4

AD-A226 854

IMPROVED ESTIMATION OF VARIANCE COMPONENTS IN MIXED MODELS

BY

ALAN E. GELFAND and DIPAK K. DEY

TECHNICAL REPORT NO. 434
AUGUST 30, 1990

Prepared Under Contract
N00014-89-J-1627 (NR-042-267)
For the Office of Naval Research

Herbert Solomon, Project Director

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

ABSTRACT

Taking Albert's (1976) formulation of a mixed model ~~ANOVA,~~ we consider improved estimation of the variance components for balanced designs under squared error loss. Two approaches are presented. One extends the ideas of Stein (1964). The other is developed from the fact that variance components can be expressed as linear combinations of chi-square scale parameters. Encouraging simulation results are presented.

1. INTRODUCTION

Albert (1976) exhibits necessary and sufficient conditions for a sum of squares decomposition under a mixed model to be an ANOVA, i.e., for the terms of the decomposition to be independent and to be distributed as multiples of chi-square. We consider improved estimation of the variance components under squared error loss in such a set-up. We make no attempt to discuss the enormous literature on this problem. See Harville (1977) for such a review. Rather, we specialize to the "balanced" case considering designs consisting of crossed and nested classifications and combinations thereof. Rules of thumb for formalizing the associated ANOVA table are thus well known (see, e.g., Searle (1971, Chap. 9)). Customary estimators of the variance components are the unbiased ones obtained as described in Searle, pp. 405-6. Under normality these estimators are UMVU (Graybill (1954), Graybill and Wortham (1956)) and, in fact, restricted maximum likelihood (REML) (Thompson (1962)). However, positive part corrections are usually taken yielding improved mean square error but sacrificing these "optimalities." Bayesian approaches to variance component estimation in this setting are discussed in, e.g., Hill (1965) and Box and Tiao (1973).

Since the positive part estimators are not smooth and, thus, not admissible under squared error loss (SEL), it is natural to seek dominating estimators. The earliest work of this type is due to Klotz, Milton and Zacks (1969) for the one-way layout. They show, for example, that the MLE of the "between" variance component (see Herbach (1959)), which is itself a positive part estimator, dominates the UMVU and, using ideas of Stein (1964), that it in turn can be dominated.

The objective of this paper is to describe two general approaches for creating improved estimators of the variance components under SEL. In Section 2 we develop a method by extending the aformentioned Stein idea. We have discussed a special case of this approach in linear regression models in Gelfand and Dey (1987b). Since the variance components are linear combinations of chi-square scale parameters, we can draw upon some literature for improved estimation of linear combinations of scale parameters. This second approach is offered in Section 3. Work of Dey and Gelfand (1987) for arbitrary scale parameter distributions and of Klonecki and Zontek (1985, 1987) for the Gamma family of distributions is pertinent here. Finally, in Section 4 we present some simulation results.

In the remainder of this section we develop notation for and features of the model we will be working with. Consider the general balanced mixed model of the form

$$Y = \mu 1_{n\times 1} + \sum_{r=1}^{p} H_r \tau_r + X\beta + \varepsilon \qquad (1.1)$$

where Y is an n x 1 vector of observations, μ is an overall mean effect or intercept, H_r are known $n \times m_r$ incidence matrices where $H_r 1_{m_r \times 1} = 1_{n\times 1}$ and $H_r^T H_r = \nu_r I$ (i.e., ν_r is the number of nonzero entries in a typical column of H_r), τ_r are independent distributed as $N(0, \sigma_r^2 I_{m_r})$, X is a known n x s design matrix involving possibly fixed effects and covariates, β is the associated s x 1 vector of coefficients and ε is an n x 1 vector of errors distributed $N(0,\sigma_e^2 I_n)$ independent of the τ_r. Thus, $Y \sim N(m,W)$ where

$$m = \mu 1_{nx1} + X\beta$$

and W is the patterned covariance matrix

$$W = \sigma_e^2 I_n + \sum_{r=1}^{p} \sigma_r^2 H_r H_r^T$$

Let $(\sigma^2)^T = (\sigma_e^2, \sigma_1^2, \ldots, \sigma_p^2)^T$. Our primary interest is in estimating the σ_r^2 individually (as it has been done historically) although we shall say something in Section 3 about simultaneous estimation.

As in Albert (1976) we consider a complete set of orthogonal projections, $P_1, P_2, \ldots, P_p, P_e, P_\mu, P_\beta$, $\sum_1^p P_i + P_e + P_\mu + P_\beta = I_n$. In particular, P_e is associated with the error, i.e., $Y^T P_e Y$ is the full model error sum of squares. P_μ is associated with the intercept $(P_\mu = \mu^{-1} 1_{nxn})$, i.e., $Y^T P_\mu Y = n\bar{Y}.^2$ where $\bar{Y}.$ is the average of the Y's. SS_H is the model sum of squares for the reduced ANOVA model, i.e., $SS_H = Y^T H(H^T H)^- H^T Y = \sum_{i=1}^{p} Y^T P_i Y + n\bar{Y}.^2$ where $H = (H_1 H_2 \ldots H_p)$. Note that we have a sum of squares for each random effect. Finally, let $SS_{\beta|H}$ be the sum of squares for the fixed effects and covariates adjusted for the ANOVA, i.e.,

$SS_{\beta|H} = Y^TP_\beta Y$. Typically, P_β is itself expressed as a sum of orthogonal pieces.

According to Albert (1976) (see also Brown (1984) and Harville (1984) in this regard), we have an ANOVA if and only if for r = 1, 2, ..., p, $H_rH_r^TP_k = \lambda_{kr}P_r$, k = 1, ..., r, $H_rH_r^TP_\mu = \lambda_{\mu r}P_\mu$, $H_rH_r^TP_e = \lambda_{er}P_e$ and $H_rH_r^TP_\beta = \lambda_{\beta r}P_\beta$ where $\lambda_{\cdot r} = 0$ or ν_r according to whether or not $H_r^TP_\cdot = 0$. Then the $Q_\cdot = Y^TP_\cdot Y$ are independent and distributed as $\gamma_\cdot\chi^2_{f_\cdot,\delta_\cdot}$ where

$\gamma_\cdot = \sigma_e^2 + \sum_{r=1}^{p}\lambda_{\cdot r}\sigma_r^2$, $f_\cdot = \text{rank}(P_\cdot)$, $\delta_k = 0$, $\delta_e = 0$, $\delta_\mu = m^TP_\mu m/2\gamma_\mu$, $\delta_\beta = m^TP_\beta m/2\gamma_\beta$.

We note that since for each r, $1_{nx1} \in M(H_r)$, $\lambda_{\mu r} = \nu_r$ and, thus, $\gamma_\mu = \sigma_e^2 + \Sigma\nu_r\sigma_r^2$. Since $H^TP_e = 0$ and $H^TP_\beta = 0$ $\lambda_{er} = 0$ and $\lambda_{\beta r} = 0$, i.e., $\gamma_e = \sigma_e^2$ and $\gamma_\beta = \sigma_e^2$. The λ_{kr} can not be determined explicitly without specifying the design.

However, for two random effects, with respective sums of squares Y^TP_kY and $Y^TP_{k'}Y$, if the latter is any nested or crossed effect involving all the factors in the former, then $\gamma_k \geq \gamma_{k'}$. This is , in fact, Rule 12 of Searle (1971 , p. 393). Obviously, $\gamma_e \leq \gamma_k \leq \gamma_\mu$ and typically there is a partial ordering amongst the γ_k.

Finally, again as in Searle (1971, p. 405), if we define $\gamma^T = (\gamma_0, \gamma_1, \ldots, \gamma_p)$ with $\gamma_0 = \sigma_e^2$ we have $\gamma = A\sigma^2$ where

$A = \begin{pmatrix} 1 & 0^T \\ 1 & \Omega \end{pmatrix}$ with $\{\Omega\}_{kr} = \lambda_{kr}$ whence

$$\sigma^2 = A^{-1}\gamma. \tag{1.2}$$

Expression (1.2) reveals a key point. The variance components are expressible as linear combinations of chi-square scale parameters. In fact, this expression is usually employed to create the familiar unbiased estimators of the σ_r^2 using $f_k^{-1}Q_k$, the unbiased estimator of γ_k.

2. IMPROVED ESTIMATORS USING STEIN'S METHOD

Consider estimating $a\gamma_k + b\gamma_{k'}$. For appropriate choices of a, b, k, k' (in fact $ab < 0$), this parameter will be a variance component. To proceed we utilize the following elementary lemma whose proof is immediate.

Lemma 2.1. Let S_1 be an estimator of θ_1 and let T_1 dominate S_1 under SEL. Let S_2 be an estimator of θ_2 where S_2 is independent of S_1 and T_1. Then in estimating $a\theta_1 + b\theta_2$, $aT_1 + bs_2$ dominates $aS_1 + bS_2$ under SEL if

$$abE_{\theta_1}(S_1 - T_1)\, E_{\theta_2}(S_2 - \theta_2) \geq 0 \tag{2.1}$$

In our applications we will meet (2.1) by having $ab < 0$, $T_1 \leq S_1$, $E_{\theta_2}(S_2) \leq \theta_2$.

We also require the following result which is a minor generalization of a theorem stated and proved in Gelfand and Dey (1987b).

Theorem 2.1. Let $S_0 \sim \phi_0\chi^2_{n_0} + S_i \sim (\gamma_0 + \phi_i)\chi^2_{n_i,\lambda_i}$, $i = 1, \ldots, t$ all independent where $\phi_i \geq 0$, $\lambda_i \geq 0$. Define $R_j = c_j^{-1}\sum_{i=1}^{j} S_i$ where $c_j = \sum_{i=1}^{j} n_i + 2$ and let $\delta_j = \min(R_0, R_1, \ldots, R_j)$.

Then in estimating γ_0 under SEL $\delta_0 << \delta_1 << \delta_2 \ldots << \delta_t$ where $\delta_i << \delta_j$ means δ_j dominates δ_i.

Lastly we need a lemma which appears, for example, in Klotz, Milton and Zacks (1969, p. 1394).

Lemma 2.2. If $T \leq S$ and, in estimating $\theta > 0$, $S << T$ under SEL then $S^+ << T^+$, under SEL where + denotes positive part.

These results will be synthesized in the following way. Assume $a\gamma_k + b\gamma_{k'} > 0$ and w.l.o.g. that $a > 0$, $b < 0$. This will be the case if $a\gamma_k + b\gamma_{k'}$ defines a variance component. Find the set of all $\gamma_r \geq \gamma_k$ (excluding $\gamma_{k'}$, regardless). This set is nonempty since at the very least $\gamma_\mu \geq \gamma_k$. Q_k and the associated set of Q_r form the S_0 and S_i, respectively, for Theorem 2.1 and enable the creation of a decreasing sequence of estimators which dominate $(f_k + 2)^{-1}Q_k$, the best invariant estimator of γ_k. The resultant δ_j suitably defined play the role of S_1 and T_1 in Lemma 2.1 and will be independent of $S_2 = (f_{k'} + 2)^{-1} Q_{k'}$, $\gamma \geq 0$ whence Lemma 2.1 holds. Finally, using Lemma 2.2, $[aS_1 + bS_2]^+ <<$ $[aT_1 + bS_2]^+$.

Remark 2.1. Theorem 2.1 allows for a variety of improved estimators for σ_e^2. Let $S_0 = Q_e$ with S_i being the Q_r, $r = 1, \ldots, p$ as well as Q_μ and Q_β. Then $t = p + 2$ and we may readily create δ_t. In fact, corresponding to any specified permutation, π, of the S_i there will be a resultant δ_t^π, i.e., there will be t! such estimators. How might we combine them to produce a permutation invariant estimator? It can be argued that the minimum of these will be "too small" and that the average is a better practical choice. See Gelfand and Dey (1987b) for details.

We illustrate using the one-way ANOVA, $Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$,

$i = 1, \ldots, I$, $j = 1, \ldots, J$, $\alpha_i \sim N(0, \sigma_\alpha^2)$, $\varepsilon_{ij} \sim N(0, \sigma_e^2)$, all

independent. In this case $p = 1$ with $Q_1 \sim (\sigma_e^2 + J\sigma_\alpha^2)\chi^2_{I-1}$,

$Q_u \sim (\sigma_e^2 + J\sigma_\alpha^2)\chi^2_{1,\ \frac{IJ\ \ ^2}{2(\sigma_e^2 + J\sigma_\alpha^2)}}$ and $Q_e \sim \sigma_e^2\chi^2_{I(J-1)}$.

In estimating σ_e^2 we may dominate the best invariant estimator $R_0 = (I(J-1) + 2)^{-1}Q_3$ using $\delta_{e,1} = \min(R_0, R_1)$ which in turn is dominated by

$$\delta_{e,2} = \min(R_0, R_1, R_2) \tag{2.2}$$

or using $\delta'_{e,1} = \min(R_0, R_1')$ which in turn is dominated by

$$\delta'_{e,2} = \min(R_0, R_1', R_2) \tag{2.3}$$

Here $R_1 = (IJ + 1)^{-1}(Q_e + Q_1)$, $R_1' = (I(J - 1) + 3)^{-1}(Q_e + Q_u)$ and $R_2 = (IJ + 2)(Q_e + Q_1 + Q_u)$. Estimators $\delta_{e,1}$ and $\delta_{e,2}$ appear in Klotz, Milton and Zacks (1969). In practice, if we suspect σ_α^2 small we would use $\delta_{e,2}$; if we suspect μ small we would use $\delta'_{e,2}$ and if we have no prior suspicions we would recommend

$$(\delta_{e,2} + \delta'_{e,2})/2 \tag{2.4}$$

Turning to σ_α^2 we may write $\sigma_\alpha^2 = J^{-1}(\sigma_1^2 - \sigma_e^2)$ whence the usual unbiased estimator is given by $J^{-1}[(I - 1)^{-1}Q_1 = (I(J - 1))^{-1}Q_e]$. By Lemma 2.1 this is immediately dominated by $J^{-1}[(I + 1)^{-1}Q_1 - (I(J - 1))^{-1}Q_e]$ which in turn is dominated by $\delta_{\alpha,1} =$

$J^{-1}[\min\{(I + 1)^{-1}Q_1, (I + 2)^{-1}(Q_1 + Q_u)\} - (I(J - 1))^{-1}Q_e]$.

Using Lemma 2.2 we arrive at the positive part version $\delta^+_{\alpha,1}$.

Alternatively, again by Lemma 2.1, the usual unbiased estimator is dominated by $J^{-1}[(I-1)^{-1}Q_1 - (I(J - 1) + 2)^{-1}Q_e]$ which is dominated by $J^{-1}[(I + 1)^{-1}Q_1 - (I(J - 1) + 2)^{-1}Q_e]$ which is dominated by $\hat{\sigma}_{\alpha,2} = J^{-1}[\min\{(I + 1)^{-1}Q_1, (I + 2)^{-1}(Q_1 + Q_e)\} - (I(J - 1) + 2)^{-1}Q_e]$. Again by Lemma 2.2, we arrive at $\hat{\sigma}^+_{\alpha,2}$. The estimator $\hat{\sigma}^+_{\alpha,1}$ appears in Klotz, Milton and Zacks. Note that while $\hat{\sigma}_{\alpha,1} << \hat{\sigma}_{\alpha,2}$ since $\hat{\sigma}_{\alpha,2} > \hat{\sigma}_{\alpha,1}$ we cannot conclude regarding $\hat{\sigma}^+_{\alpha,1}$ and $\hat{\sigma}^+_{\alpha,2}$. In fact, the simulation results in Section 4 show that neither dominates the other.

In concluding this section we remark that utilizing the ideas in Gelfand and Dey (1987a) and in Gelfand (1987), along with the aformentioned results, we can improve in the estimation of the ratio $\gamma_k/\gamma_{k'}$. This allows for improved estimation of, e.g., σ^2_r/σ^2_e. See Loh (1986) in this regard. We omit the details. Unfortunately, we cannot extend this to, e.g., the intraclass correlation coefficient since it is a non-linear function of such ratios.

3. IMPROVED ESTIMATES USING A GEOMETRIC MEANS APPROACH

In this section we develop a method for obtaining improved estimates which arises from expression (1.2), the fact that the variance components are expressible as linear combinations of chi-square scale parameters. Consider a single σ^2_r which we write as $\sigma^2_r = \sum_{k=0}^{p} c_k\gamma_k$ and let $\sum_{k=0}^{p} \lambda_k Q_k$ be a candidate estimator. Here we denote Q_e by Q_0. When can $\sum\lambda_k Q_k$ be dominated and what is the form of the dominating estimator? Dey and Gelfand (1987) discuss this problem when the γ_k are scale parameters from arbitrary distributions. Klonecki and Zontek (1985, 1987), assuming the γ_k are scale parameters from Gamma distributions, obtain conditions which enable assessment of linear admissibility for $\Sigma \xi_k Q_k$, i.e., admissibility within the class of linear estimators of σ_r^2. They also offer a slightly broader class of dominating estimators than in Dey and Gelfand (1987).

More precisely, the following result appears in Dey and Gelfand (1987).

<u>Theorem 3.1</u>. Let $Y_i \sim \theta_i f_i$, $i = 1, 2, \ldots, t$, $t \geq 2$, independent and such that $EY_i^2 < \infty$. Consider the estimator

$$\sum_{i=1}^{t} c_i Y_i + b\left(\prod_{i=1}^{t} Y_i\right)^{1/t} \tag{3.1}$$

Let $d_i = c_i - \gamma a_i$ where $a_i = E(Y_i^{1+t^{-1}} \mid \theta_i = 1)/E(Y_i^{t^{-1}} \mid \theta_i = 1)$ and $d_{(1)} = \min_i d_i$, $d_{(t)} = \max_i d_i$. Then (3.1) dominates $\sum c_i Y_i$ under SEL in estimating $\sum c_i \theta_i$ if either

(i) $d_{(1)} > 0$ and $0 < b < 2td_{(1)}$

(ii) $d_{(t)} < 0$ and $2td^{(t)} < b < 0$

where $\gamma = \prod_{i=1}^{t} E(Y_i^{t^{-1}} \mid \theta_i = 1)/E(Y_i^{2t^{-1}} \mid \theta_i = 1)$

If we denote by $G(\alpha_i, \beta_i)$ the gamma density $f_i(y) =$

$\dfrac{y^{\alpha_i - 1} e^{-y/\theta_i}}{\theta_i^{\alpha_i} \Gamma(\alpha_i)}$, then $a_i = \alpha_i + t^{-1}$ and $\rho = \Sigma_i (\alpha_i + t^{-1})/(\alpha_i + 2t^{-1})$.

Now let D be a diagonal matrix whose diagonal entries are the α_i and let G be of the form $\Sigma_d^{-1}\Sigma$ where Σ is a nonnegative definite matrix such that $(\Sigma)_{ij} \geq 0$, $(\Sigma)_{ii} > 0$ and Σ_d is a diagonal matrix whose diagonal entries are $(\Sigma)_{ii}$. Then Klonecki and Zontek (1985) show:

Theorem 3.2. If $Y_i \sim G(\alpha_i, \theta_i)$ $i = 1 \ldots t$ independent, then $\Sigma \ell_i y_i$ is linearly admissible for $\Sigma c_i \theta_i$ if and only if there exists a matrix G of the above form such that $(I + GD)\ell = Gc$ where $\ell^T = (\ell_1, \ldots, \ell_t)$, $c^T = (c_1, \ldots, c_t)$.

Theorem 3.1 is often too restrictive. If instead we allow a more general product $\Pi Y_j^{q_j}$ we can choose q_j to achieve suitably defined "d_i" all having the same sign. In fact, for a specified set of $q_j \geq 0$ such that $\Sigma q_j = 1$ Klonecki and Zontek (1987), again for gamma distributions, provide necessary and sufficient conditions for the existence of an estimator of this form which improves upon $\Sigma \ell_i Y_i$. We state a version of their Lemma 1 which is in a form parallel to Theorem 3.1.

Theorem 3.3. If $Y_i \sim G(\alpha_i, \theta_i)$ $i = 1 \ldots t$ independent there exists $b \neq 0$ such that the estimator

$$\Sigma \ell_i Y_i + b \Pi Y_j^{q_j}, \tag{3.2}$$

where $q_j \geq 0$, $\Sigma q_j = 1$, dominates $\Sigma \ell_i y_i$ under SEL, in estimating $\Sigma c_i \theta_i$ if and only if either (i) or (ii) below holds. Define $d_i^* =$

$c_i - (\alpha_i + q_i)\xi_i$, $d^*_{(1)} = \min_i d_i^*$ and $d^*_{(t)} = \max_i d_i^*$.

(i) $d^*_{(1)} \geq 0$, $q_i = 0$ if $d^*_i = 0$ and $0 < b < b^*$

(ii) $d^*_{(t)} \leq 0$, $q_i = 0$ if $d^*_i = 0$ and $-b^* < b < 0$

where $b^* = 2 \sum_{\{j:d_j \neq 0\}} \left(\frac{\alpha_j + q_j}{\alpha_j + 2q_j}\right) \left|\frac{d_j}{q_j}\right|^{q_j}$.

Remark 3.1. Theorem 3.3 holds more generally than for the gamma family. Its proof only requires specification of the increasing functions $w_i(q) \equiv E(Y_i^{1+q} \mid \theta_i = 1)/E(Y_i^q \mid \theta_i = 1)$. (In the gamma case $w_i(q) = \alpha_i + q$). Given w_i we can characterize the sets of q_j's which make the corresponding d^*_i's all have the same sign, thus enabling domination by (3.2).

Remark 3.2. The bounding of the risk difference in Theorem 3.1 is not as sharp as is possible under the Gamma assumption in Theorem 3.3; hence, the resulting bounds on b in Theorem 3.3 when all q^j are equal are more liberal than those in Theorem 3.1.

Returning to the estimation of a variance component $\sigma_r^2 = \sum_{k=0}^{p} c_k \gamma_k$ consider the estimator $\sum_{k=0}^{p} \delta_k Q_k$ where $\delta_k = c_k \cdot (f_k + 2\varepsilon_k)^{-1}$, $0 \leq \varepsilon_k \leq 1$. The terms $\delta_k Q_k$ range from the unbiased to the best invariant estimator of $c_k \gamma_k$ as ε_k ranges from 0 to 1. Since $\gamma_k \chi^2_{f_k}$ is $G(\frac{1}{2} f_k, 2\gamma_k)$,

$$d^*_k = c_k(\varepsilon_k - q_k)/(f_k + 2\varepsilon_k) \tag{3.3}$$

If $c_k = 0$ we must set $q_k = 0$. Thus, if γ_k does not appear in σ_r^2, using Theorem 3.3, Q_k does not help in estimating σ_r^2. This clearly differs from the approach in Section 2 where, for example, in estimating σ_e^2 all the Q_k can be used to improve upon the best invariant estimator. Note that with ℓ_k as defined above, from (3.3), the sign of d_k^* depends only upon $\mathrm{sgn}(c_k(\varepsilon_k - q_k))$; for specified ε_k and q_k the magnitude of c_k does not play a role with respect to whether an estimator of the form (3.2) can dominate.

From (3.3) if all $\varepsilon_k = 0$ or all $\varepsilon_k = 1$ this approach will provide a dominating estimator if and only if at least two c_k differ from 0 and all nonzero c_k have the same sign. For a variance component some pair of c_k will have opposite signs. Therefore, a dominating estimator will not be obtained if for any such pair both ε's are 0 or both ε's are 1. If we can choose q_k to make $d_{(1)}^* \leq 0$ then the dominating estimator in (3.2) will be a "shrinker." Hence, using Theorem 2.3, the positive part of (3.2) will dominate $[\sum_{k=0}^{p} \ell_k Q_k]^+$.

In this spirit it is natural to ask whether the approach of this section can be combined with that of the previous section. Can we improve upon the estimators developed through Lemma 2.1 and Theorem 2.1 using a more general version of Theorem 3.3 as suggested in Remark 3.1? The answer appears to be no since in the notation of Theorem 2.1 γ_0 is not a scale parameter for the distribution of δ_j. The reader might suggest that γ_0 could be viewed as a scale parameter for the distribution of δ_j under suitable conditioning. Following the argument leading to Theorem 3.3, while b must be chosen unconditionally, it would have to provide improvement at each conditional level. We can readily show that even in the simplest case, t = 2, no b unequal to 0 can achieve this.

As an example, we turn again to the one-way ANOVA using the notation in Section 2. Recalling $\sigma^2 = J^{-1}(\gamma_1 - \gamma_e)$ we consider dominating the estimator

$$J^{-1}[(I - 1 + 2\varepsilon_1)^{-1}Q_1 - (I(J - 1) + 2\varepsilon_e)^{-1}Q_e] \tag{3.4}$$

Thus, $d_1^* \gtrless 0$ as $q_1 \lessgtr t_1$, $d_e^* \gtrless 0$ as $q_e \gtrless \varepsilon_e$. As noted above,

this approach unfortunately does not provide a dominating estimator in the two important cases where $\varepsilon_1 = \varepsilon_e = 0$ and where $\varepsilon_1 = \varepsilon_e = 1$. Instead, we take

(i) $\varepsilon_1 = 1$, $\varepsilon_e = 0$ for which any $q_1, q_2, > 0$, $q_1 + q_2 = 1$ work

with

$$0 < b < \frac{2}{J}\,\frac{(I - 1) + 2q_1}{(I - 1) + 4q_1}\;\frac{I(J - 1) + 2q_2}{I(J - 1) + 4q_2}\;x$$

$$\left(\frac{1 - q_1}{q_1(I + 1)}\right)^{q_1}\left(\frac{1}{I(J - 1)}\right)^{q_2}$$

and (3.2) becomes

$$J^{-1}[(I + 1)^{-1}Q_1 - (I(J - 1)^{-1}Q_e] + b\,Q_1^{q_1}Q_e^{q_e} \tag{3.5}$$

(ii) $\varepsilon_1 = 0$, $\varepsilon_e = 1$ for which any $q_1, q_2 > 0$, $q_1 + q_2 = 1$ work with

$$0 > b > -\frac{2}{J}\,\frac{(I - 1) + 2q_1}{(I - 1) + 4q_1}\;\frac{I(J - 1) + 2q_2}{(I(J - 1) + 4q_2}\;x$$

$$\left(\frac{1}{I - 1}\right)^{q_1}\left(\frac{1 - q_2}{q_2[I(J - 1) + 2]}\right)^{q_2}$$

and (3.2) becomes

$$J^{-1}[(I-1)^{-1}Q_1 - (I(J-1)+2)^{-1}Q_e] + b\,Q_1^{q_1}Q_e^{q_e} \qquad (3.6)$$

We conclude this section with a remark.

Remark 3.3. Results applicable to the simultaneous estimation of variance components under unweighted SEL are given in Klonecki and Zontek (1987). In particular, extensions of Theorems 3.2 and 3.3 are given for the estimation of a vector $C^T\theta$ using L^TY. The special case C = I (not of interest here) has been extensively discussed. See, e.g., Berger (1980), Das Gupta (1986), Dey and Gelfand (1987), and Das Gupta, Dey and Gelfand (1987).

4. SIMULATION RESULTS

In the one-way ANOVA we studied improved estimation of both σ_e^2 and σ_α^2 by undertaking a substantial simulation study over various values of I, J, μ, σ_α^2 and σ_e^2. Each case received 10,000 replications. Even with so many replications, resimulation of particular cases suggests that the stated percent improvements (PI's) will only be accurate within 2%.

In estimating σ_e^2 some selected cases are presented in Table 1. In this table PI is relative to the best invariant estimator given above (2.2). Not surprisingly (2.2) outperforms (2.3) when μ is small, and vice versa when σ_α^2 is small. The estimator (2.4) seems like a good compromise. For fixed (μ, σ_α^2, σ_e^2), PI's increase in I, decrease in J. Although the PI's are small the fact that (2.2)-(2.4) are so simple to calculate encourages their use.

In the estimation of σ_α^2, the reference estimator is the positive part of the unbiased estimator.

$$\left[\frac{Q_1}{I-1} - \frac{Q_2}{I(J-1)}\right]^+ /J \qquad (4.1)$$

TABLE 1

PERCENT IMPROVEMENTS IN ESTIMATING σ_e^2

	$(\mu, \sigma_\alpha^2, \sigma_e^2)$	PI for (2.2)	(2.3)	(2.4)
I = 2, J = 5	(0, 1, 10)	2.60	2.39	2.65
	(1, .1, 10)	1.95	2.27	2.28
	(1, 1, 1)	1.39	0.78	1.24
I = 5, J = 5	(0, 1, 10)	4.09	2.98	3.72
	(1, .1, 10)	3.64	3.36	3.95
	(1, 1, 1)	0.72	0.15	0.50
I = 10, J = 5	(0, 1, 10)	4.97	3.86	4.53
	(1, .1, 10)	6.09	3.92	5.69
	(1, 1, 1)	0.15	0.03	0.10

As shown in Section 2, (4.1) is dominated by

$$\left[\frac{Q_1}{I+1} - \frac{Q_e}{I(J-1)}\right]^+ /J \tag{4.2}$$

and by

$$\left[\frac{Q_1}{I+1} - \frac{Q_e}{I(J-1)+2}\right]^+ /J \tag{4.3}$$

Neither of (4.2) and (4.3) dominates the other. However, from Section 2, $\delta_{\alpha,1}^+$ dominates (4.2), $\delta_{\alpha,2}^+$ dominates (4.3). For b sufficiently small (3.5) dominates (4.2) ignoring the positive parts. With positive parts applied to both estimators this is no longer true. Since there is no obvious optimal choice we took b at the middle of the allowable range.

In Table 2 we compare $\delta^+_{\alpha,1}$ and $\delta^+_{\alpha,2}$ with (4.1). We see enormous improvement for both, that the PI's are essentially indistinguishable and that neither of the δ's dominates the other. We would draw the same conclusions in the comparison of (4.2) and (4.3) with (4.1). Of course, if $\sigma^2_\alpha > \sigma^2_e$ then $\delta_{\alpha,1}$ will tend to be nonnegative whence the domination result in Section 2 argues for $\delta^+_{\alpha,2}$. Turning to a comparison of $\delta^+_{\alpha,2}$ with (4.3) we see that if μ is small the gain may be substantial. A comparison of $\delta^+_{\alpha,1}$ with (4.2) would yield essentially the same magnitudes of improvement. Again, since these estimators are so simple to calculate, their use is encouraged. Finally, the comparison of the positive part of (3.5) with (4.2) is discouraging when σ^2_α is smaller than σ^2_e. Modest improvement will usually occur when $\sigma^2_\alpha > \sigma^2_e$. This is reasonable since then the positive part modification is rarely applied and the dominance result comes into play.

We conclude by recommending (2.4) for σ^2_e and $\delta^+_{\alpha,2}$ for σ^2_α.

ACKNOWLEDGMENT

The authors acknowledge Brad Carlin for performing the computations.

TABLE 2

PERCENT IMPROVEMENTS IN ESTIMATING σ_α^2

	$(\mu, \sigma_\alpha^2, \sigma_e^2)$	PI's $\delta^+_{\alpha,1}$ vs (4.1)	$\delta^+_{\alpha,2}$ vs (4.1)	$\delta^+_{\alpha,2}$ vs (4.3)	(3.5) vs (4.2)
I = 2, J = 5	(0, 1, 1)	92.22	91.96	17.67	-53.73
	(0, 1, .1)	66.49	66.69	1.10	3.19
	(0, 1, 1)	70.61	71.31	2.10	3.78
	(1, .1, 1)	90.80	90.32	1.87	-53.68
	(1, 1, .1)	66.44	66.64	.90	3.29
	(1, 1, 1)	71.51	72.14	1.74	3.44
I = 5, J = 5	(0, .1, 1)	68.55	67.72	14.90	-172.38
	(0, 1, .1)	33.33	33.61	1.57	6.46
	(0, 1, 1)	35.24	37.03	0.03	-174.08
	(1, 1, .1)	32.08	32.33	0.63	6.56
	(1, 1, 1)	33.42	35.08	0.71	3.59
I = 10, J = 5	(0, .1, 1)	44.49	43.92	8.73	-336.81
	(0, 1, .1)	19.55	19.73	1.35	3.77
	(0, 1, 1)	19.39	20.88	1.33	-27.04
	(1, .1, 1)	40.47	39.21	0.00	-338.97
	(1, 1, .1)	18.81	18.91	0.10	3.60
	(1, 1, 1)	18.39	19.68	0.20	-25.85

REFERENCES

Albert, A. (1976). When is a sum of squares an analysis of variance? Ann. Statist. 4, 775-778.

Berger, J. O. (1980). Improving on inadmissible estimators in continuous exponential families with applications to simultaneous estimation of gamma scale parameters. Ann. Statist. 8, 545-571.

Box, G. E. P. and Tiao, G. C. (1973). Bayesian inference in statistical analysis. Addison Wesley, Reading, Mass.

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

Das Gupta, A., Dey, D. K. and Gelfand, A. E. (1987). A new inadmissibility theorem with applications to estimation of survival and hazard rates and means in the scale parameter family. To appear in Sankhya.

Dey, D. K. and Gelfand, A. E. (1987). Improved estimators in simultaneous estimation of scale parameters. Stanford Univ. Tech. Rpt. #399.

Gelfand, A. E. (1987). Estimation of a restricted variance ratio. To appear in Proceedings Second Intl. Tampere Conference in Statistics.

Gelfand, A. E. and Dey, D. K. (1987a). On the estimation of a variance ratio. To appear in J. Statist. Planning and Inference.

Gelfand, A. E. and Dey, D. K. (1987b). Improved estimation of the disturbance variance in a linear regression model. To appear in Journal of Econometrics.

Graybill, F. A. (1954). On quadratic estimates of variance components. Annals of Math. Statist. 25, 367-372.

Graybill, F. A. and Wortham, A. W. (1956). A note on uniformly best unbiased estimators for variance components. Journal Amer. Statist. Assn. 51, 266-268.

Harville, D. A. (1977). Maximum likelihood approaches to variance component estimation and related problems. Journal Amer. Statist. Assn. 72, 330-338.

Harville, D. A. (1984). A generalized version of Albert's theorem, with applications to the mixed linear model. In Experimental Design Statistical Models and Genetic Statistics, Essays in Honor of Oscar Kempthorne (K. Hinkelmann, ed.), 231-238, Dekker, NY.

Herbach, L. H. (195). Properties of model II-type analysis of variance tests, A: Optimum nature of the F-test for model II in the balanced case. Annals of Math. Statist. 30, 939-959.

Hill, B. M. (1965). Inference about variance components in the one-way model. Journal Amer. Statist. Assoc. 60, 806-825.

Klonecki, W. and Zontek, S. (1985). On admissible estimation of parametric functions for mixed linear models. Proc. Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer, Vol. 2, 693-728, Wadsworth, Monterey, Calif.

Klonecki, W. and Zontek, S. (1987). Inadmissibility Results for Linear Simultaneous Estimation in the Multiparameter Gamma Distribution. Institute of Mathematics, polish Academy of Sciences, Warsaw.

Klotz, J. H., Milton, R. C. and Zacks, S. (1969). Mean square efficiency of estimators of variance components. Journal Amer. Statist. Assn. 64, 1383-1402.

Loh, W. (1986). Improved estimators for ratios of variance components. Journal Amer. Statist. Assn. 81, 699-702.

Searle, S. R. (1971). Linear Models. J. Wiley and Sons, New York.

Stein, C. (1964). Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. Ann. Inst. Statist. Math. 41, 385-388.

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

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER 434	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Improved Estimation Of Variance Components In Mixed Models		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Alan E. Gelfand and Dipak K. Dey		8. CONTRACT OR GRANT NUMBER(s) N00014-89-J-1627
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, CA 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-042-267
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics & Probability Program Code 1111		12. REPORT DATE August 30, 1990
		13. NUMBER OF PAGES 22
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

variance components; ANOVA decomposition; balanced designs; squared error loss.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Taking Albert's (1976) formulation of a mixed model ANOVA, we consider improved estimation of the variance components for balanced designs under squared error loss. Two approaches are presented. One extends the ideas of Stein (1964). The other is developed from the fact that variance components can be expressed as linear combinations of chi-square scale parameters. Encouraging simulation results are presented.

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)