The Annals of Mathematical Statistics
1971, Vol. 42, No. 4, 1379–1402

FORMAL BAYES ESTIMATION WITH APPLICATION TO A RANDOM EFFECTS MODEL

BY STEPHEN PORTNOY
Harvard University

1. Introduction. In this paper we will consider some techniques of formal Bayes estimation and apply them to the estimation of components of variance in the one way layout random effects model of the analysis of variance. In particular, we will consider the following problem in canonical form: we observe

$$\bar{Y} \sim \mathfrak{N}(\mu, (\sigma_e + J\sigma_a)/IJ), S_1 \sim \sigma_e \chi^2_{I(J-1)}, \quad \text{and} \quad S_2 \sim (\sigma_e + J\sigma_a)\chi^2_{I-1}$$

where I and J are positive integers (the number of treatments and replications respectively), μ is real, and σ_e and σ_a are positive. We want to find estimates for σ_e and σ_a using essentially mean squared error as a measure of performance.

The problem of estimating σ_a and σ_e is not new, and minimum variance unbiased estimates and maximum likelihood estimates are well known. However, one can look for improvements; and in the estimation of σ_a a special problem arises; the unbiased estimate may be negative and the maximum likelihood estimate may be exactly zero. This particular problem has been considered recently by a number of investigators (see, for example, Scheffé [12] and Thompson [18]) and various interpretations for such estimates have been suggested. However, in problems where estimates are really desired, use of such estimators seems to me to be unacceptable. To solve this problem we will consider the use of formal Bayes estimators (i.e. Bayes estimators versus priors which are not necessarily finite), which will be strictly positive. We will show that certain formal Bayes estimators both of σ_a and σ_e have good mean squared error properties and can seriously be recommended.

Inferences about σ_e and σ_a from a Bayesian viewpoint have been recently presented by Box and Tiao [1], Hill [4], Stone and Springer [17], and Tiao and Tan [19]. The methods described in these papers are generally based on use of the Jeffreys' prior. We will later compare these methods with ones considered here and give reasons why the present methods should be generally preferable.

The techniques used here are special cases of more general considerations applicable whenever the statistical problem is invariant under a group of transformations which does not act transitively on the parameter space (i.e. in problems where there is not a unique best invariant procedure). The analysis of variance problem considered here is easily seen to be invariant under location and scale changes (if invariant loss functions are used); that is, under transformations on the parameter space $\{(\mu, \sigma_e, \sigma_a): -\infty < \mu < \infty, \sigma_e > 0, \sigma_a > 0\}$ described by

$$\mu \to a\mu + b, \sigma_e \to a^2\sigma_e, \sigma_a \to a^2\sigma_a.$$

Received January 19, 1970; revised February 8, 1971.

However, here invariance only reduces the parameter space to the range of the maximal invariant which we will take to be $\gamma = \sigma_e/(\sigma_e+J\sigma_a)$. In general, any invariant procedure will have a risk function which is a function only on the maximal invariant (in our case, the risk will be a function of γ). Thus, we can assume that there is a prior distribution on the space of the maximal invariant and ask for the invariant procedure minimizing the expected risk under the assumed prior distribution. Zidek [22] shows that such Bayes invariant procedures are actually formal Bayes procedures with respect to a prior measure constructed from the assumed prior and right Haar measure on the group. Zacks [21] considers such formal Bayes estimators of σ_e and σ_a and characterizes their structure. Here we will consider a particular class of such estimators and will describe certain optimality properties related to mean squared error.

We first list a general multi-parameter admissibility theorem which is later applied to the present particular problem. This theorem gives sufficient conditions for admissibility and seems to be particularly applicable in proving admissibility of these Bayes invariant procedures when the group is one dimensional. It is a generalization of the work of Stein [15] and Zidek [23] and is proved using the same argument given in Zidek [23].

In Section 2, a class of estimators of σ_a of the following form is presented:

$$\phi_{(b)}(S_1,S_2) = \frac{1}{J}\left\{\frac{S_2}{(I+3)+2b} - \frac{S_1}{I(J-1)-4-2b} + \frac{K(S_1+S_2)}{F_b(S_1/(S_1+S_2))}\right\}$$

where, letting $c = \frac{1}{2}(IJ+1)$ and $d = \frac{1}{2}(I+3)+b$,

$$K = \frac{c-1}{cd(c-d-1)}, \qquad F_b(A) = \frac{(1-I_A(c-d,d+1))B(c-d,d+1)}{A^{c-d}(1-A)^{d+1}}$$

and where $I_x(p,q)$ denotes the incomplete beta function. For mean squared error, the estimator $\phi_{(b)}$ is admissible among location invariant procedures for $-1 < b < \frac{1}{2}I(J-1)-2$. In Section 3, the particular estimate $\phi_{(-1)}$ is shown to be admissible among scale and location invariant procedures, nearly minimax, and as good as possible as $\sigma_e/\sigma_a \to 0$.

In Section 4, related estimators of σ_e are discussed and conclusions are given in Section 5. Most of these conclusions are based on calculations of mean squared errors presented in the appendix.

We now list the general admissibility results used in subsequent sections. They are most reasonably stated in the following framework of statistical decision theory: there are measurable spaces $(\mathscr{X},\mathscr{B}_{\mathscr{X}})$ (the observation space), $(\mathscr{T},\mathscr{B}_{\mathscr{T}})$ (the parameter space), and the action space is the real line, R, (together with the Borel sets). Consider a loss function, $L\colon \mathscr{T}\times\mathscr{A}\times\mathscr{X} \to [0,\infty)$, of the form

$$L(\theta,a,x) = v(\theta,x)(a-g(\theta))^2$$

where $v\colon \mathscr{T}\times\mathscr{X} \to (0,\infty)$ and $g\colon \mathscr{T} \to R$ are measurable functions. Suppose the distributions are given by densities defined by $p\colon \mathscr{X}\times\mathscr{T} \to [0,\infty)$ satisfying

$$\int p_\theta(x)\,d\mu(x) = 1 \qquad \text{for all } \theta\in T$$

where μ is a sigma-finite measure on $\mathscr{B}_{\mathscr{X}}$; and assume that

$$p_\theta(x) > 0 \qquad \text{for all } x \in \mathscr{X} \quad \text{and all } \theta \in \mathscr{T}. \tag{1.1}$$

We restrict ourselves to non-randomized decision rules, which are measurable functions $\phi: \mathscr{X} \to \mathscr{A}$, and define the risk of ϕ to be

$$R(\phi, \theta) = \int L(\theta, \phi(x), x) p_\theta(x)\, d\mu(x).$$

The rule ϕ is said to be admissible if there is no rule ϕ^* such that

$$R(\phi^*, \theta) \leqq R(\phi, \theta) \qquad \text{for all } \theta \in \mathscr{T} \tag{1.2}$$

with strict inequality for some $\theta_0 \in \mathscr{T}$. If Π is a measure on $\mathscr{B}_{\mathscr{T}}$, a rule ϕ is said to be almost admissible with respect to Π if there is no ϕ^* for which (1.2) holds with strict inequality on a set of positive Π measure. Since the loss function is strictly convex in a for every x and θ, (1.1) implies that if there is a measure Π with respect to which ϕ is almost admissible, then ϕ is admissible.

With this notation, we now state Lemma 1.1, whose proof appears in James and Stein [7] on page 371.

LEMMA 1.1. *Let* Π *be a sigma-finite measure on* $\mathscr{B}_{\mathscr{T}}$ *and let* ϕ_0 *be an arbitrary decision rule. Let* $\mathscr{C} \subset \mathscr{B}_{\mathscr{T}}$ *be a countable covering of* $\mathscr{T}$ *by sets of finite* Π *measure. If for every* $C \in \mathscr{C}$ *and every* $\varepsilon > 0$, *there is a function* $f: \mathscr{T} \to [0, \infty)$ *satisfying* (1.3), (1.4), *and* (1.5) *below, then* ϕ_0 *is almost admissible with respect to* Π; *and, hence, admissible.*

$$f(\theta) \geqq 1 \qquad \text{for all } \theta \in C. \tag{1.3}$$

$$\int R(\phi_0, \theta) f(\theta)\, d\Pi(\theta) < \infty. \tag{1.4}$$

$$K(f) \equiv \int\int v(\theta, x)(\phi_0(x) - \phi_f(x))^2 p_\theta(x) f(\theta)\, d\Pi(\theta)\, d\mu(x) < \varepsilon \tag{1.5}$$

where ϕ_f is the Bayes rule with respect to the measure $f d\Pi$.

In James and Stein [7] $K(f)$ is further calculated to be

$$K(f) = \int d\mu(x) \frac{\{\int(\phi_0(x) - g(\theta'))v(\theta', x) p_{\theta'}(x) f(\theta')\, d\Pi(\theta')\}^2}{\int v(\theta, x) p_\theta(x) f(\theta)\, d\Pi(\theta)}. \tag{1.6}$$

We now state a multiparameter generalization of the theorem of Zidek [23].

Let $\mathscr{T} = (\underline{\theta}, \bar{\theta}) \times \mathscr{T}_0$ where $(\underline{\theta}, \bar{\theta})$ is an interval (not necessarily finite) in the real line and $(\mathscr{T}_0, \mathscr{B}_{\mathscr{T}_0})$ is a measurable space. Consider a sigma-finite measure Π on $\mathscr{T}$ of the form

$$d\Pi(\theta) = \pi(\theta_1, \theta_2)\, dv(\theta_2) d\theta_1 \qquad \theta_1 \in (\underline{\theta}, \bar{\theta}) \text{ and } \theta_2 \in \mathscr{T}_0 \tag{1.7}$$

where v is a sigma-finite measure on $\mathscr{B}_{\mathscr{T}_0}$. Suppose the formal Bayes rule, ϕ_Π, is well defined by (1.8).

$$\phi_\Pi(x) = \frac{\int\int g(\theta_1, \theta_2) v(\theta_1, \theta_2, x) p_{(\theta_1, \theta_2)}(x) \pi(\theta_1, \theta_2)\, dv(\theta_2)\, d\theta_1}{\int\int v(\theta_1, \theta_2, x) p_{(\theta_1, \theta_2)}(x) \pi(\theta_1, \theta_2)\, dv(\theta_2)\, d\theta_1} \tag{1.8}$$

and define, for $\theta_1 \in (\underline{\theta}, \bar{\theta})$ and $x \in \mathscr{X}$,

(1.9) $h_1(\theta_1, x) =$

$$\int_{\theta_1}^{\bar{\theta}} \int_{\mathscr{T}_0} (\phi_\Pi(x) - g(\theta_1', \theta_2')) p_{(\theta_1', \theta_2')}(x) v(\theta_1', \theta_2', x) \pi(\theta_1', \theta_2') \, d\nu(\theta_2') \, d\theta_1'$$

$$h_2(\theta_1, x) = \int_{\mathscr{T}_0} p_{(\theta_1, \theta_2)}(x) v(\theta_1, \theta_2, x) \pi(\theta_1, \theta_2) \, d\nu(\theta_2) \tag{1.10}$$

$$\lambda(\theta_1) = \int \varphi_0 E_{(\theta_1, \theta_2)} \left\{ \frac{h_1(\theta_1, X)}{h_2(\theta_1, X)} \right\}^2 \pi(\theta_1, \theta_2) v(\theta_1, \theta_2, X) \, d\nu(\theta_2). \tag{1.11}$$

Note that condition (1.1) implies that $h_2(\theta_1, x) > 0$ for all $\theta_1 \in (\underline{\theta}, \bar{\theta})$ and $x \in \mathscr{X}$.

THEOREM 1.1. *Consider the statistical decision theory problem described above (including the restriction given in* (1.1)). *Suppose* $\lambda(\theta_1)$ (*see* (1.11)) *is a continuous function of* θ_1 *on* $(\underline{\theta}, \bar{\theta})$; *and suppose further that for every compact sub-interval* $[a_0, b_0] \subset (\underline{\theta}, \bar{\theta})$

$$\int_{a_0}^{b_0} \int_{\mathscr{T}_0} R(\phi_\Pi, (\theta_1, \theta_2)) \pi(\theta_1, \theta_2) \, d\nu(\theta_2) \, d\theta_1 < \infty. \tag{1.12}$$

Suppose also that for every $c \in (\underline{\theta}, \bar{\theta})$ *conditions* (A) *and* (B) *hold*;

(A) $\int_c^{\bar{\theta}} \int_{\mathscr{T}_0} R(\mathscr{T}_\Pi, (\theta_1, \theta_2)) \pi(\theta_1, \theta_2) \, d\nu(\theta_2) \, d\theta_1 = \infty \Rightarrow \int_c^{\bar{\theta}} \frac{d\theta_1}{\lambda(\theta_1)} < \infty.$

(B) $\int_{\underline{\theta}}^{c} \int_{\mathscr{T}_0} R(\phi_\Pi, (\theta_1, \theta_2)) \pi(\theta_1, \theta_2) \, d\nu(\theta_2) \, d\theta_1 = \infty \Rightarrow \int_{\underline{\theta}}^{c} \frac{d\theta_1}{\lambda(\theta_1)} = \infty.$

Then ϕ_Π *is admissible.*

The proof of this theorem is essentially the same as that given in Zidek [23] and is given in detail in Portnoy [10].

2. Estimation of the "between" component of variance. We now apply Theorem 1.1 to the estimation of the components of variance in the one way random effects model (Model II) in analysis of variance. In particular we consider the following statistical problem in canonical form: we observe

$$\bar{Y} \sim \mathfrak{N}(\mu, (\sigma_e + J\sigma_a)/IJ), S_1 \sim \sigma_e \chi^2_{I(J-1)}, S_2 \sim (\sigma_e + J\sigma_a)\chi^2_{I-1} \tag{2.1}$$

where μ is real, $\sigma_e > 0$, $\sigma_a > 0$ and I and J are positive integers. We wish to tiesmate σ_a with loss proportional to squared error.

First we will restrict consideration to location and scale invariant procedures; that is, in particular, estimators not depending on $\bar{Y}$. Although the work of Stein [16] indicates that such procedures are likely to be inadmissible and Zacks [21] shows that location and scale invariant estimates of σ_e are, in fact, inadmissible (among all estimators), we know that consideration of $\bar{Y}$ cannot lead to substantial improvements (see Brown [2] for some particular calculations).

So consider the reduced problem where we observe

$$S_1 \sim \sigma_e \chi^2_{I(J-1)} \quad \text{and} \quad S_2 \sim (\sigma_e + J\sigma_a)\chi^2_{I-1}. \tag{2.2}$$

For ease in integrating we will use the following parameterization:

$$\alpha = \frac{1}{2\sigma_e}, \qquad \beta = \frac{1}{2(\sigma_e + J\sigma_a)}, \qquad \gamma = \frac{\beta}{\alpha}. \tag{2.3}$$

Then the densities of S_1 and S_2 become

$$p_{\alpha\beta}(S_1, S_2) = \alpha^{\frac{1}{2}(IJ-1)}\gamma^{\frac{1}{2}(I-1)}e^{-\alpha(S_1+\gamma S_2)} \tag{2.4}$$

with respect to the measure μ given by

$$d\mu(S_1, S_2) \propto S_1^{\frac{1}{2}I(J-1)-1}S_2^{\frac{1}{2}(I-1)-1}dS_1\,dS_2. \tag{2.5}$$

We take the parameter space to be

$$\mathscr{T} = \{(\alpha,\gamma): 0 < \alpha < \infty, 0 < \gamma < 1\}. \tag{2.6}$$

We wish to estimate

$$\sigma_a = \frac{1}{2J\alpha}\left(\frac{1}{\gamma} - 1\right);$$

so we take the loss function

$$L((\alpha,\gamma),\phi) = \alpha^2\gamma^2\left(\phi - \frac{1}{2J\alpha}\left(\frac{1}{\gamma} - 1\right)\right)^2. \tag{2.7}$$

Then the problem is invariant under scale changes; that is, under the transformations

$$S_1 \to \frac{1}{c}S_1, \qquad S_2 \to \frac{1}{c}S_2; \qquad \alpha \to c\alpha \tag{2.8}$$

for $c > 0$. Thus, we can think of α as a scale parameter and of γ as a parameter indexing the orbits. The result of Zidek [22] indicates that we can find Bayes scale invariant procedures by considering prior measures with densities of the form $\alpha^{-1}f(\gamma)$. For ease of computation, we will take for the prior density

$$\pi(\alpha,\gamma) = \alpha^a\gamma^b \tag{2.9}$$

(with respect to Lebesgue measure restricted to $\mathscr{T}$) and we will calculate the Bayes rules $\phi_\pi = \phi_{(a,b)}$. We expect (from Zidek [22]) that $\phi_{(a,b)}$ should be reasonable (in this reduced problem) for $a = -1$ (the invariant Haar measure) and $b > -1$.

The estimator $\phi_{(a,b)}$ was calculated for certain a and b in Klotz, Milton, and Zacks [6] and in other papers. Similar calculations yield

$$\phi_{(a,b)}(S_1, S_2) = \frac{1}{2J}\left\{\frac{S_2}{d} - \frac{S_1}{c-d-1} + \frac{(c-1)}{cd(c-d-1)} \cdot \frac{(S_1+S_2)}{F_{(c+1,d)}(A)}\right\}. \tag{2.10}$$

where $A = S_1/(S_1+S_2)$, $c = \frac{1}{2}(IJ-1)+a+2$; $d = \frac{1}{2}(I-1)+b+2$, and

$$\begin{aligned} F_{(c+1,d)}(A) &= \int_0^1 \frac{\gamma^d\,d\gamma}{[A+\gamma(1-A)]^{c+1}} \\ &= \left(\frac{1}{A}\right)^{c-d}\left(\frac{1}{1-A}\right)^{d+1}(1-I_A(c-d,d+1))\beta(c-d,d+1) \end{aligned} \tag{2.11}$$

where $I_x(\cdot,\cdot)$ is the incomplete Beta function (see Pearson [9]; $I_x(p,q) = \int_0^x u^{p-1}(1-u)^{q-1}du/\beta(p,q))$ and $\beta(\cdot,\cdot)$ is the (complete) Beta function. We can also write

$$(2.12)\quad \phi_{(a,b)}(S_1,S_2) = \frac{1}{2J}\left[\frac{S_2}{d}\left(\frac{1-I_A(c-d,d)}{1-I_A(c-d,d+1)}\right)-\frac{S_1}{c-d-1}\left(\frac{1-I_A(c-d-1,d+1)}{1-I_A(c-d,d+1)}\right)\right]$$

which may be easier to compute from tables of the incomplete Beta function (e.g. Pearson [9]).

To prove admissibility of the appropriate $\phi_{(a,b)}$, we will not need explicit expressions for $\phi_{(a,b)}$. We will only require the following inequality:

$$(2.13)\qquad \phi_{(a,b)}(S_1,S_2) \leqq M(S_1+S_2)$$

where M is a constant depending only on a, b, I, and J, and is finite for $d > 0$ and $(c-d-1) > 0$.

This inequality follows from the fact that

$$(2.14)\qquad F_{(c+1,d)}(A) \geqq \int_0^1 \gamma^d\,d\gamma = \frac{1}{d+1}$$

(which follows since $A+\gamma(1-A) \leqq 1$ for $0 < \gamma < 1$). Further properties of $\phi_{(a,b)}$, particularly for $a = b = -1$, will be discussed in Section 3. We will now prove

THEOREM 2.1. *Consider the statistical problem described by* (2.4), (2.5), *and* (2.7). *If*

$$(2.15)\qquad a = -1 \qquad \textit{and}$$

$$(2.16)\qquad -1 < b < \tfrac{1}{2}I(J-1)-2$$

then $\phi_{(a,b)}$ *defined by* (2.10) *is admissible. That is, if* (2.15) *and* (2.16) *hold, the Bayes invariant rule* $\phi_{(a,b)}$ *is admissible in the class of location invariant estimators in the original problem.*

PROOF. We will apply Theorem 1.1 directly with $\theta_1 = \alpha$, $\theta_2 = \gamma$, $(\underline{\theta},\bar{\theta}) = (0,\infty)$, $\mathscr{T}_0 = (0,1)$, $\pi(\alpha,\gamma) = \alpha^a\gamma^b$, $v(\alpha,\gamma) = \alpha^2\gamma^2$, $X = (S_1,S_2)$, $p_\theta(x)$ given by (2.4), and ϕ_Π given by (2.10). Condition (1.12) can be checked as follows: By (2.13) (and (2.2)),

$$\begin{aligned} E_{(\alpha,\gamma)}L(\phi_\Pi,(\alpha,\gamma)) &= E_{(\alpha,\gamma)}\alpha^2\gamma^2\left(\phi_\Pi(S_1,S_2)-\frac{1}{2J\alpha}\left(\frac{1}{\gamma}-1\right)\right)^2 \\ (2.17)\qquad &\leqq \alpha^2\gamma^2E_{(\alpha,\gamma)}\phi_\Pi{}^2(S_1,S_2)+\frac{1}{4J^2}(1-\gamma)^2 \\ &\leqq M\alpha^2\gamma^2E_{(\alpha,\gamma)}(S_1+S_2)^2+\frac{1}{4J^2} \\ &\leqq M^* \end{aligned}$$

where M^* depends only on a, b, I, and J. Therefore, since π integrable on $C\times(0, 1)$ (with C compact) for $b > -1$, the integrated risk is also finite on such sets. It remains to investigate the function $\lambda(\alpha)$ given by (1.11). We have

$$(2.18)\quad \frac{h_1(\alpha, S_1, S_2)}{h_2(\alpha, S_1, S_2)} = \frac{\int_\alpha^\infty \int_0^1 \left(\phi_\Pi(S_1, S_2) - \frac{1}{2J\alpha'}\left(\frac{1-\gamma'}{\gamma'}\right)\right)\alpha'^c\gamma'^d \exp[-\alpha'(S_1+\gamma' S_2)]\,d\gamma'\,d\alpha'}{\int_0^1 \alpha^c\gamma'^d \exp[-\alpha(S_1+\gamma' S_2)]\,d\gamma'}.$$

Changing variables, let $T_1 = \alpha S_1$ and $T_2 = \alpha\gamma S_2$. Then $T_1 \sim \frac{1}{2}\chi^2_{I(J-1)}$, $T_2 \sim \frac{1}{2}\chi^2_{I-1}$, and

$$(2.19)\quad E_{(\alpha,\gamma)}\left\{\frac{h_1(\alpha, S_1, S_2)}{h_2(\alpha, S_1, S_2)}\right\}^2 = E_{(1,1)}\left\{\frac{h_1\left(\alpha, \frac{T_1}{\alpha}, \frac{T_2}{\alpha\gamma}\right)}{h_2\left(\alpha, \frac{T_1}{\alpha}, \frac{T_2}{\alpha\gamma}\right)}\right\}^2.$$

Furthermore, changing variables from (α', γ') to $x = \alpha'/\alpha$, $y = \gamma'/\gamma$, using the fact that ϕ_Π is scale invariant for any a and b,

$$(2.20)\quad B \equiv \frac{h_1\left(\alpha, \frac{T_1}{\alpha}, \frac{T_2}{\alpha\gamma}\right)}{h_2\left(\alpha, \frac{T_1}{\alpha}, \frac{T_2}{\alpha\gamma}\right)} = \frac{\frac{1}{\gamma}\int_1^\infty \int_0^{1/\gamma} \left(\phi_\Pi(\gamma T_1, T_2) - \frac{1}{2Jx}\left(\frac{1-\gamma y}{y}\right)\right)x^c y^d \exp[-(T_1+yT_2)x]\,dy\,dx}{\int_0^{1/\gamma} y^d \exp[-(T_1+yT_2)]\,dy}.$$

Furthermore from the definition of ϕ_Π, the integral over $(0, \infty)$ in the numerator of [2.20] is zero. So we also have

$$(2.21)\quad B = \frac{\frac{1}{\gamma}\int_0^1 \int_0^{1/\gamma} \left(\phi_\Pi(\gamma T_1, T_2) - \frac{1}{2Jx}\left(\frac{1-\gamma y}{y}\right)\right)x^c y^d \exp[-(T_1+yT_2)x]\,dy\,dx}{\int_0^{1/\gamma} y^d \exp[-(T_1+yT_2)]\,dy}.$$

Note that B is independent of α. Hence, $E_{(\alpha,\gamma)}\{h_1(\alpha, S_1, S_2)/h_2(\alpha, S_1, S_2)\}^2$ is a function of γ alone (i.e., is independent of α). Therefore, from (1.11),

$$(2.22)\quad \lambda(\alpha) = \int_0^1 E_{(\alpha,\gamma)}\left\{\frac{h_1(\alpha, S_1, S_2)}{h_2(\alpha, S_1, S_2)}\right\}^2 \alpha^{a+2}\gamma^{b+2}\,d\gamma$$
$$= \alpha^{a+2}\int_0^1 E_{(\alpha,\gamma)}\left(\frac{h_1(\alpha, S_1, S_2)}{h_2(\alpha, S_1, S_2)}\right)^2 \gamma^{b+2}\,d\gamma$$

and $\lambda(\alpha)$ is a continuous function of α whenever the integral in (2.22) is finite. I will later use (2.20) and (2.21) to show that

$$|B| \leqq \frac{1}{\gamma}(M_1 + M_2 T_1 + M_3 T_2) \quad \text{for } c-d-1>0, \quad d>0 \tag{2.23}$$

where M_1, M_2, M_3 are constants not depending on α and γ. From (2.23), it follows that

$$E(B^2) = E_{(1,1)}\left\{\frac{h_1\left(\alpha, \frac{T_1}{\alpha}, \frac{T_2}{\alpha\gamma}\right)}{h_2\left(\alpha, \frac{T_1}{\alpha}, \frac{T_2}{\alpha\gamma}\right)}\right\}^2 \leqq \frac{1}{\gamma^2} M^* \tag{2.24}$$

and, from (2.19), that

$$\lambda(\alpha) \leqq M^* \alpha^{a+2} \textstyle\int_0^1 \gamma^b \, d\gamma = \alpha^{a+2} M^*/(b+1) \tag{2.25}$$

which is finite for $b > -1$. To prove ϕ_{Π} admissible it remains to check conditions (A) and (B) of Theorem 1.1. From (2.10), $\phi_{\Pi}(S_1, S_2) \geqq S_2/d - S_1/(c-d-1)$, so calculating expectations,

$$\begin{aligned} R(\phi_{\Pi}, (\alpha, \gamma)) &= \alpha^2 \gamma^2 E_{(\alpha,\gamma)}\left(\phi_{\Pi}(S_1, S_2) - \frac{1}{2J\alpha\gamma}(1-\gamma)\right)^2 \\ &\geqq \frac{(I-1)\gamma^2}{2d^2} + \frac{I(J-1)}{2(c-d-1)^2}. \end{aligned} \tag{2.26}$$

Therefore, combining this with (2.17),

$$\textstyle\int_0^1 R(\phi_{\Pi}, (\alpha, \gamma)) \pi(\alpha, \gamma) \, d\gamma \approx M^{**} \alpha^a. \tag{2.27}$$

For conditions (A) and (B) we want $1/\lambda(\alpha)$ to be non-integrable whenever $\int_0^1 R(\phi_{\Pi}, (\alpha, \gamma))\pi(\alpha, \gamma) d\gamma$ is; that is, we want $1/\alpha^{a+2}$ to be non-integrable whenever α^a is. But this holds if and only if $a = -1$. Thus, ϕ_{Π} is admissible (by Theorem 1.1) for $a = -1$, $b > -1$ and $d > 0$, $c-d-1 > 0$. These conditions, by definition, are just (2.15) and (2.16); hence, to complete the proof of this theorem we need only prove (2.23), for which we will give the following rather technical and lengthy argument:

First note that for $s, t > 0$

$$\begin{aligned} \int_s^\infty u^n e^{-tu} \, du &= \frac{1}{t^{n+1}} \int_{st}^\infty v^n e^{-v} \, dv \\ &= \frac{1}{t^{n+1}} \int_{st}^\infty v^{-2}(v^{n+2} e^{-v}) \, dv \\ &\leqq \frac{1}{t^{n+1}} (st)^{n+2} e^{-st} \int_{st}^\infty \frac{1}{v^2} \, dv \qquad \text{for } st \geqq n+2 \\ &= s^{n+1} e^{-st} \qquad \text{for } st \geqq n+2. \end{aligned} \tag{2.28}$$

Therefore, taking the outer integral (over x) in the numerator of (2.20), we have, for $T_1 \geqq c+2$

(2.29) $$\gamma \cdot (\text{numerator}) \leqq \int_0^{1/\gamma} \left(\phi_\Pi(\gamma T_1, T_2) + \frac{1}{2Jy} + \frac{\gamma}{2J} \right) y^d \exp\left[-(T_1 + yT_2)\right] dy.$$

So

(2.30) $$\gamma \cdot |B| \leqq \phi_\Pi(\gamma T_1, T_2) + \frac{\gamma}{2J} + \frac{\frac{1}{2J}\int_0^{1/\gamma} y^{d-1} e^{-yT_2}\, dy}{\int_0^{1/\gamma} y^d e^{-yT_2}\, dy}.$$

Integrating by parts, we can bound the last term in (2.30) by

(2.31) $$\frac{T_2}{2Jd} + \frac{1}{2Jd} \frac{(1/\gamma)^d e^{-T_2/\gamma}}{\int_0^{1/\gamma} y^d e^{-yT_2}\, dy},$$

and since

(2.32) $$\int_0^{1/\gamma} y^d e^{-yT_2}\, dy \geqq e^{-T_2/\gamma} \int_0^{1/\gamma} y^d\, dy = \frac{1}{d+1} e^{-T_2/\gamma} (\gamma)^{-(d+1)}$$
$$\geqq \frac{1}{d+1} e^{-T_2/\gamma} (\gamma)^{-d}$$

we have, for $T_1 \geqq c+2$, $d > 0$ (using (2.13)),

(2.33) $$\gamma \cdot |B| \leqq M_1 + M_2 T_1 + M_3 T_2.$$

Now, for $T_1 < c+2$, we will use (2.21). Since $e^{-T_1 x} \leqq 1$, taking absolute values, we have

(2.34) $$\gamma \cdot |B| \leqq e^{c+2} \frac{\int_0^1 \int_0^{1/\gamma} \left| \phi_\Pi(\gamma T_1, T_2) - \frac{1-\gamma y}{2Jxy} \right| (xy)^d x^{c-d} e^{-xT_2 y}\, dy\, dx}{\int_0^{1/\gamma} y^d e^{-yT_2}\, dy}.$$

Letting $z = xy$ in the inner integral of the numerator,

(2.35) $$\gamma \cdot |B| \leqq e^{c+2} \frac{\int_0^1 \left[\int_0^{x/\gamma} \left(\phi_\Pi(\gamma T_1, T_2) + \frac{1}{2Jz} \right) z^d e^{-T_2 z}\, dz \right] x^{c-d-1}\, dx}{\int_0^{1/\gamma} y^d e^{-yT_2}\, dy}.$$

The integrand is maximized over $0 \leqq x \leqq 1$ at $x = 1$ (for $c-d-1 > 0$). So

(2.36) $$\gamma \cdot |B| \leqq e^{c+2} \frac{\int_0^{1/\gamma} \left(\phi_\Pi(\gamma T_1, T_2) + \frac{1}{Jz} \right) z^d e^{-T_2 z}\, dz}{\int_0^1 y^d e^{+T_2 y}\, dy}$$
$$\leqq M_1' + M_2' T_1 + M_3' T_2$$

where we have used the same argument as we used to get (2.33). Therefore (2.23) follows from (2.33) and (2.36) and the proof of the theorem is complete. □

3. A minimax, formal Bayes estimate of σ_a admissible among scale and location invariant estimators. In Section 2 we discussed a class of formal Bayes estimators $\phi_{(-1,b)}$ admissible among location invariant procedures. In this section we consider the limiting case as $b \to -1$. In particular we will consider the estimator $\phi_{(-1,-1)}$, which, as is often the case with limits of Bayes rules, we will show to be admissible among fully invariant procedures. Since we are now taking $a = -1$ in $\phi_{(a,b)}$, we will suppress dependence on a throughout this section and refer to ϕ_b instead of $\phi_{(a,b)}$.

We now show that $\phi_{(-1)}$ is admissible among location and scale invariant estimators. To do this, we will reduce the problem by invariance and apply Lemma 1.1 and (1.6) directly. We first discuss the reduced statistical problem.

If ϕ is scale (and location) invariant, there is a function $\Psi: R \to R$ such that ϕ can be written

$$\phi(S_1, S_2) = (S_1 + S_2)\Psi\left(\frac{S_1}{S_1 + S_2}\right). \tag{3.1}$$

(Note: we let Ψ_b denote the appropriate function of the maximal invariant corresponding to ϕ_b.) Then, the risk of any invariant rule is given by

$$\begin{aligned} R(\phi, (\alpha, \gamma)) &= E\alpha^2\gamma^2\left(\phi(S_1, S_2) - \frac{1}{2J\alpha\gamma}(1-\gamma)\right)^2 \\ &= E\left((T_1 + T_2)\Psi\left(\frac{T_1}{T_1 + T_2}\right) - \frac{1}{2J}(1-\gamma)\right)^2 \end{aligned} \tag{3.2}$$

where

$$T_1 \sim \gamma/2\chi^2_{I(J-1)} \quad \text{and} \quad T_2 \sim \tfrac{1}{2}\chi^2_{I-1}. \tag{3.3}$$

If we let

$$U = T_1 + T_2, \qquad V = \frac{T_1}{T_1 + T_2} \tag{3.4}$$

then the joint distribution of (U, V) can be found to have densities

$$\begin{aligned} p_\gamma(U, V) \propto (1/\gamma)^{\frac{1}{2}I(J-1)} U^{\frac{1}{2}(IJ-1)-1} V^{\frac{1}{2}I(J-1)-1} \\ \cdot (1-V)^{\frac{1}{2}(I-1)-1} e^{(1/\gamma)U(V+\gamma(1-V))}. \end{aligned} \tag{3.5}$$

That is, the conditional distribution of U given V is given by

$$U \mid V \sim \frac{\gamma}{2(V + \gamma(1-V))}\chi^2_{(IJ-1)}. \tag{3.6}$$

Therefore, from (3.2)

$$
\begin{aligned}
R(\varphi,(\alpha,\gamma)) &= EE\left\{\left(U\Psi(V)-\frac{1}{2J}(1-\gamma)\right)^2 \mid V\right\} \\
&= E\left\{\frac{\gamma^2(IJ-1)(IJ+1)}{4(V+\gamma(1-V))^2}\Psi^2(V)-\frac{(1-\gamma)\gamma(IJ-1)}{2J(V+\gamma(1-\gamma))}\Psi(V)+\frac{(1-\gamma)^2}{4J^2}\right\} \\
&= E\left\{\frac{\gamma^2(IJ-1)(IJ+1)}{4(V+\gamma(1-V))^2}\left[\Psi(V)-\frac{(V+\gamma(1-V))(1-\gamma)}{J\gamma(IJ+1)}\right]^2\right. \\
&\quad \left.+\frac{(1-\gamma)^2}{4J^2}-\frac{(1-\gamma)^2(IJ-1)}{4J^2(IJ+1)}\right\}.
\end{aligned} \tag{3.7}
$$

We are thus led to consider the statistical problem where we observe V having densities (parametrized by γ, $0 < \gamma < 1$)

$$
p_\gamma(V) \propto \frac{\gamma^{\frac{1}{2}(I-1)}V^{\frac{1}{2}I(J-1)-1}(1-V)^{\frac{1}{2}(I-1)-1}}{(V+\gamma(1-V))^{\frac{1}{2}(IJ-1)}} \tag{3.8}
$$

(with respect to Lebesgue measure restricted to $\{V:0 \leqq V \leqq 1\}$), and we wish to estimate $(V+\gamma(1-V))(1-\gamma)/J\gamma(IJ+1)$ with loss

$$
L(\Psi,\gamma,V) = \frac{\gamma^2(IJ-1)(IJ+1)}{4(V+\gamma(1-V))^2}\left[\Psi-\frac{(V+\gamma(1-V))(1-\gamma)}{J\gamma(IJ+1)}\right]^2. \tag{3.9}
$$

It is clear from (3.7) that $\phi(S_1, S_2)$ is admissible among invariant procedures (with squared error as loss) if and only if the corresponding estimate $\Psi(V)$ is admissible in the above problem.

We remark that $\Psi_{(b)}$, corresponding to the Bayes invariant procedure ϕ_b, is actually the Bayes procedure (in the reduced problem) with respect to the prior distribution γ^b Thus, to prove admissibility of $\psi_{(-1)}$, and hence of $\phi_{(-1)}$, we can apply Lemma 1.1 with parameter $\theta = \gamma(\mathscr{T} = (0, 1))$, Π equal to Lebesgue measure (restricted to (0, 1)), p_θ given by (3.8) and L given by (3.9) (and we will use $f(\gamma) = \gamma^b$). To apply Lemma 1.1, we will need a bound for $(\psi_{(-1)}-\psi_b)^2$. The following technical lemma provides this bound:

LEMMA 3.1. *Let*

$$
\psi_b\left(\frac{S_1}{S_1+S_2}\right) = \left(\frac{1}{S_1+S_2}\right)\phi_b(S_1,S_2),
$$

where ϕ_b *is given by* (2.10) (*with* $a = -1$). *Then there are* $p_0 > 0$ *and constants* K_1 *and* K_2 (*independent of* V *and* p) *such that, for* $p < p_0$

$$
|\psi_{(-1)}(V)-\psi_{(-1+p)}(V)| \leqq p(K_1V+K_2(1-V)) \tag{3.10}
$$

for all V, $0 < V < 1$.

The proof of this lemma involves long and tedious calculations which are given in detail in Portnoy [10]. We now use this lemma to prove

THEOREM 3.2. *If $\phi_{(-1)}$ is defined by* (2.10) (*for $a = b = -1$*), *then it is admissible among scale and location invariant estimators in the original problem of estimating σ_a with squared error loss.*

PROOF. As previously remarked, it is sufficient to prove that $\Psi_{(-1)}$ is admissible in the reduced problem (described by (3.8) and (3.9)). To do this, we will apply Lemma 1.1 with Π equal to Lebesgue measure (restricted to (0, 1)) and the covering of the parameter space consisting of just one set, $\{\mathscr{T}\}$. In particular, we will show that for any $\varepsilon > 0$, conditions (1.3), (1.4), and (1.5) are satisfied by $f(\gamma)$ of the form $f(\gamma) = \gamma^b$ for $-1 < b < -\frac{1}{2}$. First note that, for $-1 < b < -\frac{1}{2}$, $\gamma^b \geqq 1$ (for $0 < \gamma < 1$); and, hence, (1.3) is satisfied. Furthermore, we showed in Section 2 (see (2.17)) that $R(\phi_{(a,b)}, (\alpha, \gamma))$ was bounded for any a, b, I, and J for which $c-d-1 > 0$ and $d > 0$. Hence, if $I > 2$, $J > 2$, $R(\phi_{(-1)}, (\alpha, \gamma))$ is bounded. Therefore, from (3.7) it follows (since the expectation of the last two terms in the final expression are bounded) that $E_\gamma L(\Psi_{(-1)}, \gamma, V)$ is bounded where L is given by (3.9) (and the distribution by (3.8)). That is, the risk function in the reduced problem is bounded. Therefore, since $f(\gamma)$ is integrable, it follows that condition (1.4) holds. To check condition (1.5) we evaluate $K(f)$ (with p defined by $b = -1+p$) as follows:

$$\text{(3.11)}\quad K(f) = \int_0^1 \int_0^1 p_\gamma(V) \left\{ \frac{\gamma^2(IJ-1)(IJ+1)}{4(V+\gamma(1-V))^2} [\Psi_{(-1)}(V) - \Psi_{(-1+p)}(V)]^2 \right\} \cdot \gamma^{(-1+p)} \, dV \, d\gamma$$

$$\leqq K_3 p^2 \int_0^1 E_\gamma \left(\frac{\gamma^2(K_1 V + K_2(1-V))^2}{(V+\gamma(1-V))^2} \right) \gamma^{p-1} \, d\gamma$$

where the last inequality uses Lemma 2.1. Thus, for p small enough,

$$\text{(3.12)}\quad K(f) \leqq K_4 p^2 \textstyle\int_0^1 \gamma^{p-1} \, d\gamma = K_4 p \leqq \varepsilon.$$

Therefore, condition (1.5), and hence, Lemma 1.1 holds; thus $\psi_{(-1)}$ is admissible and the desired result follows.

Since $\phi_{(-1)}$ is a limit of Bayes procedures in the reduced problem, it might be expected to be a minimax among invariant rules for loss given by (2.7). However, calculations in Table I in the appendix indicate that this is not true. Nonetheless, these calculations also show that the risk of $\phi_{(-1)}$ is rarely greater than the lower bound for the minimax risk (for loss (2.7)) given in the following lemma.

LEMMA 3.3. *With the notation of Section* 2, *let $\phi(S_1, S_2) = (S_1+S_2)\psi(S_1/S_1+S_2))$ be an invariant rule and suppose the loss is given* (*for $M \geqq 0$*) *by*

$$\text{(3.13)}\quad L^{(M)}(\phi, (\alpha, \gamma)) = \frac{4\alpha^2\gamma}{(1+\gamma M)^2} \left(\phi - \frac{1-\gamma}{2J\alpha\gamma} \right)^2$$

$$\left(\text{i.e.,}\quad L^{(M)}(\phi, (\sigma_e, \sigma_a)) = \frac{1}{((M+1)\sigma_e + J\sigma_a)^2} (\phi - \sigma_a)^2\right).$$

If $R^{(M)}(\phi, \gamma)$ denotes the risk of ϕ for loss $L^{(M)}$, then, for any $M > 0$,

$$\lim_{\gamma\to 0} R^{(M)}(\phi, \gamma) \geqq \frac{2}{J^2(I+1)}. \tag{3.14}$$

PROOF. First note that since ϕ is invariant, $R^{(M)}$ is indeed a function of γ alone; and that since the densities (2.4) form an exponential family, $R^{(M)}(\phi, \gamma)$ is continuous for $\gamma \in (0, 1)$. Now, by invariance,

$$E_{(\alpha,\gamma)} \frac{4\alpha^2\gamma^2}{(1+\gamma M)^2}\left\{(S_1+S_2)\psi\left(\frac{S_1}{S_1+S_2}\right) - \frac{(1-\gamma)}{2J\alpha\gamma}\right\}^2 \tag{3.15}$$

$$= E\frac{1}{(1+\gamma M)^2}\left\{(\gamma T_1+T_2)\psi\left(\frac{\gamma T_1}{\gamma T_1+T_2}\right) - \frac{1-\gamma}{J}\right\}^2$$

where $T_1 \sim \chi^2_{I(J-1)}$ and $T_2 \sim \chi^2_{I-1}$. So, by Fatou's lemma,

$$\lim_{\gamma\to 0} R^{(M)}(\phi, \gamma) \geqq E \liminf_{\gamma\to 0}\left(\frac{1}{1+\gamma M}\right)^2\left((\gamma T_1+T_2)\psi\left(\frac{\gamma T_1}{\gamma T_1+T_2}\right) - \frac{1-\gamma}{J}\right)^2 \tag{3.16}$$

$$= E\left(CT_2 - \frac{1}{J}\right)^2 \geqq \frac{2}{J^2(I+1)}$$

where $C = \liminf_{\gamma\to 0} \psi(\gamma)$. □

We now note that this lower bound is not attained by $\phi_{(-1)}$ for loss (2.7) (i.e., for $L^{(M)}$ with $M = 0$). In particular, in Portnoy [10] it is shown that although

$$\lim_{\gamma\to 0} R^{(0)}(\phi_{(-1)}, \gamma) = \frac{2}{J^2(I+1)}, \tag{3.17}$$

it is also found that

$$\lim_{\gamma\to 0} \frac{d}{d\gamma} R^{(0)}(\phi_{(-1)}, \gamma) = \frac{8}{J^2(I+1)(IJ-I-2)} > 0. \tag{3.18}$$

Thus, $R^{(0)}(\phi_{(-1)}, \gamma) > 2/(J^2(I+1))$ for some $\gamma > 0$ (this is shown explicitly in calculations in the appendix). Hence, we can not conclude that $\phi_{(-1)}$ is minimax for loss (2.7). However, for loss $L^{(M)}$ (given by (3.13)) the risk $R^{(M)}(\phi_{(-1)}, \gamma)$ satisfies

$$R^{(M)}(\phi_{(-1)}, \gamma) = \frac{1}{(1+\gamma M)^2} f(\gamma) \tag{3.19}$$

where $f(\gamma) = R^{(0)}(\phi_{(-1)}, \gamma)$. Thus, taking derivatives and letting $\gamma \to 0$,

$$\lim_{\gamma\to 0} \frac{d}{d\gamma} R^{(M)}(\phi_{(-1)}, \gamma) = -2Mf(0) + f'(0) \tag{3.20}$$

where $f(0)$ and $f'(0)$ are given by (3.17) and (3.18) respectively. Therefore, for M large enough $R^{(M)}(\phi_{(-1)}, \gamma)$ is decreasing in γ for $0 < \gamma \leqq \gamma_0$ for some $\gamma_0 > 0$.

Furthermore, since $f(\gamma)$ is bounded (say, by K_0), it follows from (3.19) that for $\gamma > \gamma_0$,

$$R^{(M)}(\phi_{(-1)}, \gamma) \leqq \frac{1}{(1+\gamma M)^2} f(\gamma) \leqq \frac{1}{(1+\gamma_0 M)^2} K_0 \leqq \lim_{\gamma \to 0} R^{(M)}(\phi_{(-1)}, \gamma) \tag{3.21}$$

for M large enough (depending on γ_0 and K_0). Therefore, $R^{(M)}(\phi_{(-1)}, \gamma)$ is maximized at $\gamma = 0$; and hence, $\phi_{(-1)}$ is minimax among invariant procedures for loss $L^{(M)}$ for large enough M. We can now apply Kiefer's theorem (Kiefer [5]) to infer that $\phi_{(-1)}$ is minimax among all procedures (in the original problem) for loss $L^{(M)}$.

It should be noted that $2/(J^2(I+1))$ is merely a lower bound for the minimax risk and that it may not be attained. This would indicate that $\phi_{(-1)}$ is nearly minimax for loss (2.7). Furthermore, $\phi_{(-1)}$ does do as well as possible at $\gamma = 0$, and thus is optimal in this sense.

4. Estimates of the "within" component. We now discuss the problem of estimating σ_e in the previously described analysis of variance problem. As before, we observe

$$S_1 \sim \frac{1}{2\alpha} \chi^2_{I(J-1)} \quad \text{and} \quad S_2 \sim \frac{1}{2\alpha\gamma} \chi^2_{I-1} \tag{4.1}$$

where we are using the same parameterization

$$\alpha = \frac{1}{2\sigma_e} \quad \text{and} \quad \gamma = \frac{\sigma_e}{\sigma_e + J\sigma_a}. \tag{4.2}$$

Again we take the loss function to be

$$L(\phi, (\alpha, \gamma)) = \alpha^2 \gamma^2 \left(\phi - \frac{1}{2\alpha}\right)^2 \tag{4.3}$$

and consider formal Bayes estimators with respect to prior measures

$$d\Pi(\alpha, \gamma) = \alpha^a \gamma^b \, d\alpha \, d\gamma. \tag{4.4}$$

Again letting

$$c = \tfrac{1}{2}(IJ-1) + a + 2 \quad \text{and} \quad d = \tfrac{1}{2}(I-1) + b + 2 \tag{4.5}$$

we find (as before)

$$\phi_{(a,b)}(S_1, S_2) = \frac{S_1}{2(c-d-1)} \left[1 - \frac{1}{cAF_{(c+1,d)}(A)}\right] \tag{4.6}$$

where $A = S_1/(S_1+S_2)$ and $F_{(c+1,d)}(A)$ is given by (2.11). We also have, directly,

$$\phi_{(a,b)}(S_1, S_2) = \frac{S_1}{2(c-d-1)} \left[\frac{1 - I_A(c-d-1, d+1)}{1 - I_A(c-d, d+1)}\right] \tag{4.7}$$

where $I_A(\cdot, \cdot)$ is the incomplete beta function (again, see Pearson [9]).

Similar proofs to those in Section 2 can now be used to show admissibility of the appropriate $\phi_{(a,b)}$. However, in this case there is a slight difference. From (4.6) we have

$$\phi_{(a,b)}(S_1, S_2) \leqq \frac{S_1}{2(c-d-1)}. \tag{4.8}$$

Therefore, since $\phi_{(a,b)}$ is positive, the risk satisfies

$$R(\phi_{(a,b)}, (\alpha, \gamma)) = \alpha^2\gamma^2 E\left(\phi_{(a,b)}(S_1, S_2) - \frac{1}{2\alpha}\right)^2 \leqq M\gamma^2 \tag{4.9}$$

where M is a fixed constant depending on a, b, I, J and finite whenever $c-d-1 > 0$. That is, the risk is integrable with respect to Π as long as $b > -3$ (instead of $b > -1$). Furthermore, using similar reasoning to that in Section 2, we can find in this case (using notation of Sections 1 and 2),

$$E_{(\alpha,\gamma)}\left\{\frac{h_1(S_1, S_2, \alpha)}{h_2(S_1, S_2, \alpha)}\right\}^2 \leqq M^* \tag{4.10}$$

(in contrast to (2.24)). Therefore, Theorem 1.1 implies that $\phi_{(a,b)}$ is admissible for $a = -1$ and $b > -3$.

As before, we can show that the limiting estimator $\phi_{(-1,-3)}$ is admissible among scale (and location) invariant estimators. Furthermore, although using loss function (4.3) will force the risk to tend to zero as γ approaches zero, we can consider the following, perhaps more reasonable, loss function:

$$L^*(\phi, (\sigma_e, \sigma_a)) = \frac{1}{\sigma_e^2}(\phi - \sigma_e)^2. \tag{4.11}$$

Again, for this loss function, $\phi_{(-1,-3)}$ does as well as possible at $\gamma = 0$; and there will be related loss functions for which $\phi_{(-1,-3)}$ will be minimax. In fact, numerical calculations in the appendix seem to indicate that $\phi_{(-1,-3)}$ is actually minimax for the loss function given by (4.11), although I have been unable to prove this.

5. Conclusions. There is, I think, sufficient evidence to seriously recommend the use of a formal Bayes rule for estimating σ_a in the model II analysis of variance problem considered here. Although one may claim that point estimation of σ_a in this problem is not a serious or useful statistical consideration, there are problems, I think, where point estimates are really desired. For example, one may want to use the data to plan future experiments, or one may want to compare the data with some other results and perhaps estimate a correlation coefficient. In either of these problems, the use of a negative, or even a zero, estimate is unacceptable. Thus, a Bayes or formal Bayes estimate should be used. The formal Bayes estimates considered in Section 2 are scale and location invariant estimates with what I feel are adequate mean squared error properties. The theoretical considerations in Chapter 2 certainly indicate that no estimate can be substantially better than $\phi_{(-1)}$ (given by (2.10) for $a = b = -1$). Numerical calculations

(Table I) indicate that other estimates are better than $\phi_{(-1)}$ only very near $\sigma_e/(\sigma_e+J\sigma_a) = 1$. Elsewhere, $\phi_{(-1)}$ is just as good and is actually a definite improvement in small and moderate sample sizes for $\sigma_e/(\sigma_e+J\sigma_a) < \cdot 5$. Thus there seems to be adequate reason to seriously recommend the formal Bayes estimators of σ_a.

Recommendation of the formal Bayes estimators of σ_e does not seem quite so strongly indicated, primarily because in this problem there are already reasonably good estimators. In this case the maximum likelihood estimator, (A.7) in the appendix, or the related estimate, (A.8) in the appendix, have nearly as good mean squared error properties as any other estimate. The numerical calculations (Table II in the appendix) do indicate that the formal Bayes estimators can offer some improvement in certain cases. Furthermore, they do have the desirable property that formal Bayes estimates of σ_e are strictly less than formal Bayes estimates of $(\sigma_e+J\sigma_a)$. However, the fact that the formal Bayes estimators are more difficult to calculate, together with the fact that there are other estimators approximately as good seem to be sufficient reason not to seriously recommend their use in this case.

Actually, the most noticeable feature of the calculations listed in Table I is the extremely poor performance (in terms of mean squared error) of the posterior expected value of σ_a for the Jeffrey's prior suggested by Tiao and Tan [19] (see (A.5) in Table I). Calculations in Klotz, Milton and Zacks [6] show that much of this large mean squared error is due to the extremely large bias of (A.5). Nonetheless, the expected posterior variance is still 5 to 10 times larger than the mean squared error of $\phi_{(-1)}$. The behavior of the posterior distribution suggested by Tiao and Tan becomes clear if we note that the estimator (A.3), the mode of the Tiao–Tan posterior, is, on the average, reasonably close to the true value. Thus, the posterior distribution with respect to the Jeffreys' prior is centred approximately correctly but has far too large mean and variance (on the average).

For the case of the present estimator, $\phi_{(-1)}$, we actually find that the prior distribution (for $a = -1$, $b = -1$) corresponds exactly to the Jeffreys' prior, $d\sigma_e d\sigma_a/(\sigma_e(\sigma_e+J\sigma_a))$. However, $\phi_{(-1)}$ is not the posterior mean, but the posterior expected value of $\sigma_a/(\sigma_e+J\sigma_a)^2$ over the posterior expected value of $1/(\sigma_e+J\sigma_a)^2$, the denominator coming from the normalizing factor in the loss function. This is the same as taking the posterior mean for the prior $d\sigma_e d\sigma_a/((\sigma_e+J\sigma_a)^3\sigma_e)$. The reasonable size of the mean squared error of $\phi_{(-1)}$ shows that this latter posterior distribution is centered at least as well and has substantially smaller variance (on the average) than the posterior for the Jeffreys' prior. Thus, if one wishes to make inferences based on a posterior distribution, one can seriously recommend using the prior $d\sigma_e d\sigma_a/((\sigma_e+J\sigma_a)^3\sigma_e)$ instead of the Jeffreys' prior. This serves to justify, in my opinion, the use of squared error as loss: by using squared error and by taking an appropriate limit of what might be called Bayes invariant priors, one is assured of finding a posterior distribution which, on the average, is reasonably centered and has reasonably small variance. As this example shows, the Jeffreys' prior can lead to posterior distributions with mean and variance far too large.

One other comment should be made: in this case, the appropriate limit of Bayes invariant priors was exactly the Jeffreys' prior. This need not always be the case. In particular, the appropriate limit may be any of a large class of priors; and in this case was chosen strictly for ease of calculation. In other problems (e.g. the estimation of parameters of a gamma distribution with unknown shape and scale) the Jeffreys' prior may be computationally unfeasible; but there may be a reasonable limit of Bayes invariant priors (see Portnoy [11]).

In conclusion, we have presented a method (not yet completely well defined) for finding reasonable formal Bayes procedures. In problems invariant under a group, $\mathfrak{G}$ (hopefully as small as possible), with a parameter space of the form $\mathfrak{G} \times \mathscr{T}$, we take Haar measure on $\mathfrak{G}$ and multiply by an appropriate limit of probability measures on $\mathscr{T}$ (or a probability measure itself). We then take an invariant version of the squared error loss function (or some other reasonable loss function), chosen so that for some procedure, the risk is bounded above zero and less than infinity (so that the concept of a minimax procedure makes sense). We finally make inferences using the posterior distribution with respect to the prior distribution modified by any appropriate normalizing functions in the loss function. These techniques, in problems like the present one, should lead to reasonable statistical procedures, especially in problems where the existing procedures are inadequate.

APPENDIX

Numerical Calculations of Mean Squared Errors for Estimators of σ_a and σ_e

A computer program was used to calculate expectations with respect to the following distributions:

$$S_1 \sim \sigma_e \chi^2_{I(J-1)} \quad \text{and} \quad S_2 \sim (\sigma_e + J\sigma_a)\chi^2_{I-1}.$$

In Table I, the following expectation is listed as a function of $\sigma_e/(\sigma_e+J\sigma_a)$ for estimators ϕ of σ_a:

$$\text{Risk} = \frac{J^2}{(\sigma_e+J\sigma_a)^2} E_{(\sigma_e,\sigma_a)}(\phi-\sigma_a)^2.$$

This form of the mean squared error was chosen so that the results could be directly compared with those in Klotz, Milton and Zacks [6]. In Table I, we compare the above risk of formal Bayes estimators ϕ_b defined by (2.10) for $a = -1$ and listed values of b with values for the following estimators given in Klotz, Milton and Zacks [6]:

(A.1) $$\frac{1}{J}\left(\frac{S_1}{I} - \frac{S_1}{I(J-1)}\right)^+$$

(A.2) $$\frac{1}{J}\left(\frac{S_2}{I+1} - \frac{S_1}{I(J-1)}\right)^+$$

$$\text{(A.3)} \qquad \frac{1}{J}\left(\frac{S_2}{I+2}-\frac{S_1}{I(J-1)+2}\right)^+$$

$$\text{(A.4)} \qquad \frac{1}{J}\left(\frac{S_2}{I+1}-\frac{S_1}{I(J-1)-2}\right)^+$$

(A.5) is the expected value of σ_a under the Tiao–Tan posterior using the Jeffreys' prior (see (2.6) or (2.7) on page 39 of Tiao–Tan [19]); where $(x)^+ = \max(x, 0)$. Note that (A.1) is the maximum likelihood estimator, (A.2) is suggested by Zacks [20], (A.3) is the σ_a component of the mode of the posterior distribution (2.6) in Tiao–Tan [19], and (A.4) is the mode of the posterior distribution (2.7) in Tiao–Tan [19].

TABLE 1

Risk of Estimators of σ_a: $\dfrac{J^2}{(\sigma_e+J\sigma_a)^2}E(\phi-\phi_a)^2$

$I = 4 \quad J = 2$

$\frac{\sigma_e}{\sigma_e+J\sigma_a}$	A.1	A.2	A.3	A.4	A.5	$b=-1$	$b=-\frac{1}{2}$
1.0	.2744	.1408	.1220	.0604	20.6496	.3764	.2340
.8182	.2588	.1379	.1107	.0807	17.2529	.2159	.1140
.6667	.2775	.1703	.1364	.1392	14.8606	.1483	.0775
.5385	.3101	.2170	.1787	.2133	13.1576	.1366	.0897
.4286	.3455	.2663	.2261	.2886	11.9439	.1582	.1306
.3333	.3776	.3111	.2720	.3561	11.0876	.1988	.1870
.2500	.4033	.3478	.3131	.4091	10.5013	.2489	.2502
.1765	.4216	.3744	.3481	.4422	10.1264	.3017	.3138
.1111	.4323	.3911	.3764	.4520	9.9263	.3515	.3722
.0526	.4369	.3987	.3988	.4363	9.8829	.3910	.4174
0.0	.4375	.4000	.4167	.4000	10.0000	.4000	.4167

$I = 4 \quad J = 4$

$\frac{\sigma_e}{\sigma_e+J\sigma_a}$	A.1	A.2	A.3	A.4	A.5	$b=-1$	$b=-\frac{1}{2}$	$b=0$
1.0	.2080	.0960	.0640	.0704	18.4016	.3824	.2400	.1616
.6923	.2452	.1477	.1155	.1354	13.5725	.1316	.0643	.0369
.5000	.3119	.2319	.2044	.2338	11.5275	.1219	.0906	.0875
.3684	.3625	.2978	.2761	.3085	10.5910	.1737	.1631	.1742
.2727	.3957	.3415	.3269	.3574	10.1593	.2360	.2377	.2588
.2000	.4157	.3692	.3610	.3858	9.9756	.2926	.3026	.3300
.1492	.4273	.3853	.3833	.4006	9.9165	.3380	.3535	.3859
.0968	.4335	.3941	.3976	.4064	9.9180	.3709	.3899	.4256
.0588	.4364	.3982	.4066	.4068	9.9446	.3917	.4122	.4494
.0270	.4373	.3997	.4126	.4039	9.9757	.4008	.4206	.4572
0.0	.4375	.4000	.4167	.4000	10.0000	.4000	.4167	.4490

(TABLE 1—*continued*)

$\frac{\sigma_e}{\sigma_e+J\sigma_a}$	$I=4$ $J=8$							
	A.1	A.2	A.3	A.4	A.5	$b=-1$	$b=-\frac{1}{2}$	$b=0$
1.0	.1856	.0832	.0448	.0704	17.5424	.3904	.2432	.1664
.5294	.2967	.2170	.1927	.2170	11.4048	.1063	.0753	.0709
.3333	.3744	.3156	.3044	.3211	10.2633	.1933	.1900	.2056
.2258	.4089	.3610	.3596	.3676	9.9963	.2757	.2830	.3077
.1579	.4246	.3820	.3865	.3883	9.9439	.3306	.3439	.3732
.1111	.4317	.3919	.4004	.3972	9.9473	.3644	.3805	.4128
.0769	.4350	.3967	.4078	.4007	9.9626	.3841	.4017	.4353
.0508	.4367	.3988	.4118	.4017	9.9773	.3944	.4126	.4466
.0303	.4372	.3996	.4143	.4014	9.9882	.3992	.4171	.4509
.0137	.4375	.3999	.4156	.4007	9.9954	.4004	.4178	.4510
0.0	.4375	.4000	.4167	.4000	10.0000	.4000	.4167	.4490

$\frac{\sigma_e}{\sigma_e+J\sigma_a}$	$I=4$ $J=10$							
	A.1	A.2	A.3	A.4	A.5	$b=-1$	$b=-\frac{1}{2}$	$b=0$
1.0	.1800	.0800	.0400	.0700	17.3900	.3900	.2500	.1700
.4737	.3186	.2438	.2244	.2465	10.9418	.1191	.0970	.0997
.2857	.3903	.3367	.3329	.3418	10.1046	.2283	.2296	.2500
.1892	.4178	.3733	.3769	.3784	9.9613	.3068	.3170	.3440
.1304	.4291	.3885	.3965	.3927	9.9490	.3520	.3667	.3970
.0909	.4340	.3950	.4060	.3983	9.9617	.3769	.3937	.4261
.0625	.4360	.3979	.4109	.4006	9.9753	.3903	.4077	.4412
.0411	.4370	.3993	.4136	.4010	9.9859	.3969	.4145	.4481
.0244	.4374	.3998	.4151	.4010	9.9929	.3996	.4172	.4503
.0110	.4375	.4000	.4160	.4004	9.9972	.4003	.4173	.4502
0.0	.4375	.4000	.4167	.4000	10.0000	.4000	.4167	.4490

$\frac{\sigma_e}{\sigma_e+J\sigma_a}$	$I=10$ $J=4$							
	A.1	A.2	A.3	A.4	A.5	$b=-1$	$b=-\frac{1}{2}$	$b=0$
1.0	.1056	.0672	.0544	.0560	.7696	.1968	.1472	.1136
.6923	.1363	.1079	.0937	.1041	.4500	.0729	.0483	.0341
.5000	.1713	.1531	.1431	.1563	.4188	.0888	.0775	.0756
.3684	.1866	.1742	.1680	.1804	.4171	.1254	.1232	.1294
.2727	.1911	.1815	.1782	.1878	.4271	.1547	.1583	.1702
.2000	.1917	.1830	.1823	.1884	.4362	.1725	.1792	.1948
.1492	.1912	.1829	.1841	.1867	.4420	.1810	.1890	.2053
.0968	.1906	.1825	.1850	.1850	.4452	.1835	.1916	.2084
.0588	.1902	.1820	.1859	.1837	.4468	.1834	.1909	.2068
.0270	.1900	.1819	.1866	.1826	.4480	.1826	.1891	.2038
0.0	.1900	.1818	.1875	.1818	.4490	.1818	.1875	.2012

(TABLE 1—*continued*)

	$I = 10 \quad J = 10$						
$\frac{\sigma_e}{\sigma_e + J\sigma_a}$	A.1	A.2	A.3	A.4	A.5	$b = -1$	$b = -\frac{1}{2}$
1.0	.0900	.0600	.0400	.0500	.7400	.1900	.1400
.4737	.1717	.1551	.1496	.1579	.4183	.0997	.0886
.2857	.1888	.1786	.1798	.1811	.4337	.1543	.1569
.1892	.1899	.1819	.1855	.1833	.4434	.1753	.1797
.1304	.1905	.1819	.1862	.1834	.4466	.1810	.1876
.0909	.1902	.1818	.1868	.1828	.4479	.1821	.1891
.0625	.1902	.1818	.1870	.1824	.4482	.1824	.1884
.0411	.1901	.1818	.1871	.1822	.4485	.1822	.1880
.0244	.1901	.1818	.1872	.1820	.4487	.1820	.1878
.0110	.1900	.1818	.1874	.1819	.4489	.1819	.1877
0.0	.1900	.1818	.1875	.1818	.4490	.1818	.1875

In Table II we list the following expectation as a function of $\sigma_a/(\sigma_e+\sigma_a)$ for estimators ϕ of σ_e:

$$\text{Risk} = \frac{1}{\sigma_e^2} E(\phi - \sigma_e)^2. \tag{A.6}$$

Here, we compare the risk of formal Bayes estimators ϕ_b defined by (4.6) for $a = -1$ and listed values of b with values for the following estimators:

$$\min\left(\frac{S_1}{I(J-1)}, \frac{S_1+S_2}{IJ}\right) \tag{A.7}$$

$$\min\left(\frac{S_1}{I(J-1)+2}, \frac{S_1+S_2}{IJ+1}\right) \tag{A.8}$$

$$\min\left(\frac{S_1}{I(J-1)+2}, \frac{S_1+S_2}{IJ+4}\right) \tag{A.9}$$

$$\min\left(\frac{S_1}{I(J-1)+1}, \frac{S_1+S_2}{IJ+4}\right). \tag{A.10}$$

Note that (A.7) is the maximum likelihood estimator, (A.9) is the σ_e component of the mode of the distribution (2.6) in Tiao–Tan [19], and (A.10) is the mode of a posterior distribution considered by Stone and Springer [17].

TABLE 2

Mean Squared Errors of Estimates of σ_e: $\frac{1}{\sigma_e^2} E(\varphi-\sigma_e)^2$

$I = 4 \quad J = 2$

$\frac{\sigma_a}{\sigma_e+\sigma_a}$	A.7	A.8	A.9	A.10	$b = -3.0$	$b = -2.5$	$b = -2.0$	$b = -1.0$
0.0	.27350	.31433	.34070	.32058	.33333	.31363	.29638	.26900
0.1	.28327	.31481	.33176	.30862	.32880	.30730	.28837	.25847
0.2	.29761	.31612	.32490	.29961	.32484	.30183	.28161	.25047
0.3	.31639	.31802	.32018	.29408	.32153	.29744	.27657	.24650
0.4	.33928	.32034	.31759	.29252	.31898	.29444	.27397	.24914
0.5	.36572	.32289	.31706	.29531	.31734	.29329	.27494	.26293
0.6	.39490	.32551	.31843	.30268	.31682	.29470	.28130	.29656
0.7	.42561	.32806	.32141	.31451	.31776	.29976	.29624	.36834
0.8	.45602	.33037	.32556	.33005	.32061	.31034	.32590	.52379
0.9	.48317	.33226	.33011	.34730	.32600	.32971	.38387	.91512
1.0	.50000	.33333	.33333	.36000	.33333	.36000	.50000	3.00000

$I = 4 \quad J = 4$

$\frac{\sigma_a}{\sigma_e+\sigma_a}$	A.7	A.8	A.9	A.10	$b = -3.0$	$b = -2.5$	$b = -2.0$	$b = -1.0$
0.0	.12840	.13683	.15081	.14635	.14286	.13741	.13315	.12723
0.1	.13362	.13737	.14338	.13740	.13990	.13371	.12894	.12275
0.2	.13981	.13838	.13978	.13349	.13821	.13184	.12723	.12249
0.3	.14579	.13942	.13847	.13291	.13747	.13145	.12773	.12666
0.4	.15111	.14033	.13844	.13432	.13743	.13225	.13016	.13549
0.5	.15560	.14108	.13906	.13675	.13791	.13400	.13428	.14923
0.6	.15928	.14169	.13996	.13959	.13875	.13648	.13985	.16809
0.7	.16219	.14216	.14092	.14241	.13983	.13946	.14655	.19220
0.8	.16439	.14251	.14179	.14491	.14101	.14270	.15400	.22130
0.9	.16592	.14274	.14248	.14687	.14213	.14583	.16144	.25376
1.0	.16667	.14286	.14286	.14793	.14286	.14793	.16667	.28000

(TABLE 2—*continued*)

$\frac{\sigma_a}{\sigma_e+\sigma_a}$	I = 4 J = 8							
	A.7	A.8	A.9	A.10	$b=-3.0$	$b=-2.5$	$b=-2.0$	$b=-1.0$
0.0	.06314	.06503	.07003	.06901	.06667	.06527	.06422	.06283
0.1	.06561	.06542	.06633	.06480	.06539	.06381	.06272	.06166
0.2	.06760	.06584	.06560	.06437	.06515	.06379	.06315	.06366
0.3	.06893	.06613	.06564	.06493	.06529	.06433	.06436	.06720
0.4	.06981	.06632	.06586	.06565	.06555	.06504	.06581	.07134
0.5	.07040	.06645	.06609	.06629	.06583	.06574	.06723	.07550
0.6	.07080	.06654	.06629	.06681	.06609	.06637	.06852	.07937
0.7	.07108	.06659	.06644	.06721	.06631	.06690	.06961	.08277
0.8	.07126	.06663	.06656	.06749	.06648	.06732	.07048	.08557
0.9	.07138	.06666	.06663	.06769	.06660	.06762	.07111	.08760
1.0	.07143	.06667	.06667	.06778	.06667	.06778	.07143	.08876

$\frac{\sigma_a}{\sigma_e+\sigma_a}$	I = 4 J = 10							
	A.7	A.8	A.9	A.10	$b=-3.0$	$b=-2.5$	$b=-2.0$	$b=-1.0$
0.0	.05039	.05157	.05501	.05437	.05263	.05174	.05107	.05018
0.1	.05227	.05189	.05227	.05132	.05172	.05073	.05009	.04961
0.2	.05356	.05218	.05194	.05129	.05167	.05090	.05067	.05155
0.3	.05432	.05235	.05204	.05175	.05182	.05137	.05164	.05426
0.4	.05478	.05246	.05220	.05221	.05202	.05186	.05262	.05698
0.5	.05508	.05253	.05234	.05258	.05220	.05229	.05347	.05942
0.6	.05527	.05257	.05244	.05286	.05234	.05264	.05417	.06148
0.7	.05540	.05260	.05252	.05305	.05246	.05291	.05472	.06314
0.8	.05548	.05262	.05258	.05319	.05255	.05312	.05514	.06441
0.9	.05553	.05263	.05262	.05328	.05260	.05326	.05542	.06529
1.0	.05556	.05263	.05263	.05332	.05263	.05332	.05556	.06574

(TABLE 2—*continued*)

	$I = 10 \quad J = 4$							
$\frac{\sigma_a}{\sigma_e + \sigma_a}$	A.7	A.8	A.9	A.10	$b = -3.0$	$b = -2.5$	$b = -2.0$	$b = -1.0$
0.0	.05405	.05795	.06075	.05871	.06250	.06022	.05832	.05537
0.1	.05766	.05909	.05809	.05573	.05920	.05674	.05475	.05203
0.2	.06158	.06063	.05900	.05737	.05851	.05635	.05485	.05365
0.3	.06421	.06163	.06048	.05984	.05915	.05763	.05705	.05862
0.4	.06563	.06215	.06154	.06170	.06022	.05948	.06000	.06510
0.5	.06629	.06238	.06212	.06275	.06122	.06118	.06277	.07143
0.6	.06655	.06246	.06238	.06324	.06192	.06241	.06481	.07647
0.7	.06664	.06249	.06247	.06342	.06221	.06311	.06601	.07968
0.8	.06666	.06250	.06250	.06347	.06246	.06341	.06653	.08097
0.9	.06667	.06250	.06250	.06348	.06249	.06346	.06664	.08161
1.0	.06667	.06250	.06250	.06348	.06250	.06348	.06667	.08163

	$I = 10 \quad J = 10$							
$\frac{\sigma_a}{\sigma_e + \sigma_a}$	A.7	A.8	A.9	A.10	$b = -3.0$	$b = -2.5$	$b = -2.0$	$b = -1.0$
0.0	.02053	.02104	.02160	.02134	.02174	.02140	.02113	.02073
0.1	.02169	.02151	.02126	.02107	.02112	.02087	.02071	.02067
0.2	.02210	.02169	.02159	.02160	.02141	.02133	.02142	.02210
0.3	.02219	.02173	.02170	.02178	.02159	.02163	.02188	.02307
0.4	.02221	.02174	.02173	.02184	.02170	.02179	.02206	.02350
0.5	.02222	.02174	.02174	.02185	.02173	.02183	.02219	.02368
0.6	.02222	.02174	.02174	.02186	.02174	.02185	.02221	.02373
0.7	.02222	.02174	.02174	.02186	.02174	.02185	.02222	.02376
0.8	.02222	.02174	.02174	.02186	.02174	.02186	.02222	.02376
0.9	.02222	.02174	.02174	.02186	.02174	.02186	.02222	.02376
1.0	.02222	.02174	.02174	.02186	.02174	.02186	.02222	.02376

REFERENCES

[1] Box, G. E. P. and Tiao, G. C. (1967). Bayesian analysis of a three component hierarchial design model. *Biometrika* **54** 109–125.

[2] Brown, L. (1968). Inadmissibility of the usual estimators of scale parameters in problems with unknown scale and location parameters. *Ann. Math. Statist.* **39** 29–48.

[3] Ferguson, T. (1967). *Mathematical Statistics: A Decision Theoretic Approach.* Academic Press, New York.

[4] Hill, Bruce (1965). Inference about variance components in a one way model. *J. Amer. Statist. Assoc.* **60** 806–825.

[5] Kiefer, J. (1957). Invariance, minimax sequential estimation and continuous time processes. *Ann. Math. Statist.* **28** 573–601.

[6] Klotz, J., Milton, R. C., and Zacks, S. (1969). Mean square efficiency of estimators of variance components. *J. Amer. Statist. Assoc.* **64** 1383–1402.

[7] James, W. and Stein, C. (1960). Estimation with quadratic loss. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* Univ. of California Press, 361–379.

[8] LeCam, L. (1965). An extension of Wald's theory of statistical decision functions. *Ann. Math. Statist.* **26** 69–81.

[9] Pearson, K. (1934). *Tables of the Incomplete Beta-Function.* Cambridge University Press.

[10] Portnoy, S. (1969a). Formal Bayes estimation with application to a random effects analysis of variance model. Tech. Report No. 2, Department of Statistics, Stanford University.

[11] Portnoy, S. (1969b). Estimation of parameters of velocity distributions: II. Estimating gamma scale and shape parameters. Tech. Report No. 13, Department of Statistics, Stanford University.

[12] Scheffé, H. (1961). *The Analysis of Variance.* Wiley, New York.

[13] Stein, C. (1955). A necessary and sufficient condition for admissibility. *Ann. Math. Statist.* **26** 518–522.

[14] Stein, C. (1959). The admissibility of Pitman's estimator of a single location parameter. *Ann. Math. Statist.* **30** 970–979.

[15] Stein, C. (1965). Approximation of improper prior measures by prior probability measures. *Bernoulli, Bayes, Laplace Anniversary Volume.* Springer-Verlag, New York.

[16] Stein, C. (1965). Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. *Ann. Inst. Statist. Math.* **16** 155–160.

[17] Stone, M. and Springer, B. G. F. (1965). A paradox involving quasi prior distributions. *Biometrika* **54** 623–627.

[18] Thompson, W. A. Jr. (1965). The problem of negative estimates of variance components. *Ann. Math. Statist.* **33** 273–289.

[19] Tiao, G. C. and Tan, W. Y. (1965). Bayesian analysis of random effect models in analysis of variance, I. Posterior distribution of variance components. *Biometrika* **52** 37–53.

[20] Zacks, S. (1967). More efficient estimators of variance components. Technical Report No. 4, Department of Statistics, Kansas State University.

[21] Zacks, S. (1970). Bayes equivariant estimators of variance components. *Ann. Inst. of Statist. Math.* **22** 27–40.

[22] Zidek, J. (1969). A representation of Bayes invariant procedures in terms of Haar measure. *Ann. Inst. of Statist. Math.* **21** 291–308.

[23] Zidek, J. (1970). Sufficient conditions for the admissibility under squared error loss of formal Bayes estimators. *Ann. Math. Statist.* **41** 446–456.