for the specified $\mathbf{V}$, $\mathfrak{B}$ pair, there exists no other linear combination of independent noncentral chi-squares, with distinct coefficients, having the same distribution as $\mathbf{y}'\mathfrak{B}\mathbf{y}$.

In Sections 2–4, we study the determination of the $\mathbf{C}$, $\mathbf{S}$, and $\mathbf{W}$ matrices for estimators of the variance components associated with the one-way random classification. While the techniques used here for the one-way classification can also be applied to estimators of the components associated with higher classifications, the complexity of the problem is much greater for the more-complicated models. In Section 2, results that are applicable to any quadratic form in the one-way data are given. Some special results are also presented on those quadratic estimators having a certain invariance property to be called 'μ-invariance.' It is shown that an estimator exhibiting this property is distributed as a linear combination of independent central chi-squares, and consequently it is not necessary to find a $\mathbf{W}$ matrix for such an estimator. In Section 3, attention is restricted to a certain subclass of quadratic estimators which includes all those in common usage, and, in Section 4, we further restrict ourselves to some cases where the subclass numbers display a very simple type of imbalance.

2. General results. In the one-way random classification, data are taken as having the linear model

$$y_{ij} = \mu + \alpha_i + \epsilon_{ij}; \quad i = 1, \cdots, a; j = 1, \cdots, n_i .$$

We assume, without loss of generality, that $n_i = \nu_r$ whenever i is an element of the set $G_r = \{1 + \sum_{s=0}^{r-1} \zeta_s, \cdots, (\zeta_r - 1) + \sum_{s=0}^{r-1} \zeta_s, \sum_{s=0}^{r} \zeta_s\}$ where $\zeta_0 = 0$ and $\zeta_1, \cdots, \zeta_b, \nu_1, \cdots, \nu_b$ are positive integers having $\sum_{s=1}^{b} \zeta_s = a$ and $\nu_1 < \nu_2 < \cdots < \nu_b$. Note that ζ_r represents the number of n_i's that are equal to ν_r. μ is a fixed general mean, while the α_i and the ϵ_{ij} will be regarded as independent, normal random variables with zero means and variances $\sigma_\alpha^2 > 0$ and $\sigma_\epsilon^2 > 0$, respectively. It is assumed (for estimability purposes) that $a \geqq 2$ and that $n_i > 1$ for some i.

To apply the transformation described earlier to the problem of expressing a quadratic form in these data as a linear combination of independent noncentral chi-squares, we set $n = \sum_{s=1}^{b} \zeta_s \nu_s$ and

$$\mathbf{y}' = (y_{11}, \cdots, y_{1n_1}, \cdots, y_{a1}, \cdots, y_{an_a}).$$

Then, $\boldsymbol{\mu}' = (\mu, \cdots, \mu)$ and, denoting by $\mathbf{V}_i$ an $n_i \times n_i$ matrix having diagonal elements $\sigma_\alpha^2 + \sigma_\epsilon^2$ and off-diagonal elements σ_α^2, $\mathbf{V}$ is the matrix with $\mathbf{V}_1, \cdots, \mathbf{V}_a$ down its diagonal and zeros elsewhere. It is assumed that we are dealing with an arbitrary quadratic estimator; i.e., with arbitrary $\mathfrak{B}$.

Now that $\mathbf{V}$ has been specified, we proceed to determine appropriate $\mathbf{C}$ and $\mathbf{S}$ matrices.

Since each of the $\mathbf{V}_i$ has an intraclass-covariance structure, we have that the eigenvalues of V are (after reordering from the order corresponding to the original data vector y) $\sigma_\epsilon^2 + \nu_1\sigma_\alpha^2, \cdots, \sigma_\epsilon^2 + \nu_b\sigma_\alpha^2$, and σ_ϵ^2 with multiplicities $\zeta_1, \cdots, \zeta_b$ and $n - a$, respectively. Once the eigenvalues of V are known, a C matrix can easily be determined by following the procedure outlined in Section 32 of Browne [2]. Here, we take

$$(2) \qquad \mathbf{C} = \begin{bmatrix} \mathbf{C}_1 & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ & & & & \mathbf{C}_a \end{bmatrix};$$

where the $n_i \times n_i$ matrix

$$\mathbf{C}_i = \begin{bmatrix} 1/2^{\frac{1}{2}} & 1/6^{\frac{1}{2}} & \cdots & 1/[n_i(n_i - 1)]^{\frac{1}{2}} & 1/n_i^{\frac{1}{2}} \\ -1/2^{\frac{1}{2}} & 1/6^{\frac{1}{2}} & & 1/[n_i(n_i - 1)[^{\frac{1}{2}} & 1/n_i^{\frac{1}{2}} \\ 0 & -2/6^{\frac{1}{2}} & & 1/[n_i(n_i - 1)]^{\frac{1}{2}} & 1/n_i^{\frac{1}{2}} \\ 0 & 0 & & 1/[n_i(n_i - 1)]^{\frac{1}{2}} & 1/n_i^{\frac{1}{2}} \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ 0 & 0 & & 1/[n_i(n_i - 1)]^{\frac{1}{2}} & 1/n_i^{\frac{1}{2}} \\ 0 & 0 & & -(n_i - 1)/[n_i(n_i - 1)]^{\frac{1}{2}} & 1/n_i^{\frac{1}{2}} \end{bmatrix}.$$

Then,

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_1 & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ & & & & \mathbf{S}_a \end{bmatrix},$$

where $\mathbf{S}_i$ is an $n_i \times n_i$ diagonal matrix having diagonal elements $\sigma_\epsilon, \cdots, \sigma_\epsilon, (\sigma_\epsilon^2 + n_i\sigma_\alpha^2)^{\frac{1}{2}}$.

It will be convenient to partition $\boldsymbol{\beta}$, P, and W into submatrices. We take

$$\boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_{11} & \cdots & \boldsymbol{\beta}_{1a} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \boldsymbol{\beta}_{a1} & \cdots & \boldsymbol{\beta}_{aa} \end{bmatrix}$$

where $\boldsymbol{\beta}_{uv}$ has dimension $n_u \times n_v$. We denote the jsth element of $\boldsymbol{\beta}_{uv}$ by $\beta_{uj,vs}$. Similarly, we subdivide P and W into $n_u \times n_v$ submatrices $\mathbf{P}_{uv}$ and $\mathbf{W}_{uv}$ having elements $p_{uj,vs}$ and $w_{uj,vs}$. Also, we partition the random vector z into $n_i \times 1$ subvectors $\mathbf{z}_i$ with elements z_{ij} and use d_{ij} to denote the ijth diagonal element of D.

From (1), we now have that, when **C** is given by (2),

$$E\{z_{ij}\} = \mu \sum_{t=1}^{a} n_t^{\frac{1}{2}}(\sigma_\epsilon^2 + n_t\sigma_\alpha^2)^{-\frac{1}{2}} w_{tn_t,ij}$$

for any appropriate **W** matrix.

A quadratic estimator $\mathbf{y}'\boldsymbol{\beta}\mathbf{y}$ in the one-way normal data is said to be *μ-invariant* if, for every data vector and every constant, the estimate is not changed by adding the constant to each of the data, or, equivalently, if the variance of the estimator does not depend on the parameter μ (Harville, [4]). The class of 'μ-invariant' estimators includes the ordinary analysis-of-variance estimator of σ_α^2 as well as other commonly-used estimators.

REMARK 1. If $\mathbf{y}'\boldsymbol{\beta}\mathbf{y}$ is μ-invariant, then, when **C** is given by (2), the vector $\mathbf{q} = (\mathbf{q}_1', \cdots, \mathbf{q}_a')'$, whose $n_t \times 1$ subvector $\mathbf{q}_t$ has elements $q_{ts} = 0$, $s = 1, \cdots, n_t - 1$, $q_{tn_t} = n_t^{\frac{1}{2}}(\sigma_\epsilon^2 + n_t\sigma_\alpha^2)^{-\frac{1}{2}}$, is an eigenvector of **P**.

PROOF. Since a μ-invariant quadratic estimator has $\beta_{ij,\cdot\cdot} = 0$ (a dot in place of a subscript indicates summation over that subscript) for all i, j (Harville, [4]); the lemma is clear upon noting that, when **C** is given by (2),

$$p_{ij,tn_t} = \sigma_\epsilon(\sigma_\epsilon^2 + n_t\sigma_\alpha^2)^{\frac{1}{2}}[j(j+1)n_t]^{-\frac{1}{2}}[\textstyle\sum_{s=1}^{j} \beta_{is,t\cdot} - j\beta_{i(j+1),t\cdot}], j = 1, \cdots, n_i - 1,$$

$$p_{in_i,tn_t} = [(\sigma_\epsilon^2 + n_i\sigma_\alpha^2)(\sigma_\epsilon^2 + n_t\sigma_\alpha^2)/(n_in_t)]^{\frac{1}{2}}\beta_{i\cdot,t\cdot}. \quad \square$$

REMARK 2. If $\mathbf{y}'\boldsymbol{\beta}\mathbf{y}$ is μ-invariant, then it is distributed as a linear combination of independent central chi-square variables.

PROOF. Taking **C** to be given by (2), we can take the vector **q** (properly normalized) of Remark 1 to be a column, say the tsth column, of **W**. We have $d_{ts} = 0$ and; for all i, j pairs such that $i \neq t$ or, when $i = t, j \neq s$; $E\{z_{ij}\} = 0$. $\square$

It follows from Remark 2 and from Theorem 1 of Baldessari [1] that, if the quadratic estimator is known to be μ-invariant; then, no matter what our choice for **C**, there is no need to determine the matrix **W**. Rather, once an appropriate **C** matrix has been determined, to express the estimator as a linear combination of independent noncentral chi-squares requires only that we find the eigenvalues of **P**.

3. Special results for a restricted class of estimators. In this section, we restrict attention to those (not necessarily μ-invariant) quadratic estimators $\mathbf{y}'\boldsymbol{\beta}\mathbf{y}$ in the one-way data such that, for some real numbers ϕ, τ_u, θ_u, γ_{uv}; $u, v = 1, \cdots, b$; for which $\tau_u - \theta_u = \phi$ for all u; we have for $i \in G_u$, $t \in G_v$ with $t \neq i$; $\beta_{ir,ts} = \gamma_{uv}$, $\beta_{ir,ir} = \tau_u$, and, if $s \neq r$, $\beta_{ir,is} = \theta_u$. From Harville [4], we have that if a quadratic estimator does not have this form, then we can find a second quadratic estimator having the same expectation and uniformly smaller variance that does.

Still taking **C** to be given by (2), we proceed to give results on finding the eigenvalues of **P** and on determining an appropriate **W** matrix. Setting $\lambda_u = \tau_u + (\nu_u - 1)\theta_u$, and using the matrix equality $\mathbf{P}_{it} = \mathbf{S}_i\mathbf{C}_i'\boldsymbol{\beta}_{it}\mathbf{C}_t\mathbf{S}_t$; we now have

$$\mathbf{P} = \mathbf{K} + \phi\sigma_\epsilon^2\mathbf{I}$$

where

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_{11} & \cdots & \mathbf{K}_{1a} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \mathbf{K}_{a1} & \cdots & \mathbf{K}_{aa} \end{bmatrix},$$

while $\mathbf{K}_{it}$ represents an $n_i \times n_t$ matrix all of whose elements are zero except the lower-right-corner element which equals

$$e_u = \lambda_u(\sigma_\epsilon^2 + \nu_u\sigma_\alpha^2) - \phi\sigma_\epsilon^2$$

when $i = t \varepsilon G_u$ and equals

$$k_{uv} = [\nu_u\nu_v(\sigma_\epsilon^2 + \nu_u\sigma_\alpha^2)(\sigma_\epsilon^2 + \nu_v\sigma_\alpha^2)]^{\frac{1}{2}}\gamma_{uv}$$

when $i \varepsilon G_u$, $t \varepsilon G_v$ with $t \neq i$.

By an elementary matrix theorem, the eigenvalues of $\mathbf{P} = \mathbf{K} + \phi\sigma_\epsilon^2\mathbf{I}$ can be obtained by adding $\phi\sigma_\epsilon^2$ to the eigenvalues of $\mathbf{K}$. Clearly, $\mathbf{K}$ has the eigenvalue zero with multiplicity $n - a$, and its remaining eigenvalues are identical to those of the $a \times a$ matrix

$$\mathbf{L} = \mathbf{M} - \phi\sigma_\epsilon^2\mathbf{I}$$

where m_{it}, the itth element of the matrix $\mathbf{M}$, is given by $m_{it} = e_u + \phi\sigma_\epsilon^2$ for $i = t \varepsilon G_u$ and $m_{it} = k_{uv}$ for $i \varepsilon G_u$, $t \varepsilon G_v$ with $t \neq i$. Here, the eigenvalues of $\mathbf{L}$ can be obtained by subtracting $\phi\sigma_\epsilon^2$ from the eigenvalues of $\mathbf{M}$. By elementary matrix considerations, $\mathbf{M}$ has the eigenvalue $e_u + \phi\sigma_\epsilon^2 - k_{uu}$ with multiplicity $\zeta_u - 1$, $u = 1, \cdots, b$; and its remaining eigenvalues are indentical to the eigenvalues of the $b \times b$ matrix $\mathbf{N}$ having ith diagonal element

$$e_i + \phi\sigma_\epsilon^2 + (\zeta_i - 1)k_{ii}$$

and itth off-diagonal element $\zeta_t k_{it}$.

Denoting the eigenvalues of $\mathbf{N}$ by $\rho_1, \cdots, \rho_b$, it follows from the above discussion that the eigenvalues of $\mathbf{P} = \mathbf{K} + \phi\sigma_\epsilon^2\mathbf{I}$ are $\phi\sigma_\epsilon^2$, with multiplicity $n - a$; $e_u + \phi\sigma_\epsilon^2 - k_{uu}$, with multiplicity $\zeta_u - 1$, $u = 1, \cdots, b$; and $\rho_1, \cdots, \rho_b$, each with multiplicity one. If the quadratic estimator is μ-invariant, then it is clear that $\mathbf{N}$ is not of full rank and that one of its eigenvalues is zero.

If our quadratic estimator is not μ-invariant, we must find a $\mathbf{W}$ matrix corresponding to $\mathbf{P} = \mathbf{K} + \phi\sigma_\epsilon^2\mathbf{I}$. This matrix can be constructed from the eigenvalues of $\mathbf{P}$ by following a procedure like that outlined by Browne [2]. Here, we take $w_{ij,ij} = 1$, $i = 1, \cdots, a$, $j = 1, \cdots, n_i - 1$, and all other elements of $\mathbf{W}$, save w_{in_i,tn_t}, $i, t = 1, \cdots, a$, to be zero. We must still assign a value to the $a \times a$ matrix whose itth element is w_{in_i,tn_t}. Denoting this matrix by $\mathbf{W}^*$ and letting $\mathbf{X}$ represent any $b \times b$ matrix whose columns are real eigenvectors of $\mathbf{N}$ and whose elements x_{ji} satisfy

$$\sum_{j=1}^{b} \zeta_j x_{ji}^2 = 1 \quad \text{and}$$

$$\sum_{j=1}^{b} \zeta_j x_{ji} x_{jt} = 0$$

for every i, t with $t \neq i$ (the existence of $\mathbf{X}$ can be readily established); we take

$$\mathbf{W}^* = (\mathbf{X}^*, \mathbf{T}):$$

where, using $\mathbf{x}_{ji}^*$ to denote a $1 \times \zeta_j$ vector all of whose elements equal x_{ji} and letting $\mathbf{h}_j$ represent a j-dimensional row vector having all elements equal to $[j(j+1)]^{-\frac{1}{2}}$, the ith column of the $a \times b$ matrix $\mathbf{X}^*$ is $(\mathbf{x}_{1i}^*, \cdots, \mathbf{x}_{bi}^*)'$, and

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_1 & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ & & & & \mathbf{T}_b \end{bmatrix},$$

with the jth column of the $\zeta_i \times (\zeta_i - 1)$ matrix $\mathbf{T}_i$ being given by

$$(\mathbf{h}_j, -[j/(j+1)]^{\frac{1}{2}}, 0, \cdots, 0)'.$$

For this $\mathbf{W}$ matrix, we have, for $i = 1, \cdots, b$ with $j = n_i$,

$$E\{z_{ij}\} = \mu \sum_{t=1}^{b} \zeta_t \nu_t^{\frac{1}{2}} (\sigma_\epsilon^2 + \nu_t \sigma_\alpha^2)^{-\frac{1}{2}} x_{ti};$$

and, for all other i, j pairs, $E\{z_{ij}\} = 0$.

4. Special results for the case $b = 2$. In this section, we restrict attention to estimators that are of the type considered in the previous section and which in addition are μ-invariant, and look at the special case $b = 2$. We have, from Section 2, that these estimators can be expressed as linear combinations of independent central chi-square variables and, from Section 3, that the problem of determining the combination, when $b = 2$, reduces to that of finding the single nonzero eigenvalue of the 2×2 matrix $\mathbf{N}$.

From the characteristic equation for $\mathbf{N}$; we have, using the μ-invariance property, that this eigenvalue is

$$e_1 + e_2 + 2\phi\sigma_\epsilon^2 + (\zeta_1 - 1)k_{11} + (\zeta_2 - 1)k_{22} = -\gamma_{12}(n\sigma_\epsilon^2 + a\nu_1\nu_2\sigma_\alpha^2).$$

REFERENCES

[1] Baldessari, B. (1967). The distribution of a quadratic form of normal random variables. *Ann. Math. Statist.* **38** 1700–1704.

[2] Browne, E. T. (1958). *Introduction to the Theory of Determinants and Matrices.* Univ. of North Carolina Press.

[3] Eisenhart, C. (1947). The assumptions underlying the analysis of variance. *Biometrics* **3** 1–21.

[4] Harville, D. A. (1969). Quadratic unbiased estimation of variance components for the one-way classification. Biometrika **56** 313–326.

[5] Press, S. J. (1966). Linear combinations of non-central chi-square variates. *Ann. Math. Statist.* **37** 480–487.