The Annals of Statistics
1991, Vol. 19, No. 4, 1771–1785

EXACT CONFIDENCE SETS FOR VARIANCE COMPONENTS IN UNBALANCED MIXED LINEAR MODELS

BY ALAN P. FENECH AND DAVID A. HARVILLE[1]

University of California, Davis, and Iowa State University

We present a general procedure for obtaining an exact confidence set for the variance components in a mixed linear model. The procedure can be viewed as a generalization of the ANOVA method used with balanced models. Our procedure uses, as pivotal quantities, quadratic forms that are distributed independently as chi-squared variables. These quadratic forms are constructed with reference to spaces that are orthogonal with respect to the covariance matrix of the observation vector, which is a function of the variance components. For balanced models, these pivotal quantities simplify to multiples of the sums of squares used in the ANOVA method. An exact confidence set for the vector of ratios of the effect variances to the error variance is also presented, based on the same collection of quadratic forms.

Computing formulas for calculating approximations to these confidence sets are presented, and the results of their application to several two-way data sets are given.

1. Introduction. Suppose y is an $n \times 1$ observable vector that follows the general mixed linear model

$$y = X_0\beta_0 + X_1\beta_1 + \cdots + X_k\beta_k + \beta_{k+1}, \tag{1.1}$$

where X_i is an $n \times m_i$ known matrix $(i = 0, \ldots, k)$, β_0 is an $m_0 \times 1$ vector of unknown parameters and for $i = 1, \ldots, k+1$, β_i is an $m_i \times 1$ unobservable random vector whose distribution is $N(0, \sigma_i^2 I)$ where σ_i^2 is unknown. Here the random vectors $\beta_1, \ldots, \beta_{k+1}$ are jointly independent and $\sigma_i^2 \geq 0$, $1 \leq i \leq k$, $\sigma_{k+1}^2 > 0$. Define $\sigma = (\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2, \sigma_{k+1}^2)^t$, $\gamma_i = \sigma_i^2/\sigma_{k+1}^2$, $i = 1, \ldots, k+1$, and $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_k, \gamma_{k+1})^t$; here $\gamma_{k+1} \equiv 1$.

In this article we develop exact $100(1-\alpha)\%$ confidence sets for the vector σ of variance components and the vector γ of variance ratios. Such sets can be of direct interest; they can also be used to define generally conservative simultaneous confidence intervals for functions of σ, or of γ [Khuri (1981)].

Let $\chi^2(r)$ represent a central chi-squared random variable with r degrees of freedom. A confidence set for γ or σ is easy to construct if there exist $k+1$ known quadratic forms in y, y^tA_iy, $1 \leq i \leq k+1$, such that:

(i) $y^tA_1y, \ldots, y^tA_{k+1}y$, are jointly independent;

(ii) $y^tA_iy/\Delta_i^t\sigma \sim \chi^2(r_i)$ for a known positive integer r_i and a known linear form $\Delta_i^t\sigma = \sum_{j=1}^{k+1}\Delta_{ij}\sigma_j^2$;

(iii) the $k+1$ vectors Δ_i are linearly independent.

Received August 1988; revised August 1990.

[1]Research supported in part by the Office of Naval Research Contract N00014-85-K-0418.

AMS 1980 *subject classifications*. Primary 62J10; secondary 62F25.

Key words and phrases. Analysis of variance, confidence regions, mixed linear models, unbalanced data, variance components.

Assuming the existence of such quadratic forms and taking a_i and b_i to be constants chosen to satisfy $P\{a_i \leq \chi^2(r_i) \leq b_i\} = 1 - \alpha_i$, where $\prod_{i=1}^{k+1}(1 - \alpha_i) = 1 - \alpha$, a $100(1 - \alpha)\%$ confidence set for σ is the set $S(y\colon \sigma)$ that consists of all possible σ values that simultaneously satisfy the inequalities

$$a_i \leq y^t A_i y / \Delta_i^t \sigma \leq b_i, \qquad 1 \leq i \leq k + 1. \tag{1.2}$$

Further, suppose that $\Delta_{k+1} = (0, 0, 0, \ldots, 0, 1)^t$, so that $\Delta_{k+1}^t \sigma = \sigma_{k+1}^2$. Then an exact $100(1 - \alpha)\%$ confidence set $S(y\colon \gamma)$ for γ can be obtained as follows. Observing that the random vector $F = (F_1, F_2, \ldots, F_k)^t$, where

$$F_i = \frac{y^t A_i y / r_i}{y^t A_{k+1} y / r_{k+1}} \frac{1}{\Delta_i^t \gamma}$$

follows a known multivariate F distribution, and choosing the constants c_i and d_i to satisfy $P\{c_i \leq F_i \leq d_i,\ 1 \leq i \leq k\} = 1 - \alpha$, $S(y\colon \gamma)$ is the set of all possible γ vectors that simultaneously satisfy the inequalities

$$c_i \leq \frac{y^t A_i y / r_i}{y^t A_{k+1} y / r_{k+1}} \frac{1}{\Delta_i^t \gamma} \leq d_i, \qquad 1 \leq i \leq k. \tag{1.3}$$

The sets $S(y\colon \sigma)$ and $S(y\colon \gamma)$ are computationally easy to work with, as each is (with probability 1) a closed bounded polyhedron. See Khuri (1981) for a detailed discussion of confidence sets like these.

For balanced classificatory models, quadratic forms $y^t A_i y$ satisfying (i), (ii) and (iii) exist. They are the ordinary effect and residual sums of squares in the usual ANOVA table. These quadratic forms correspond to an orthogonal decomposition of $C(X_0)^\perp$, so their ranks necessarily sum to $n - \operatorname{rank}(X_0)$. Other cases where such a collection of quadratic forms exist have been found; see Graybill and Hultquist (1961), Broemeling (1969) and Broemeling and Bee (1976). Brown (1984) gave general conditions for determining the existence of quadratic forms $y^t A_i y$ satisfying (i), (ii) and (iii).

Unfortunately, for many models of the form (1.1), quadratic forms $y^t A_i y$ satisfying (i) through (iii) may, if they exist at all, be hard to find or may be such that they correspond to an orthogonal decomposition of a proper subspace of $C(X_0)^\perp$. In the following sections, we present a general procedure for forming an exact $100(1 - \alpha)\%$ confidence set for σ, or for γ, for a model of the form (1.1). It is based on $(k + 1)$ statistically independent quadratic forms in y that are central chi-squared in distribution. These quadratic forms are used as pivotal statistics; they differ from those just considered because they are allowed to depend on σ or on γ. By construction these quadratic forms correspond to an orthogonal decomposition of $C(X_0)^\perp$, with respect to the inner product defined by $\sum_1^k X_i X_i^t \sigma_i^2 + I\sigma_{k+1}^2$. Previously, Pincus (1977) and Tjur (1984) considered quadratic forms associated with various other orthogonal decompositions of $C(X_0)^\perp$. Our proposed confidence sets for σ and for γ are defined by sets of inequalities like (1.2) and (1.3). We give lemmas that indicate how to calculate these confidence sets and present the results of their application to two-way additive models. For balanced models, these sets are identical to those presented as $S(y\colon \sigma)$ and $S(y\colon \gamma)$ earlier.

In principle the likelihood ratio can be used to generate a confidence set for σ, or for γ. Taking T to be a matrix whose rows are a basis for $C(X_0)^{\perp}$, the likelihood function associated with Ty depends only on σ. An approximate confidence set for σ could be obtained from the likelihood ratio associated with Ty by assuming that this statistic follows its limiting chi-squared distribution. However, the quality of the approximation is generally unknown and the calculation of the confidence set is a substantial numerical problem. Another large-sample approach assumes that the restricted maximum likelihood estimator of σ or of γ (the estimator obtained by maximizing the likelihood function associated with Ty) follows a limiting multivariate normal distribution. The quality of this approximation is also generally unknown.

Finally, Hartley and Rao (1967), Section 9, define an exact $100(1-\alpha)\%$ confidence set for γ. If $k = 1$, this set, our proposed set, and Wald's confidence set [as generalized by Seely and El-Bassiouni (1983)] are essentially the same. If $k \geq 2$, Hartley and Rao's set and our set are both obtained from the same set of pivotal statistics, but use these statistics in very different ways, as we discuss subsequently.

2. The proposed confidence sets.

2.1. *For* σ.

Notation. Let $\Omega = \{\sigma\colon \sigma_i^2 \geq 0 \text{ for } 1 \leq i \leq k \text{ and } \sigma_{k+1}^2 > 0\}$. The family of models (1.1) can be expressed as $y \sim N(X_0\beta_0, \Sigma_1^k X_i X_i^t \sigma_i^2 + I\sigma_{k+1}^2)$ or $y \sim N(X_0\beta_0, \sigma_{k+1}^2 V)$, where $V = \Sigma_1^k X_i X_i^t \gamma_i + I$. For $0 \leq i \leq k$, define $X_i^* = (X_0 X_1, \dots, X_i)$ and let $X_{k+1}^* = I$. For $1 \leq i \leq k+1$, let $r_i = \operatorname{rank}(X_i^*) - \operatorname{rank}(X_{i-1}^*)$. We proceed to develop a confidence set for σ.

An orthogonal decomposition of $C(X_0)^{\perp}$. For $0 \leq i \leq k$, let U_i^* be the orthogonal complement of the column space of X_i^* with respect to the ordinary Euclidean inner product. Let $U_{k+1} = U_k^*$ and for $1 \leq i \leq k$ let U_i be the orthogonal complement of U_i^* within U_{i-1}^* with respect to the inner product defined by V: $U_i = \{x\colon x \in U_{i-1}^* \text{ and } x^t V y = 0\ \forall\, y \in U_i^*\}$. The dimension of U_i equals the difference in the dimensions of U_{i-1}^* and U_i^*, which is r_i.

Write

$$C(X_0)^{\perp} = U_1 \oplus U_2 \oplus \cdots \oplus U_k \oplus U_{k+1}. \tag{2.1}$$

Our development is only applicable if the dimension of each U_i exceeds 0, and we subsequently assume this to be the case; that is, we assume $r_i \geq 1$ for $1 \leq i \leq k+1$. By construction the spaces U_i are pairwise orthogonal with respect to $x^t V y$. Except for U_{k+1}, the spaces U_i can be regarded as depending formally on σ. (In fact, U_k is invariant to the choice of σ; and if the $TX_i X_i^t T^t$ commute in pairs, where the rows of T are an orthonormal basis of U_0^*, all the U_i are invariant to the choice of σ.)

The quadratic forms. For $1 \le i \le k+1$, let H_i be a matrix whose rows are a basis for U_i. Define for $1 \le i \le k+1$,

$$\tilde{F}_i(y, \sigma) = [H_i(\sigma)y]^t \big[H_i(\sigma)\sigma_{k+1}^2 V H_i(\sigma)^t\big]^{-1} [H_i(\sigma)y].$$

In this definition, we write $H_i(\sigma)$ for H_i to emphasize the dependence of H_i on $\sigma_{k+1}^2 V$ and hence on σ. From a geometric point of view, $\tilde{F}_i(y, \sigma)$ is the squared length of the projection of $(\sigma_{k+1}^2 V)^{-1} y$ onto U_i with respect to the inner product $x^t(\sigma_{k+1}^2 V)y$ so the function $\tilde{F}_i(y, \sigma)$ does not depend on the choice of H_i. These quadratic forms will be the pivotal statistics used to define our confidence set for σ.

We use the following result in establishing our confidence set for σ.

THEOREM 2.1. *For the true value of* σ,

(i) *the* $k+1$ *random variables* $\tilde{F}_i(y, \sigma)$ *are jointly independent and*
(ii) $\tilde{F}_i(y, \sigma) \sim \chi^2(r_i)$.

PROOF. Consider the spaces U_i defined using $x^t V y$. Recalling that the r_i rows of H_i are a basis for U_i,

$$\begin{aligned} H_i y &\sim N\big(H_i X_0 \beta_0, \sigma_{k+1}^2 H_i V H_i^t\big) \\ &\sim N\big(0, \sigma_{k+1}^2 H_i V H_i^t\big) \end{aligned}$$

(since the rows of H_i are in U_{i-1}^*, which is contained in U_0^*). By construction, the U_i for $1 \le i \le k+1$ are pairwise orthogonal with respect to V. So the vectors $H_i y$ for $1 \le i \le k+1$ are jointly independent. Now

$$\tilde{F}_i(y, \sigma) = [H_i y]^t \big[\sigma_{k+1}^2 H_i V H_i^t\big]^{-1} [H_i y]$$

is clearly central chi-squared in distribution with degrees of freedom r_i; and the $\tilde{F}_i(y, \sigma)$ are jointly independent. □

Before defining our confidence set for σ, note that $\tilde{F}_{k+1}(y, \sigma)$ is a familiar quantity. First, recalling that the rows of H_{k+1} are a basis for U_{k+1},

$$\begin{aligned} \tilde{F}_{k+1}(y, \sigma) &= [H_{k+1}y]^t \big[\sigma_{k+1}^2 H_{k+1} V H_{k+1}^t\big]^{-1} [H_{k+1}y] \\ &= [H_{k+1}y]^t \big[H_{k+1} H_{k+1}^t\big]^{-1} [H_{k+1}y] / \sigma_{k+1}^2, \end{aligned}$$

since $H_{k+1} X_i = 0$ for all i. This function is just the ordinary residual sum of squares obtained from the model (1.1) divided by σ_{k+1}^2. Letting

$$F_{k+1}(y) = [H_{k+1}y]^t \big[H_{k+1} H_{k+1}^t\big]^{-1} [H_{k+1}y],$$

$\tilde{F}_{k+1}(y, \sigma) = F_{k+1}(y)/\sigma_{k+1}^2$, the usual pivotal for obtaining an interval for σ_{k+1}^2.

The confidence set $S(y\colon \sigma)$. Supposing that the constants a_i and b_i are chosen so that $\prod_{i=1}^{k+1} P\{a_i \le \chi^2(r_i) \le b_i\} = 1 - \alpha$, the set

$$S(y\colon \sigma) = \big\{\sigma\colon a_i \le \tilde{F}_i(y, \sigma) \le b_i \text{ for } 1 \le i \le k+1\big\} \tag{2.2}$$

is an exact $100(1-\alpha)\%$ confidence set for σ.

2.2. *For* γ. Regarding the elements of H_i as functions of γ rather than σ, define for $1 \le i \le k+1$, $F_i(y, \gamma) = [H_i y]^t [H_i V H_i^t]^{-1} [H_i y]$. These quadratic forms will be used to define the pivotal statistics for our confidence set for γ. [Note that $\tilde{F}_i(y, \sigma) = F_i(y, \gamma)/\sigma_{k+1}^2$ for all i, so that $F_{k+1}(y, \gamma) = F_{k+1}(y)$, identified earlier as the ordinary residual sum of squares for model (1.1).]

The confidence set $S(y\colon \gamma)$. Consider for $1 \le i \le k$ the functions $G_i(y, \gamma) = (F_i(y, \gamma)/r_i)/(F_{k+1}(y)/r_{k+1})$. Since $\tilde{F}_i(y, \sigma) = F_i(y, \gamma)/\sigma_{k+1}^2$ for all i, Theorem 2.1 implies that for the true value of γ, the distribution of the random vector $(G_1(y, \gamma), G_2(y, \gamma), \dots, G_k(y, \gamma))^t$ is that of a multivariate F random vector $(F_1, F_2, \dots, F_k)^t$ with numerator degrees of freedom $r_1, r_2, \dots, r_k$ and denominator degrees of freedom r_{k+1}. Supposing the constants c_i and d_i are chosen so that $P\{c_i \le F_i \le d_i \text{ for } 1 \le i \le k\} = 1 - \alpha$, the set

$$S(y\colon \gamma) = \{\gamma\colon c_i \le G_i(y, \gamma) \le d_i \text{ for } 1 \le i \le k\} \tag{2.3}$$

is an exact $100(1-\alpha)\%$ confidence set for γ.

2.3. *Connections with the ANOVA method.* For $1 \le i \le k$, let W_i be the orthogonal complement of U_i^* with respect to U_{i-1}^* using the ordinary Euclidean inner product, and let $W_{k+1} = U_k^*$. Let T_i be a matrix whose rows are an orthonormal basis for W_i. Define $T = (T_1^t, \dots, T_{k+1}^t)^t$, so that the rows of T are an orthonormal basis for U_0^*.

Following Brown (1984), we define an ANOVA(σ) to be a partition of the quadratic form $(Ty)^t Ty$ into $y^t A_1 y + \cdots + y^t A_s y$, where $y^t A_1 y, \dots, y^t A_s y$ are s known quadratic forms in y such that:

(i) $y^t A_1 y, \dots, y^t A_s y$ are jointly independent;

(ii) $y^t A_i y / \Delta_i^t \sigma \sim \chi^2(f_i)$ for a known positive integer f_i and a known linear form $\Delta_i^t \sigma = \sum_{j=1}^{k+1} \Delta_{ij} \sigma_j^2$; and

(iii) the s vectors $\Delta_1, \dots, \Delta_s$ are distinct (beyond a known multiplier).

Brown showed that an ANOVA(σ) exists if and only if the k matrices $TX_i X_i^t T^t$ commute in pairs, in which case the s quadratic forms $y^t A_i y$, $1 \le i \le s$, are unique up to order.

It can be shown by construction that, if an ANOVA(σ) exists, then

$$\tilde{F}_i(y, \sigma) = \sum_{j \in I_i} y^t A_j y / \Delta_j^t \sigma$$

for some partition of the integers $1, \dots, s$ into $k+1$ sets $I_1, \dots, I_{k+1}$. If, in addition, $s = k+1$, then the s quadratic forms $y^t A_i y$, $1 \le i \le s$, are, aside from order, the $k+1$ sums of squares $(T_i y)^t T_i y$, $1 \le i \le k+1$. Thus, if an ANOVA(σ) exists and if $s = k+1$, then the proposed confidence sets $S(y\colon \sigma)$ and $S(y\colon \gamma)$ are the same as those produced by the ANOVA method.

3. Calculating $S(y\colon \sigma)$ and $S(y\colon \gamma)$. In this section we present results which facilitate the calculation of $S(y\colon \sigma)$ and $S(y\colon \gamma)$. These results exploit the fact that the functions $\tilde{F}_i(y, \sigma)$ and $F_i(y, \gamma)$ depend respectively for y fixed only on σ_j^2, $j \ge i$, and γ_j, $j \ge i$.

Preliminary observations. Recall that $U_{k+1} = U_k^*$ and that for $1 \le i \le k$, U_i is the complement of U_i^* within U_{i-1}^* with respect to the inner product associated with V. Define $V_k = I$ and for $1 \le i \le k-1$,

$$V_i = \sum_{j=i+1}^{k} X_j X_j^t \gamma_j + I.$$

LEMMA 3.1.

(i) *For $1 \le i \le k$, U_i is the complement of U_i^* within U_{i-1}^* with respect to the inner product defined by V_i.*

(ii) *For $1 \le i \le k$, $F_i(y, \gamma)$ depends on γ only through values of $\gamma_k, \gamma_{k-1}, \dots, \gamma_i$.*

(iii) *For $1 \le i \le k+1$, $\tilde{F}_i(y, \sigma)$ depends on σ only through values of $\sigma_{k+1}^2, \sigma_k^2, \dots, \sigma_i^2$.*

PROOF. For $y \in U_i^*$, $x^t V y = x^t V_i y$, so that (i) is true.

Recall that $F_i(y, \gamma) = [H_i y]^t [H_i V H_i^t]^{-1} [H_i y]$, where the rows of H_i are a basis for U_i. Since the rows of H_i are in U_{i-1}^*, $H_i V H_i^t = H_i(X_i X_i^t \gamma_i + V_i) H_i^t$. Therefore,

$$F_i(y, \gamma) = [H_i y]^t \big[H_i\big(X_i X_i^t \gamma_i + V_i\big) H_i^t\big]^{-1} [H_i y]. \tag{3.1}$$

For $i = k$, U_i and V_i are invariant to γ; for $i < k$, U_i and V_i depend on γ only through $\gamma_{i+1}, \dots, \gamma_k$. So (ii) is true.

Suppose $1 \le i \le k$. Then

$$\begin{aligned} \tilde{F}_i(y, \sigma) &= F_i(y, \gamma)/\sigma_{k+1}^2 \\ &= [H_i y]^t \big[H_i\big(X_i X_i^t \sigma_i^2 + \sigma_{k+1}^2 V_i\big) H_i^t\big]^{-1} [H_i y]. \end{aligned} \tag{3.2}$$

For $i = k$, U_i and V_i are invariant to σ; for $1 \le i \le k$, U_i and $\sigma_{k+1}^2 V_i$ depend on σ only through $\sigma_{i+1}^2, \dots, \sigma_k^2, \sigma_{k+1}^2$. And $\tilde{F}_{k+1}(y, \sigma) = F_{k+1}(y)/\sigma_{k+1}^2$. So (iii) is true. □

In light of Lemma 3.1, we subsequently write $F_i(y, \gamma_k, \dots, \gamma_i)$ and $\tilde{F}_i(y, \sigma_{k+1}^2, \dots, \sigma_i^2)$ for $F_i(y, \gamma)$ and $\tilde{F}_i(y, \sigma)$, respectively.

Alternative descriptions of $S(y\colon \gamma)$ and $S(y\colon \sigma)$. Results (3.1) and (3.2) lead to a sequential approach to calculating $S(y\colon \sigma)$ and $S(y\colon \gamma)$. Consider for example $S(y\colon \gamma)$. Letting $\hat{\sigma}_{k+1}^2 = F_{k+1}(y)/r_{k+1}$, define $S_i(y\colon \gamma) = \{\gamma\colon c_i r_i \hat{\sigma}_{k+1}^2 \le F_i(y, \gamma_k, \dots, \gamma_i) \le d_i r_i \hat{\sigma}_{k+1}^2\}$. Then $S(y\colon \gamma) = \bigcap_1^k S_i(y\colon \gamma)$.

Note that, according to (3.1),

$$F_k(y, \gamma_k) = (H_k y)^t \{H_k X_k X_k^t H_k^t \gamma_k + H_k H_k^t\}^{-1} (H_k y), \tag{3.3}$$

where H_k is any matrix whose rows are a basis for U_k. Recall that U_k is independent of γ, so γ_k appears in (3.3) only as a coefficient of the symmetric positive definite matrix $H_k X_k X_k^t H_k^t$. Choosing H_k to satisfy $H_k H_k^t = I$, there exists an orthogonal $r_k \times r_k$ matrix O_k and a diagonal positive definite matrix D_k such that

$$F_k(y, \gamma_k) = (O_k H_k y)^t \{D_k \gamma_k + I\}^{-1} (O_k H_k y). \tag{3.4}$$

It follows that the values of γ_k represented in the set

$$S_k(y\colon \gamma) = \{\gamma\colon c_k r_k \hat{\sigma}^2_{k+1} \le F_k(y, \gamma_k) \le d_k r_k \hat{\sigma}^2_{k+1}\}$$

form an interval. As discussed by Harville and Fenech (1985), the end points of this interval can be computed efficiently by making use of representation (3.4).

Note also that

$$\begin{aligned} &F_{k-1}(y, \gamma_k, \gamma_{k-1}) \\ &\quad = (H_{k-1} y)^t \{H_{k-1} X_{k-1} X^t_{k-1} H^t_{k-1} \gamma_{k-1} + H_{k-1} V_{k-1} H^t_{k-1}\}^{-1} (H_{k-1} y). \end{aligned} \tag{3.5}$$

Here the rows of H_{k-1} are any basis for U_{k-1}, which depends at most on γ_k, while V_{k-1} depends only on γ_k. And so γ_{k-1} enters (3.5) only as a coefficient of the symmetric positive definite matrix $H_{k-1} X_{k-1} X^t_{k-1} H^t_{k-1}$. Choosing H_{k-1} to satisfy $H_{k-1} V_{k-1} H^t_{k-1} = I$, there exists an orthogonal $r_{k-1} \times r_{k-1}$ matrix O_{k-1} and a diagonal positive definite matrix D_{k-1} such that

$$F_{k-1}(y, \gamma_k, \gamma_{k-1}) = (O_{k-1} H_{k-1} y)^t \{D_{k-1} \gamma_{k-1} + I\}^{-1} (O_{k-1} H_{k-1} y). \tag{3.6}$$

For γ_k fixed, γ_{k-1} enters (3.6) only as a coefficient of D_{k-1}. And so the values of the vector (γ_k, γ_{k-1}) represented in the set $S_k(y\colon \gamma) \cap S_{k-1}(y\colon \gamma)$ are those obtained by associating with each γ_k represented in $S_k(y\colon \gamma)$, the interval of γ_{k-1} values defined by

$$\begin{aligned} c_{k-1} r_{k-1} \hat{\sigma}^2_{k+1} &\le (O_{k-1} H_{k-1} y)^t \{D_{k-1} \gamma_{k-1} + I\}^{-1} (O_{k-1} H_{k-1} y) \\ &\le d_{k-1} r_{k-1} \hat{\sigma}^2_{k+1}. \end{aligned}$$

By continuing in this way, this nested description can be extended to any of the sets

$$S_k(y\colon \gamma) \cap S_{k-1}(y\colon \gamma) \cap \cdots \cap S_i(y\colon \gamma), \qquad 1 \le i \le k,$$

including in the case $i = 1$, $S(y\colon \gamma)$. These descriptions lead in a natural way to an algorithm—to be described at the end of this section—for calculating these sets.

A similar computationally useful description of the set $S(y\colon \sigma)$ is as follows. Letting $S_i(y\colon \sigma) = \{\sigma\colon a_i \le \tilde{F}_i(y, \sigma^2_{k+1}, \sigma^2_k, \dots, \sigma_i^2) \le b_i\}$, $S(y\colon \sigma) = \bigcap_1^{k+1} S_i(y\colon \sigma)$. Note that the values of σ^2_{k+1} represented in the set

$$S_{k+1}(y\colon \sigma) = \left\{\sigma\colon a_{k+1} \le \frac{F_{k+1}(y)}{\sigma^2_{k+1}} \le b_{k+1}\right\}$$

consist of the interval $F_{k+1}(y)/b_{k+1} \le \sigma^2_{k+1} \le F_{k+1}(y)/a_{k+1}$. Now $S_k(y\colon \sigma) = \{\sigma\colon a_k \le \tilde{F}_k(y, \sigma^2_{k+1}, \sigma^2_k) \le b_k\}$ or equivalently $S_k(y\colon \sigma) = \{\sigma\colon a_k \sigma^2_{k+1} \le F_k(y, \gamma_k) \le b_k \sigma^2_{k+1}\}$. And so the values of $(\sigma^2_{k+1}, \sigma^2_k)$ represented in the set $S_{k+1}(y\colon \sigma) \cap S_k(y\colon \sigma)$ are those obtained by associating with each σ^2_{k+1}

represented in $S_{k+1}(y\colon \sigma)$, the interval of values of σ_k^2 defined by $a_k\sigma_{k+1}^2 \le F_k(y, \gamma_k) \le b_k\sigma_{k+1}^2$. Further,

$$\begin{aligned} S_{k-1}(y\colon \sigma) &= \left\{\sigma\colon a_{k-1} \le \tilde{F}_{k-1}\left(y, \sigma_{k+1}^2, \sigma_k^2, \sigma_{k-1}^2\right) \le b_{k-1}\right\} \\ &= \left\{\sigma\colon a_{k-1}\sigma_{k+1}^2 \le F_{k-1}(y, \gamma_k, \gamma_{k-1}) \le b_{k-1}\sigma_{k+1}^2\right\}. \end{aligned}$$

And so the values of $(\sigma_{k-1}^2, \sigma_k^2, \sigma_{k+1}^2)$ represented in $S_{k-1}(y\colon \sigma) \cap S_k(y\colon \sigma) \cap S_{k+1}(y\colon \sigma)$ are those obtained by associating with each $(\sigma_{k+1}^2, \sigma_k^2)$ represented in $S_{k+1}(y\colon \sigma) \cap S_k(y\colon \sigma)$, the interval of σ_{k-1}^2 values defined by

$$a_{k-1}\sigma_{k+1}^2 \le F_{k-1}(y, \gamma_k, \gamma_{k-1}) \le b_{k-1}\sigma_{k+1}^2.$$

All sets of the form $\bigcap_{j=i}^{k+1} S_j(y\colon \sigma)$ admit such a nested description, including as the special case $i = 1$, $S(y\colon \sigma)$.

Two results on calculations. To exploit the descriptions of the sets $S(y\colon \sigma)$ and $S(y\colon \gamma)$ given in the preceding subsection, we require a representation of the function $F_i(y, \gamma_k, \gamma_{k-1}, \dots, \gamma_i)$ that makes simple the calculation of the interval of γ_i values satisfying $l \le F_i(y, \gamma_k, \gamma_{k-1}, \dots, \gamma_i) \le u$ for $y, \gamma_k, \dots, \gamma_{i+1}$ fixed. The following theorem provides the requisite representation.

THEOREM 3.1. *For $1 \le i \le k$, let $P[X_{i-1}^*] = X_{i-1}^*(X_{i-1}^{*t}V_i^{-1}X_{i-1}^*)^- X_{i-1}^{*t}V_i^{-1}$, and define $T_i = X_i^t V_i^{-1}[I - P[X_{i-1}^*]]$ and $C_i = T_i X_i$.*

(i) *The matrix C_i is symmetric nonnegative definite of* rank r_i.

(ii) *Taking O_i to be an $m_i \times r_i$ matrix and D_i an $r_i \times r_i$ diagonal matrix such that $C_i = O_i D_i O_i^t$, where $O_i^t O_i = I$,*

$$F_i(y, \gamma_k, \gamma_{k-1}, \dots, \gamma_i) = \left(D_i^{-1/2}O_i^t T_i y\right)^t \{D_i\gamma_i + I\}^{-1}\left(D_i^{-1/2}O_i^t T_i y\right). \tag{3.7}$$

COMMENTS. C_k and T_k are independent of γ, and, for $1 \le i < k$, C_i and T_i depend on γ only through $\gamma_{i+1}, \dots, \gamma_k$. So γ_i enters expression (3.7) for $F_i(y, \gamma_k, \gamma_{k-1}, \dots, \gamma_i)$ only as a coefficient of the matrix D_i. The matrix $P[X_{i-1}^*]$ is the projection matrix for $C(X_{i-1}^*)$ with respect to the inner product $x^t V_i y$; see Lemma A1.

PROOF. It follows from $C_i = X_i^t(I - P[X_{i-1}^*])V_i^{-1}(I - P[X_{i-1}^*])X_i$ that C_i is symmetric nonnegative definite of rank r_i. Now,

$$T_i V T_i^t = T_i\{X_i X_i^t \gamma_i + V_i\}T_i^t = T_i X_i X_i^t T_i^t \gamma_i + T_i V_i T_i^t = C_i^2\gamma_i + C_i.$$

So $(D_i^{-1/2}O_i^t)(T_i V T_i^t)(O_i D_i^{-1/2}) = D_i\gamma_i + I$, implying that the rows of the $r_i \times n$ matrix $D_i^{-1/2}O_i^t T_i$ are linearly independent. Further, a direct application of Lemma A 2 indicates that the rows of T_i span U_i, and hence that the rows of $D_i^{-1/2}O_i^t T_i$ are in U_i. Thus, the rows of $D_i^{-1/2}O_i^t T_i$ are a basis for U_i. We conclude that

$$F_i(y, \gamma_k, \gamma_{k-1}, \dots, \gamma_i) = \left(D_i^{-1/2}O_i^t T_i y\right)^t \{D_i\gamma_i + I\}^{-1}\left(D_i^{-1/2}O_i^t T_i y\right). \qquad \square$$

In using expression (3.7) to evaluate $F_i(y, \gamma_k, \gamma_{k-1}, \dots, \gamma_i)$ (for $y, \gamma_k, \dots, \gamma_{i+1}$ fixed), we form the vector $T_i y$ and the matrix C_i from the expressions given in the following theorem.

THEOREM 3.2. *For fixed i, $1 \le i \le k$, let (H, a) be any solution to the system*

$$(X_{i-1}^{*t} V_i^{-1} X_{i-1}^*)(H, a) = (X_{i-1}^{*t} V_i^{-1} X_i, X_{i-1}^{*t} V_i^{-1} y). \tag{3.8}$$

Then $C_i = X_i^t V_i^{-1} X_i - X_i^t V_i^{-1} X_{i-1}^* H$ and $T_i y = X_i^t V_i^{-1} y - X_i^t V_i^{-1} X_{i-1}^* a$.

PROOF. That a solution to system (3.8) exists follows from well-known results on generalized least squares equations. Suppose that (H, a) is any solution to this system. Then $X_{i-1}^{*t} V_i^{-1}(X_i - X_{i-1}^* H) = 0$, which since $X_i = X_{i-1}^* H + (X_i - X_{i-1}^* H)$ implies that $X_{i-1}^* H = P[X_{i-1}^*] X_i$. By Lemma A1 then, $X_{i-1}^* H = X_{i-1}^*(X_{i-1}^{*t} V_i^{-1} X_{i-1}^*)^- X_{i-1}^{*t} V_i^{-1} X_i$. Substituting this expression for $X_{i-1}^* H$ into $X_i^t V_i^{-1} X_i - X_i^t V_i^{-1} X_{i-1}^* H$, we obtain

$$X_i^t V_i^{-1} X_i - X_i^t V_i^{-1} P[X_{i-1}^*] X_i = C_i.$$

A similar argument shows that $T_i y = X_i^t V_i^{-1} y - X_i^t V_i^{-1} X_{i-1}^* a$. □

To evaluate C_i and $T_i y$ (for $y, \gamma_k, \dots, \gamma_{i+1}$ fixed) from the expressions given in Theorem 3.2, we first evaluate the matrix

$$X_i^{*t} V_i^{-1}(X_i^*, y) = \begin{pmatrix} X_{i-1}^{*t} V_i^{-1} X_{i-1}^* & X_{i-1}^{*t} V_i^{-1} X_i & X_{i-1}^{*t} V_i^{-1} y \\ X_i^t V_i^{-1} X_{i-1}^* & X_i^t V_i^{-1} X_i & X_i^t V_i^{-1} y \end{pmatrix}.$$

The following theorem gives a formula that can be used to evaluate $X_i^{*t} V_i^{-1}(X_i^*, y)$ recursively.

THEOREM 3.3. *For $1 \le j \le k-1$,*

$$\begin{aligned} &X_j^{*t} V_j^{-1}(X_j^*, y) \\ &\quad = X_j^{*t} V_{j+1}^{-1}(X_j^*, y) \\ &\qquad - \gamma_{j+1} X_j^{*t} V_{j+1}^{-1} X_{j+1}\left(I + \gamma_{j+1} X_{j+1}^t V_{j+1}^{-1} X_{j+1}\right)^{-1} X_{j+1}^t V_{j+1}^{-1}(X_j^*, y). \end{aligned} \tag{3.9}$$

PROOF. Identity (3.9) is an immediate consequence of the identity

$$V_j^{-1} = V_{j+1}^{-1} - \gamma_{j+1} V_{j+1}^{-1} X_{j+1}\left(I + \gamma_{j+1} X_{j+1}^t V_{j+1}^{-1} X_{j+1}\right)^{-1} X_{j+1}^t V_{j+1}^{-1}$$

[Henderson and Searle (1981)]. □

Since $V_k = I$, formula (3.9) reduces, in the special case $j = k-1$, to

$$\begin{aligned} &X_{k-1}^{*t} V_{k-1}^{-1}(X_{k-1}^*, y) \\ &\quad = X_{k-1}^{*t}(X_{k-1}^*, y) - \gamma_k X_{k-1}^{*t} X_k\left(I + \gamma_k X_k^t X_k\right)^{-1} X_k^t(X_{k-1}^*, y). \end{aligned} \tag{3.10}$$

By using formula (3.10), $X_{k-1}^{*t} V_{k-1}^{-1}(X_{k-1}^*, y)$ can be determined from the array

$X_k^{*t}(X_k^*, y)$. More generally, by repeatedly using formula (3.9) (with $j = k-1, \dots, i$), $X_i^{*t}V_i^{-1}(X_i^*, y)$ can be computed recursively (in $k - i$ steps) from $X_k^{*t}(X_k^*, y)$.

Algorithms. The proposed algorithms operate on the elements of the array $X_k^{*t}(X_k^*, y)$ and produce the set $S(y\colon \gamma)$ in k steps and the set $S(y\colon \sigma)$ in $k+1$ steps.

The first step of the algorithm for producing $S(y\colon \gamma)$ consists of (i) computing C_k and $T_k y$ from the formulas of Theorem 3.2, (ii) using the iterative algorithm of Harville and Fenech (1985), in conjunction with representation (3.4) for $F_k(y, \gamma_k)$, to compute the lower and upper end points, say l_k and u_k, of the interval of γ_k-values that satisfy the inequality $c_k r_k \hat{\sigma}_{k+1}^2 \le F_k(y, \gamma_k) \le d_k r_k \hat{\sigma}_{k+1}^2$, and (iii) imposing a grid of equally spaced values of γ_k, say $\gamma_k(1) \le \cdots \le \gamma_k(s_k)$, on the interval $[l_k, u_k]$.

The second step of the algorithm consists, for the ith of the γ_k-values $\gamma_k(1), \dots, \gamma_k(s_k)$, of (i) computing $X_{k-1}^{*t}V_{k-1}^{-1}(X_{k-1}^*, y)$ from formula (3.10) and then computing C_{k-1} and $T_{k-1}y$ from the formulas of Theorem 3.2, (ii) using Harville and Fenech's iterative algorithm, in conjunction with representation (3.6) for $F_{k-1}(y, \gamma_k, \gamma_{k-1})$, to compute the end points, say $l_{k-1}(i)$ and $u_{k-1}(i)$, of the interval of γ_{k-1}-values that satisfy the inequality

$$c_{k-1} r_{k-1} \hat{\sigma}_{k+1}^2 \le F_{k-1}(y, \gamma_k(i), \gamma_{k-1}) \le d_{k-1} r_{k-1} \hat{\sigma}_{k+1}^2,$$

and (iii) imposing a grid of equally spaced values of γ_{k-1}, say $\gamma_{k-1}(i, 1) \le \cdots \le \gamma_{k-1}[i, s_{k-1}(i)]$, on the interval $[l_{k-1}(i), u_{k-1}(i)]$, $i = 1, \dots, s_k$.

Continuing in this way, we obtain after $k-1$ steps, a grid of $(k-1)$-dimensional points $[\gamma_2(i_2, \dots, i_k), \gamma_3(i_3, \dots, i_k), \dots, \gamma_k(i_k)]$, where $1 \le i_2 \le s_2(i_3, \dots, i_k), 1 \le i_3 \le s_3(i_4, \dots, i_k), \dots, 1 \le i_k \le s_k$, that covers the set of $(\gamma_2, \dots, \gamma_k)$-values represented in the set $\bigcap_2^k S_i(y\colon \gamma)$. The kth and final step of the algorithm consists, for the point with index $i_2, \dots, i_k$, of (i) computing $X_1^{*t}V_1^{-1}(X_1^*, y)$ recursively from formula (3.9) and then computing C_1 and $T_1 y$ from the formulas of Theorem 3.2 and (ii) using Harville and Fenech's iterative algorithm, in conjunction with representation (3.7), to compute the end points of the interval of γ_1-values that satisfy the inequality

$$c_1 r_1 \hat{\sigma}_{k+1}^2 \le F_1(y, \gamma_k(i_k), \gamma_{k-1}(i_{k-1}, i_k), \dots, \gamma_2(i_2, \dots, i_k), \gamma_1) \le d_1 r_1 \hat{\sigma}_{k+1}^2,$$
$$i_2 = 1, \dots, s_2(i_3, \dots, i_k), i_3 = 1, \dots, s_3(i_4, \dots, i_k), \dots, i_k = 1, \dots, s_k.$$

An algorithm for producing $S(y\colon \sigma)$ can be obtained by modifying the algorithm for producing $S(y\colon \gamma)$. The output of these algorithms can be used to display graphically $S(y\colon \sigma)$ or $S(y\colon \gamma)$, as illustrated in Section 4, or projections of $S(y\colon \sigma)$ or $S(y\colon \gamma)$.

4. An example. In this section we illustrate the nature of the set $S(y\colon \gamma)$ by applying our algorithm to the subfamily of mixed models consisting of two-way additive random effects models:

$$y_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk}, \tag{4.1}$$

TABLE 1
Cell sample size

Data set II

2	3	1	4		1	3	1			1	3
	3	2	1		4	2	1	3		1	4
	1			1	1	3	3	4	4	3	4
					2	1	4	4	4	3	4
1	1				1	2	3	4	4	4	4
					1	4	3	4	4	4	4
2		3	4	1	2	1	3		4	1	2
				1	1	2	1		1	1	1
				2	3	1		1	2	1	1

Data set III

3	2										
	1	4									
		3	3								
			2	4							
				1	4						
					3	3					
						4	2				
							1	1			
								1	2	3	1

where μ is an unknown constant, the α_i, β_j and e_{ijk} are jointly independent normal variables with zero means and $\mathrm{var}(\alpha_i) = \sigma_1^2 \geq 0$, $\mathrm{var}(\beta_j) = \sigma_2^2 \geq 0$, $\mathrm{var}(e_{ijk}) = \sigma_3^2 > 0$. Here $1 \leq i \leq I$, $1 \leq j \leq J$ and $1 \leq k \leq m_{ij}$; if $m_{ij} = 0$, cell (ij) in the two-way table is empty. Let $n = \Sigma\Sigma\, m_{ij}$. Assume that the design is connected. Writing (4.1) in the form (1.1) with the random effects ordered as in (4.1), $\dim(U_1) = I - 1$, $\dim(U_2) = J - 1$ and $\dim(U_3) = n - I - J + 1$. This subfamily (4.1) covers a wide range of mixed models, from balanced models which have no empty cells to sparse models with few occupied cells.

For model (4.1), consider the vector $\gamma = (\gamma_1, \gamma_2)^t$, where $\gamma_i = \sigma_i^2/\sigma_3^2$. We computed $S(y\colon \gamma)$ without complication for various simulated data sets. We present graphs of the results for three such data sets.

The first data set, called data set I, was generated by a simulation program from model (4.1) taking $\gamma_1 = 5$, $\gamma_2 = 2$, $I = 9$, $J = 12$ and $m_{ij} \equiv 4$. By deleting cases from this data set, two other data sets were obtained (data sets II and III). The cell sizes for data sets II and III are given in Table 1 (where a blank is to be interpreted as an empty cell).

Figure 1 displays the confidence set $S(y\colon \gamma)$ for data sets I, II and III; each set has a confidence level of 81%. The horizontal boundaries of all three confidence sets are straight lines, as follows from the definition of $S(y\colon \gamma)$.

If—as in the case of data set I—the data are balanced (i.e., $m_{ij} \equiv m$), then the confidence set $S(y\colon \gamma)$ is the same as that produced by the ANOVA

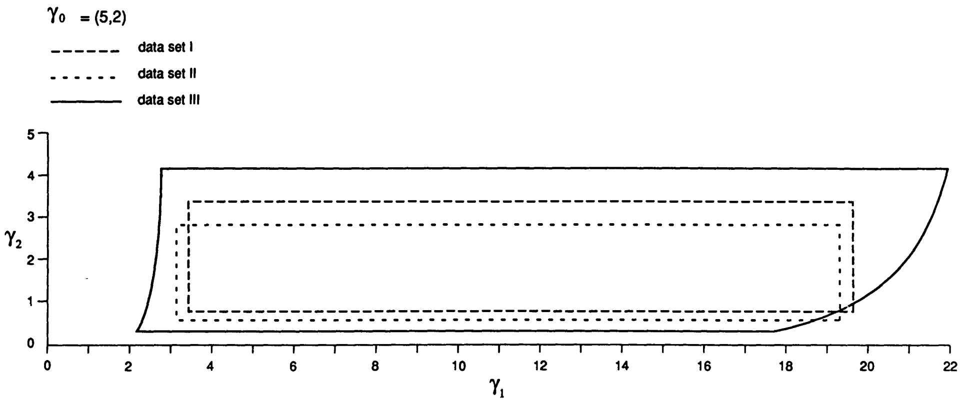

FIG. 1.

method. The sum of squares in the ANOVA are $SS_1 = \Sigma\Sigma\Sigma\,(\bar{y}_{i..} - \bar{y}_{...})^2$, $SS_2 = \Sigma\Sigma\Sigma\,(\bar{y}_{.j.} - \bar{y}_{...})^2$ and $SS_3 = \Sigma\Sigma\Sigma\,(y_{ijk} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2$. For balanced data, those sums of squares are statistically independent, with $SS_1/(Jm\sigma_1^2 + \sigma_3^2) \sim \chi^2(I-1)$, $SS_2/(Im\sigma_2^2 + \sigma_3^2) \sim \chi^2(J-1)$ and $SS_3/\sigma_3^2 \sim \chi^2(n - I - J + 1)$, in which case

$$(4.2) \qquad S(y\colon \gamma) = \left\{\gamma\colon c_1 \le \frac{MS_1/MS_3}{Jm\gamma_1 + 1} \le d_1,\, c_2 \le \frac{MS_2/MS_3}{Im\gamma_2 + 1} \le d_2\right\},$$

where $MS_1 = SS_1/(I-1)$, $MS_2 = SS_2/(J-1)$ and $MS_3 = SS_3/(n - I - J + 1)$.

For data sets II and III, the vertical boundaries of $S(y\colon \gamma)$ are not straight lines. The curvature in the vertical boundaries of $S(y\colon \gamma)$ for data set II is so slight as to be imperceptible in Figure 1. For mixed models of the form (4.1), $S(y\colon \gamma)$ differs noticeably from a rectangle only in cases of extreme imbalance.

5. Some remarks.

1. Generally there is more than one way to express a particular mixed model as a special case of the general mixed linear model (1.1). For example, in the additive two-way crossed model (4.1), the elements of β_1 in (1.1) could be the effects of the first factor or, alternatively, the effects of the second factor. Except for those situations where the spaces U_i associated with the identified sets of random effects are unchanged by the different possible orderings of these effects (e.g., balanced additive models), the sets $S(y\colon \sigma)$ and $S(y\colon \gamma)$ will vary with the order in which the effects are assigned to $\beta_1, \dots, \beta_k$.

2. Harville (1988) defined a function, say $g(y)$, of y to be X_i^*-invariant [under model (1.1)] if $g(y + X_i^* b) = g(y)$ for every vector b (of appropriate dimension) and every value of y. An X_i^*-invariant statistical procedure (a statistical procedure that depends on y only through the values of X_i^*-invariant statistics) is such that its statistical properties do not depend on β_0

or on $\sigma_1^2, \dots, \sigma_i^2$ or $\gamma_1, \dots, \gamma_i$. It is clear that, for any fixed value of σ, $\tilde{F}_i(y, \sigma)$ is X_{i-1}^*-invariant. Similarly, for any fixed value of γ, $F_i(y, \sigma)$ is X_{i-1}^*-invariant.

The set $S(y\colon \sigma)$ is one in a general class of exact $100(1 - \alpha)\%$ confidence sets for σ, which we now describe. Letting σ_0 represent the true value of σ, define $\tilde{F}_i^*(y, \sigma) = y^t B_i(\sigma) y$, $i = 1, \dots, k+1$, to be quadratic forms (in y), whose matrices $B_1(\sigma), \dots, B_{k+1}(\sigma)$ may vary with σ and that have the following three properties:

(i) $\tilde{F}_1^*(y, \sigma_0), \dots, \tilde{F}_{k+1}^*(y, \sigma_0)$ are jointly independent;
(ii) $\tilde{F}_i^*(y, \sigma_0) \sim \chi^2(r_i^*)$ for some positive integer r_i^*, $i = 1, \dots, k+1$;
(iii) for any fixed value of σ, $\tilde{F}_i^*(y, \sigma)$ is X_{i-1}^*-invariant, $i = 1, \dots, k+1$.

Define $S^*(y\colon \sigma)$ to be the set $S^*(y\colon \sigma) = \{\sigma\colon a_i \le \tilde{F}_i^*(y, \sigma) \le b_i \text{ for } 1 \le i \le k+1\}$. Then clearly $S^*(y\colon \sigma)$ is an exact $100(1 - \alpha)\%$ confidence set for σ.

The proposed set $S(y\colon \sigma)$ can be derived by limiting attention to sets of the general form $S^*(y\colon \sigma)$ and by successively choosing $B_{k+1}(\sigma), B_k(\sigma), \dots, B_1(\sigma)$ to maximize $r_{k+1}^*, r_k^*, \dots, r_1^*$, respectively [i.e., by choosing $B_{k+1}(\sigma)$ to maximize r_{k+1}^*, choosing $B_k(\sigma)$ to maximize r_k^* given that $B_{k+1}^*(\sigma)$ has been chosen to maximize r_{k+1}^* and in general choosing $B_i(\sigma)$ to maximize r_i^* given that $B_{k+1}(\sigma), B_k(\sigma), \dots, B_{i+1}(\sigma)$ have been successively chosen to maximize $r_{k+1}^*, r_k^*, \dots, r_{i+1}^*$, respectively]. To see this, observe that, for any fixed value of σ, $\tilde{F}_i^*(y, \sigma)$ is expressible as a quadratic form in $H_i(\sigma)y, \dots, H_{k+1}(\sigma)y$ [Harville (1988), Section 4].

3. Hartley and Rao (1967), Section 9, constructed a pivotal statistic $G(y, \gamma)$ for exact inference on γ. There is a simple relationship between $G(y, \gamma)$ and our pivotals $G_i(y, \gamma)$; specifically,

$$G(y, \gamma) = \sum_1^k r_i G_i(y, \gamma) \Big/ \sum_1^k r_i. \tag{5.1}$$

For example, if the model is (4.1) and if $m_{ij} \equiv m$, then

$$G(y, \gamma) = \frac{\omega_1 MS_1/MS_3}{Jm\gamma_1 + 1} + \frac{\omega_2 MS_2/MS_3}{Im\gamma_2 + 1},$$

where $\omega_1 = (I - 1)/(I + J - 2)$, $\omega_2 = (J - 1)/(I + J - 2)$ and Hartley and Rao's confidence set for γ is

$$\left\{\gamma\colon q_1 \le \frac{\omega_1 MS_1/MS_3}{Jm\gamma_1 + 1} + \frac{\omega_2 MS_2/MS_3}{Im\gamma_2 + 1} \le q_2\right\}, \tag{5.2}$$

where the $q_i > 0$ are selected percentiles from the appropriate F distribution. The confidence set (5.2) is quite different from $S(y\colon \gamma)$, (4.2), which (in this special case) is rectangular. In particular, $|\gamma|$ may be unbounded over (5.2), and (5.2) may have infinite area.

4. Calculating the $100(1 - \alpha)\%$ confidence set $S(y\colon \gamma)$ requires the constants c_i and d_i satisfying $P\{c_i \le F_i \le d_i \text{ for } 1 \le i \le k\} = 1 - \alpha$; see (2.3). In

the special case $r_1 = \cdots = r_k$, these constants can be obtained from published tables [Johnson and Kotz (1972), Section 40.8]. More generally, they can be determined iteratively by employing numerical integration or simulation. Moreover, an approximate $100(1-\alpha)\%$ confidence set for γ can be obtained by replacing c_i and d_i in (2.3) with e_i and f_i, where $P\{e_i \le F_i \le f_i\} = (1-\alpha_i)$ with $\prod_1^k(1-\alpha_i) = 1-\alpha$. In the special case where the e_i's are all assigned the value 0, Kimball's inequality [Miller (1981), page 101] implies that this region is conservative, that is, its probability of coverage equals or exceeds $(1-\alpha)$; in the case of balanced classificatory models, this region simplifies to that proposed by Broemeling (1969).

5. The proposed confidence sets for α and γ can be modified to obtain exact confidence sets for $\sigma_i^2, \sigma_{i+1}^2, \ldots, \sigma_{k+1}^2$ and for $\gamma_i, \gamma_{i+1}, \ldots, \gamma_k$, as is evident from the discussion in Section 3.

6. By following the approach described, for example, by Spjotvoll (1972) and Khuri (1981), the proposed confidence region for the variance components $\sigma_1^2, \ldots, \sigma_{k+1}^2$ or variance ratios $\gamma_1, \ldots, \gamma_k$ can be transformed into a generally conservative confidence region for a function of the variance components or variance ratios or, more generally, for a family of such functions.

7. Note that, by using general relationships between confidence regions and tests of hypothesis or significance, the proposed procedure for forming confidence regions can be used to test, against appropriate alternatives, a null hypothesis of the general form

$$H_0\colon \sigma_1^2 = \bar{\sigma}_1^2, \ldots, \sigma_{k+1}^2 = \bar{\sigma}_{k+1}^2 \quad \text{or} \quad H_0\colon \gamma_1 = \bar{\gamma}_1, \ldots, \gamma_k = \bar{\gamma}_k,$$

where the $\bar{\sigma}_i^2$ and $\bar{\gamma}_i$ represent specified constants.

APPENDIX

For any matrix G, G^- denotes a matrix satisfying $GG^-G = G$. Let A be a symmetric positive definite matrix, in which case $(x, y) = x^tAy$ is an inner product for the vector space $\mathbb{R}^n$.

LEMMA A1. *Suppose X is an $n \times k$ matrix whose column space is V. Then the projection matrix for V using x^tAy is*

$$P[V] = X(X^tAX)^- X^tA.$$

PROOF. Omitted.

LEMMA A2. *Suppose X_0 and X_1 are $n \times m_0$ and $n \times m_1$ matrices, respectively. Let $C(X_0)^\perp$ and $C(X_0X_1)^\perp$ denote the complements of $C(X_0)$ and $C(X_0X_1)$, respectively, within $\mathbb{R}^n$ with respect to the Euclidean inner product. The row space of the matrix $X_1^tA^{-1} - X_1^tA^{-1}X_0(X_0^tA^{-1}X_0)^-X_0^tA^{-1}$ is the complement of $C(X_0, X_1)^\perp$ within $C(X_0)^\perp$ with respect to the inner product x^tAy.*

PROOF. Omitted.

Acknowledgments. The suggestions and observations of the reviewers of this article contributed substantially and positively to its development. We thank them for their efforts.

REFERENCES

BROEMELING, L. D. (1969). Confidence regions for variance ratios of random models. *J. Amer. Statist. Assoc.* **64** 660–664.

BROEMELING, L. D. and BEE, D. E. (1976). Simultaneous confidence intervals for parameters of a balanced incomplete block design. *J. Amer. Statist. Assoc.* **71** 425–428.

BROWN, K. G. (1984). On analysis of variance in the mixed model. *Ann. Statist.* **12** 1488–1499.

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

HARTLEY, H. O. and RAO, J. N. K. (1967). Maximum-likelihood estimation for the mixed analysis of variance model. *Biometrika* **54** 93–108.

HARVILLE, D. A. (1988). Invariant inference for variance components. In *Probability and Statistics: Essays in Honor of Franklin A. Graybill* (J. N. Srivastava, ed.) 117–133. North-Holland, Amsterdam.

HARVILLE, D. A. and FENECH, A. P. (1985). Confidence intervals for a variance ratio, or for heritability in an unbalanced mixed linear model. *Biometrics* **41** 137–152.

HENDERSON, H. V. and SEARLE, S. R. (1981). On deriving the inverse of a sum of matrices. *SIAM Rev.* **23** 53–60.

JOHNSON, N. L. and KOTZ, S. (1972). *Distributions in Statistics: Continuous Multivariate Distributions.* Wiley, New York.

KHURI, A. I. (1981). Simultaneous confidence intervals for functions of variance components in random models. *J. Amer. Statist. Assoc.* **76** 878–885.

MILLER, R. G. (1981). *Simultaneous Statistical Inference*, 2nd ed. Springer, New York.

PINCUS, R. (1977). On tests in variance components models. *Math. Operationsforsch. Statist. Ser. Statist.* **8** 251–255.

SEELY, J. F. and EL-BASSIOUNI, Y. (1983). Applying Wald's variance component test. *Ann. Statist.* **11** 197–201.

SPJOTVOLL, E. (1968). Confidence intervals and tests for variance ratios in unbalanced variance components models. *Rev. Internat. Statist. Inst.* **36** 37–42.

THOMPSON, W. A., JR. (1955). On the ratio of variances in the mixed incomplete block model. *Ann. Math. Statist.* **26** 721–733.

TJUR, T. (1984). Analysis of variance models in orthogonal designs. *Internat. Statist. Rev.* **52** 33–81.

WALD, A. (1947). A note on regression analysis. *Ann. Math. Statist.* **18** 586–589.

DIVISION OF STATISTICS
UNIVERSITY OF CALIFORNIA, DAVIS
DAVIS, CALIFORNIA 95616

DEPARTMENT OF STATISTICS
IOWA STATE UNIVERSITY
AMES, IOWA 50011