The Annals of Statistics
1996, Vol. 24, No. 2, 825–840

SHRINKAGE ESTIMATION IN THE TWO-WAY MULTIVARIATE NORMAL MODEL[1]

BY LI SUN

Simon Fraser University

A two-way multivariate normal model is proposed and attention is focused on estimation of the mean values when the common variance of the observations is unknown. A class of empirical Bayes estimators is proposed and mean-squared errors are given. A lower bound on the mean-squared error is found and related to risk asymptotics. A James–Stein-type estimator is derived and compared with its competitor—a modal estimator that is obtained from a hierarchical prior for the unknown parameters.

1. Introduction. Consider the following two-way model:

$$y_{ijk} = \theta_{ij} + \varepsilon_{ijk}, \tag{1.1}$$

where $\varepsilon_{ijk} \sim_{\text{i.i.d.}} N(0, \tau^2)$ with unknown τ^2 for $k = 1, \dots, n$, $i = 1, \dots, r$, $j = 1, \dots, s$. Our interest lies in estimation of the treatment means θ_{ij}, $i = 1, \dots, r$, $j = 1, \dots, s$.

Following Lindley (1972), suppose the treatment means θ_{ij} in each cell (i, j) are composed of four factors, overall effect μ, row effect α, column effect β and interaction effect γ, that is,

$$\theta_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij}, \qquad i = 1, \dots, \tau, j = 1, \dots, s.$$

At this stage, suppose the overall mean μ is fixed, with α, β and γ being independent random effects distributed as follows:

$$\alpha_i \sim_{\text{i.i.d.}} N(0, \sigma_A^2), \qquad \beta_j \sim_{\text{i.i.d.}} N(0, \sigma_B^2), \qquad \gamma_{ij} \sim_{\text{i.i.d.}} N(0, \sigma_{AB}^2).$$

Consequently, the distributions of θ_{ij} are normal with common mean

$$E\left[\theta_{ij} \mid \mu, \sigma_{AB}^2, \sigma_A^2, \sigma_B^2\right] = \mu \tag{1.2}$$

Received December 1993; revised April 1995.
[1]Research supported by Natural Sciences and Engineering Research Council of Canada.
AMS 1991 *subject classifications*. Primary 62F15; secondary 62C10, 62F11, 62F12, 62J07.
Key words and phrases. Empirical Bayes estimates, hierarchical priors, James–Stein estimator, mean-squared error, modal estimator, noncentral chi-square distribution, shrinkage of estimates.

and covariances

$$\operatorname{cov}\left[\theta_{i_1 j_1}, \theta_{i_2 j_2} \mid \mu, \sigma_{AB}^2, \sigma_A^2, \sigma_B^2\right]$$

$$(1.3)\qquad = \begin{cases} \sigma_{AB}^2 + \sigma_A^2 + \sigma_B^2, & i_1 = i_2, j_1 = j_2, \\ \sigma_A^2, & i_1 = i_2, j_1 \neq j_2, \\ \sigma_B^2, & i_1 \neq i_2, j_1 = j_2, \\ 0, & i_1 \neq i_2, j_1 \neq j_2. \end{cases}$$

Denote

$$(1.4)\quad \boldsymbol{\theta} = (\theta_{11}, \dots, \theta_{1s}, \dots, \theta_{r1}, \dots, \theta_{rs})', \qquad \boldsymbol{\phi} = (\mu, \tau^2, \sigma_{AB}^2, \sigma_A^2, \sigma_B^2)'.$$

In this article, we will use the conventional dot notation such as $\bar{y}_{ij\cdot} = n^{-1}\sum_{k=1}^{n} y_{ijk}$, $\bar{y}_{\cdot j \cdot} = r^{-1}\sum_{i=1}^{r} \bar{y}_{ij\cdot} = (rn)^{-1}\sum_{i=1}^{r}\sum_{k=1}^{n} y_{ijk}$ and $\bar{\theta}_{i\cdot} = s^{-1}\sum_{j=1}^{s}\theta_{ij}$ and so forth. A two-way array will always be represented by a vector ordered as is $\boldsymbol{\theta}$ in (1.4).

Applying Lindley and Smith's (1972) approach to the above hierarchical structure, we find that the posterior distribution $(\boldsymbol{\theta} \mid \mathbf{y}, \boldsymbol{\phi})$ given $\boldsymbol{\phi}$ is multivariate normal with mean

$$E[\boldsymbol{\theta} \mid \mathbf{y}, \boldsymbol{\phi}] = \boldsymbol{\theta}^* = (\theta_{11}^*, \dots, \theta_{1s}^*, \dots, \theta_{r1}^*, \theta_{rs}^*)',$$

where

$$(1.5)\quad \begin{aligned} \theta_{ij}^* = E[\theta_{ij} \mid \mathbf{y}, \boldsymbol{\phi}] &= \left(1 - \frac{\tau^2}{\tau^2 + n\sigma_{AB}^2}\right)(\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot j\cdot} + \bar{y}_{\cdot\cdot\cdot}) \\ &\quad + \left(1 - \frac{\tau^2}{\tau^2 + n\sigma_{AB}^2 + ns\sigma_A^2}\right)(\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot\cdot\cdot}) \\ &\quad + \left(1 - \frac{\tau^2}{\tau^2 + n\sigma_{AB}^2 + nr\sigma_B^2}\right)(\bar{y}_{\cdot j\cdot} - \bar{y}_{\cdot\cdot\cdot}) \\ &\quad + \left(\frac{\tau^2}{\tau^2 + n\sigma_{AB}^2 + ns\sigma_A^2 + nr\sigma_B^2}\right)(\mu - \bar{y}_{\cdot\cdot\cdot}) + \bar{y}_{\cdot\cdot\cdot}. \end{aligned}$$

The variance matrix $\operatorname{var}[\boldsymbol{\theta} \mid \mathbf{y}, \boldsymbol{\phi}]$ is found to be $(\tau^{-2}I + C^{-1})^{-1}$, where I is the identity matrix and C is the covariance matrix determined by (1.3). The derivation can be found, for example, in Sun (1992). A similar discussion is given by Lindley (1972).

The posterior mean $E[\boldsymbol{\theta} \mid \mathbf{y}, \boldsymbol{\phi}]$ does not provide a valid estimator for $\boldsymbol{\theta}$ in the circumstances when the components of $\boldsymbol{\phi}$ are unknown. However, the empirical Bayes method developed by Efron and Morris (1973) can be used to find a class of empirical Bayes estimators by utilizing the marginal distribution of $(\mathbf{y}|\boldsymbol{\phi})$. Sun [22] shows that this marginal distribution has the following form:

$$
\begin{aligned}
p(\mathbf{y}|\boldsymbol{\phi}) \propto\ & (\tau^2)^{-(n-1)rs/2}\left(\tau^2+n\sigma_{AB}^2\right)^{-(r-1)(s-1)/2}\left(\tau^2+n\sigma_{AB}^2+ns\sigma_A^2\right)^{-(r-1)/2} \\
& \times\left(\tau^2+n\sigma_{AB}^2+nr\sigma_B^2\right)^{-(s-1)/2}\left(\tau^2+n\sigma_{AB}^2+ns\sigma_A^2+nr\sigma_B^2\right)^{-1/2} \\
& \times\exp\left\{-\frac{1}{2}\left[\frac{S^2}{\tau^2}+\frac{T_{AB}^2}{\tau^2+n\sigma_{AB}^2}+\frac{T_A^2}{\tau^2+n\sigma_{AB}^2+ns\sigma_A^2}\right.\right. \\
& \qquad\qquad\left.\left.+\frac{T_B^2}{\tau^2+n\sigma_{AB}^2+nr\sigma_B^2}\right]\right\} \\
& \times\exp\left\{-\frac{1}{2}\left[\frac{nrs(\bar{y}_{\cdots}-\mu)^2}{\tau^2+n\sigma_{AB}^2+ns\sigma_A^2+nr\sigma_B^2}\right]\right\},
\end{aligned} \tag{1.6}
$$

where

$$
\begin{aligned}
S^2 &= \sum_{i=1}^{r}\sum_{j=1}^{s}\sum_{k=1}^{n}\left(y_{ijk}-\bar{y}_{ij\cdot}\right)^2, \\
T_{AB}^2 &= n\sum_{i=1}^{r}\sum_{j=1}^{s}\left(\bar{y}_{ij\cdot}-\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdot j\cdot}+\bar{y}_{\cdots}\right)^2, \\
T_A^2 &= ns\sum_{i=1}^{r}\left(\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdots}\right)^2, \\
T_B^2 &= nr\sum_{j=1}^{s}\left(\bar{y}_{\cdot j\cdot}-\bar{y}_{\cdots}\right)^2.
\end{aligned} \tag{1.7}
$$

It is obvious that $\bar{y}_{\cdots}$, S^2, T_{AB}^2, T_A^2 and T_B^2 are sufficient and complete statistics for this distributional family indexed by $\boldsymbol{\phi}$.

When $\boldsymbol{\phi}$ is unknown, a typical empirical Bayes estimator, $\boldsymbol{\delta}$, will have $[(i-1)s+j]$th component

$$
\begin{aligned}
\delta_{ij} =\ & (1-\rho_{AB})\left(\bar{y}_{ij\cdot}-\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdot j\cdot}+\bar{y}_{\cdots}\right)+(1-\rho_A)\left(\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdots}\right) \\
& +(1-\rho_B)\left(\bar{y}_{\cdot j\cdot}-\bar{y}_{\cdots}\right)+\bar{y}_{\cdots}
\end{aligned} \tag{1.8}
$$

for $i=1,\dots,r$, $j=1,\dots,s$, where ρ_{AB}, ρ_A and ρ_B are some functions of $(S^2, T_{AB}^2, T_A^2, T_B^2)$, to be written generically as $\rho_* = \rho_*(S^2, T_{AB}^2, T_A^2, T_B^2)$. This type of empirical Bayes estimator is obviously suggested by the form of $E[\theta_{ij}|\mathbf{y},\boldsymbol{\phi}]$ in (1.5).

As a special case, when $\rho_{AB}=\rho_A=\rho_B=0$, $\delta_{ij}=\bar{y}_{ij}$ yields the maximum likelihood estimate (MLE) of θ_{ij}, $i=1,\dots,r$, $j=1,\dots,s$. Another special case is when $\rho_{AB}=\rho_A=\rho_B=1$, corresponding to a common estimate—the overall sample average $\bar{y}_{\cdots}$ for each component of $\boldsymbol{\theta}$. Other particular forms of the estimators can be obtained by letting $\rho_{AB}=0$, $\rho_A=\rho_B=1$, $\rho_{AB}=1$,

$\rho_A = 0$, $\rho_B = 1$ and so on. In any case, an estimator determined by (1.8) is referred to as a *shrinkage estimator*.

Estimation problems for multivariate normal means have long been of interest in the literature. See, for example, Stein (1956, 1966, 1981), James and Stein (1961), Box and Tiao (1968), Baranchik (1970), Strawderman (1971), Efron and Morris (1973), Leonard (1976), Berger (1980, 1982), Casella and Hwang (1982), Berger and Wolpert (1983), Morris (1983) and Green and Strawderman (1991).

The two-way model with interactions is a classic model. Some interesting discussions are given by Stein (1966), Giri (1970), Lindley (1972) and Smith (1973). This paper shows how to implement the standard empirical Bayes technique and proves minimaxity of the resultant empirical Bayes estimators. Some useful properties of the risk function are established, and the dominance results are generalized to other interesting estimators, such as a positive part estimator and a hierarchical Bayes modal estimator.

In Section 2, some useful stochastic properties of the statistics in (1.7) are given. We discuss the mean-squared error (MSE) and provide a class of improved shrinkage estimators in Section 3. Asymptotic results are presented in Section 4. The discussion is expanded in Section 5 to the Bayes modal estimator which arises by assuming a hierarchical prior for the hyperparameter $\boldsymbol{\phi}$.

2. Preliminary results. In this section, we will provide needed distributional properties of the statistics $(S^2, T_{AB}^2, T_A^2, T_B^2)$.

We write $X =_d Y$ if X has the same distribution as Y. To facilitate the discussion, we introduce the notation

$$
\begin{aligned}
h_1 &= \sum_{i=1}^{r}\sum_{j=1}^{s}\left(\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot j\cdot} + \bar{y}_{\cdot\cdot\cdot}\right)\left(\theta_{ij} - \bar{\theta}_{i\cdot} - \bar{\theta}_{\cdot j} + \bar{\theta}_{\cdot\cdot}\right),\\
h_2 &= s\sum_{i=1}^{r}\left(\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot\cdot\cdot}\right)\left(\bar{\theta}_{i\cdot} - \bar{\theta}_{\cdot\cdot}\right),\\
h_3 &= r\sum_{j=1}^{s}\left(\bar{y}_{\cdot j\cdot} - \bar{y}_{\cdot\cdot\cdot}\right)\left(\bar{\theta}_{\cdot j} - \bar{\theta}_{\cdot\cdot}\right)
\end{aligned}
\tag{2.1}
$$

and

$$
\begin{aligned}
v_1^2 &= \sum_{i=1}^{r}\sum_{j=1}^{s}\left(\theta_{ij} - \bar{\theta}_{i\cdot} - \bar{\theta}_{\cdot j} + \bar{\theta}_{\cdot\cdot}\right)^2,\\
v_2^2 &= s\sum_{i=1}^{r}\left(\bar{\theta}_{i\cdot} - \bar{\theta}_{\cdot\cdot}\right)^2,\\
v_3^2 &= r\sum_{j=1}^{s}\left(\bar{\theta}_{\cdot j} - \bar{\theta}_{\cdot\cdot}\right)^2.
\end{aligned}
\tag{2.2}
$$

LEMMA 1. *Let* $(S^2, T_{AB}^2, T_A^2, T_B^2)$ *be given in* (1.7). *We have*:

(i) $(T_{AB}^2/n,\ h_1) =_d (x_1^2 + z_1^2,\ v_1 x_1)$, *where* $x_1 \sim N(v_1, \tau^2/n)$ *and* $z_1^2 \sim (\tau^2/n)\chi^2_{(r-1)(s-1)-1}$,

(ii) $(T_A^2/n, h_2) =_d (x_2^2 + z_2^2, v_2 x_2)$, *where* $x_2 \sim N(v_2, \tau^2/n)$ *and* $z_2^2 \sim (\tau^2/n)\chi^2_{r-2}$,

(iii) $(T_B^2/n, h_3) =_d (x_3^2 + z_3^2, v_3 x_3)$, *where* $x_3 \sim N(v_3, \tau^2/n)$ *and* $z_3^2 \sim (\tau^2/n)\chi^2_{s-2}$.

The distributions of $(x_1, x_2, x_3, z_1, z_2, z_3)$ *are independent.*

PROOF. The proof is straightforward; see Stein (1966), for example. □

LEMMA 2. *Suppose* $x \sim N(\mu, \sigma^2)$ *and* $h(x)$ *satisfies* $E[|h(x)|] < \infty$ *and* $E[|xh(x)|] < \infty$. *We have*

$$E[xh(x)] = \mu E[h(x)] + \sigma^2\, \partial E[h(x)]/\partial\mu.$$

PROOF. Since $E[|h(x)|]$ and $E[|xh(x)|]$ are finite, $\partial E[h(x)]/\partial\mu = E[(x-\mu)h(x)/\sigma^2]$, which leads to the result. □

THEOREM 1. *Suppose* $K_l \sim \text{Poisson}(\lambda_l)$ *are independent for* $l = 1, 2, 3$, *where* $\lambda_l = nv_l^2/2\tau^2$ *and the* v_l *are given in* (2.2). *For any real function* g *of* $(S^2, T_{AB}^2, T_A^2, T_B^2)$ *and* h_l *in* (2.1), *we have, given* $\boldsymbol{\theta}$ *and* τ^2,

$$\begin{aligned} &E\left[g\left(S^2, T_{AB}^2, T_A^2, T_B^2\right)\right] \\ &\qquad = E\left[g\left(\tau^2\chi_N^2, \tau^2\chi_{L_1}^2, \tau^2\chi_{L_2}^2, \tau^2\chi_{L_3}^2\right)\right], \\ &E\left[g\left(S^2, T_{AB}^2, T_A^2, T_B^2\right)h_l\right] \\ &\qquad = (2\tau^2/n)E\left[K_l g\left(\tau^2\chi_N^2, \tau^2\chi_{L_1}^2, \tau^2\chi_{L_2}^2, \tau^2\chi_{L_3}^2\right)\right], \end{aligned}$$

where the four chi-square random variables are independent when K_1, K_2, K_3 *are fixed, and where* $L_1 = 2K_1 + (r-1)(s-1)$, $L_2 = 2K_2 + r - 1$ *and* $L_3 = 2K_3 + s - 1$.

PROOF. Recalling the definition of the noncentral chi-squared distribution, if $a \sim \chi_s^2(\lambda)$, then, for any real function $f(\cdot)$,

$$E[f(a)] = E\left[f(\chi^2_{2K+s})\right] = \sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!}E\left[f(\chi^2_{2k+s})\right], \tag{2.3}$$

where $K \sim \text{Poisson}(\lambda)$, provided that the summation is convergent.

It follows from (1.6) and (1.7) that $S^2, T_{AB}^2, T_A^2, T_B^2$ are independent for given $\boldsymbol{\theta}, \tau^2$ and it is well known that $S^2 =_d \tau^2\chi_N^2$. Note that Lemma 1 implies

$$\begin{aligned} T_{AB}^2 &=_d nx_1^2 + \tau^2\chi^2_{(r-1)(s-1)-1} =_d \tau^2\chi^2_{(r-1)(s-1)}(\lambda_1), \\ T_A^2 &=_d nx_2^2 + \tau^2\chi^2_{r-2} =_d \tau^2\chi^2_{(r-1)}(\lambda_2), \\ T_B^2 &=_d nx_3^2 + \tau^2\chi^2_{s-2} =_d \tau^2\chi^2_{(s-1)}(\lambda_3), \end{aligned} \tag{2.4}$$

where $x_l \sim N(v_l, \tau^2/n)$ and $\lambda_l = nv_l^2/2\tau^2$, $l = 1, 2, 3$. The independence of $S^2, T_{AB}^2, T_A^2, T_B^2$ implies the independence of the four chi-square variables. It follows from (2.3) that

$$E\left[g\left(S^2, T_{AB}^2, T_A^2, T_B^2\right)\right] = E\left[g\left(\tau^2\chi_N^2, \tau^2\chi_{L_1}^2, \tau^2\chi_{L_2}^2, \tau^2\chi_{L_3}^2\right)\right].$$

To prove the second part of the theorem, we take $l = 1$ as an example. Note that (S^2, T_A^2, T_B^2) and (T_{AB}^2, h_1) are independent. Defining $\tilde{E}[\cdot \mid S^2, T_A^2, T_B^2]$, we have

$$E\left[g\left(S^2, T_{AB}^2, T_A^2, T_{AB}^2\right)h_l | S^2, T_A^2, T_B^2\right] = \tilde{E}\left[f\left(T_{AB}^2\right)h_1\right], \tag{2.5}$$

where $f(T_{AB}^2) = g(S^2, T_{AB}^2, T_A^2, T_B^2)$ when (S^2, T_A^2, T_B^2) are fixed. From Lemma 1,

$$\tilde{E}\left[f\left(T_{AB}^2\right)h_1\right] = v_1\tilde{E}\left[x_1 f\left(nx_1^2 + \tau^2 z_1^2\right)\right], \tag{2.6}$$

where $x_1 \sim N(v_1, \tau^2/n)$ and $z_1^2 \sim \chi^2_{(r-1)(s-1)-1}$, independently. It follows from Lemma 2 that

$$\begin{aligned} &v_1\tilde{E}\left[x_1 f\left(nx_1^2 + \tau^2 z_1^2\right)\right] \\ &\quad = v_1^2\tilde{E}\left[f\left(nx_1^2 + \tau^2 z_1^2\right)\right] + v_1\frac{\tau^2}{n}\frac{\partial}{\partial v_1}\tilde{E}\left[f\left(nx_1^2 + \tau^2 z_1^2\right)\right]. \end{aligned} \tag{2.7}$$

Now the last term above becomes

$$\begin{aligned} &\frac{v_1\tau^2}{n}\frac{\partial}{\partial v_1}\tilde{E}\left[f\left(\tau^2\chi_{L_1}^2\right)\right] \\ &\quad = \frac{v_1\tau^2}{n}\frac{\partial}{\partial v_1}\sum_{k_1=0}^{\infty} e^{-\lambda_1}\frac{\lambda_1^{k_1}}{k_1!}\tilde{E}\left[f\left(\tau^2\chi_{L_1}^2\right)\right] \\ &\quad = \frac{v_1\tau^2}{n}\sum_{k_1=1}^{\infty} e^{-\lambda_1}\left(\frac{k_1\lambda_1^{k_1-1}}{k_1!} - \frac{\lambda_1^{k_1}}{k_1!}\right)\frac{nv_1}{\tau^2}\tilde{E}\left[f\left(\tau^2\chi_{L_1}^2\right)\right] \\ &\quad = \frac{\tau^2}{n}\tilde{E}\left[2K_1 f\left(\tau^2\chi_{L_1}^2\right)\right] - v_1^2\tilde{E}\left[f\left(\tau^2\chi_{L_1}^2\right)\right]. \end{aligned}$$

From (2.5)–(2.7),

$$E\left[g\left(S^2, T_{AB}^2, T_A^2, T_{AB}^2\right)h_l | S^2, T_A^2, T_B^2\right] = \frac{\tau^2}{n}\tilde{E}\left[2K_1 f\left(\tau^2\chi_{L_1}^2\right)\right].$$

Taking expectation with respect to S^2, T_A^2, T_B^2 on the both sides leads to

$$\begin{aligned} &E\left[g\left(S^2, T_{AB}^2, T_A^2, T_{AB}^2\right)h_l\right] \\ &\quad = (2\tau^2/n)E\left[K_1 g\left(\tau^2\chi_N^2, \tau^2\chi_{L_1}^2, \tau^2\chi_{L_2}^2, \tau^2\chi_{L_3}^2\right)\right], \end{aligned}$$

which is the result of the second part of the theorem when $l = 1$. For $l = 2, 3$, the proofs can be similarly shown. □

3. MSE and improved shrinkage estimators. For an estimator $\boldsymbol{\delta}$, the mean-squared error (MSE) is defined as a function of $\boldsymbol{\theta}$ as

$$\text{MSE}(\boldsymbol{\delta}, \boldsymbol{\theta}) = E\left[\|\boldsymbol{\delta} - \boldsymbol{\theta}\|^2\right] = E\left[\sum_{i=1}^{r}\sum_{j=1}^{s}\left(\delta_{ij} - \theta_{ij}\right)^2\right], \tag{3.1}$$

where the expectation is taken over the distribution of $(\mathbf{y}|\boldsymbol{\theta}, \tau^2)$.

For estimators which have the form of (1.8), we denote

$$\rho_*(\underset{\sim}{K}) = \rho_*\left(\tau^2\chi_N^2, \tau^2\chi_{L_1}^2, \tau^2\chi_{L_2}^2, \tau^2\chi_{L_3}^2\right) \tag{3.2}$$

for $*$ covering AB, A and B.

THEOREM 2. *Assuming K_1, K_2, K_3 and L_1, L_2, L_3 given in Theorem 1, the mean-squared error of $\boldsymbol{\delta}$ can be written as*

$$\begin{aligned}
&\text{MSE}(\boldsymbol{\delta}, \boldsymbol{\theta}) \\
&\quad = \frac{\tau^2}{n}E\Big[\left(\rho_{AB}(\underset{\sim}{K})\right)^2\chi_{L_1}^2 - 2\rho_{AB}(\underset{\sim}{K}) \\
&\qquad\qquad \times\left(\chi_{L_1}^2 - 2K_1\right)\Big] \\
&\qquad + \frac{\tau^2}{n}E\left[\left(\rho_A(\underset{\sim}{K})\right)^2\chi_{L_2}^2 - 2\rho_A(\underset{\sim}{K})\left(\chi_{L_2}^2 - 2K_2\right)\right] \\
&\qquad + \frac{\tau^2}{n}E\left[\left(\rho_B(\underset{\sim}{K})\right)^2\chi_{L_3}^2 - 2\rho_B(\underset{\sim}{K})\left(\chi_{L_3}^2 - 2K_3\right)\right] + \frac{rs\tau^2}{n}.
\end{aligned} \tag{3.3}$$

Moreover, if $r > 3$, $s > 3$, then all MSE functions are bounded from below by

$$B_{rs} = (\tau^2/n)(rs - 6 - U), \tag{3.4}$$

where

$$\begin{aligned}
U = {} & E\left[\frac{((r-1)(s-1)-2)^2}{2K_1 + (r-1)(s-1) - 2}\right] \\
& + E\left[\frac{(r-3)^2}{2K_2 + r - 3}\right] + E\left[\frac{(s-3)^2}{2K_3 + s - 3}\right].
\end{aligned} \tag{3.5}$$

PROOF. Since $\theta_{ij} = (\theta_{ij} - \bar{\theta}_{i\cdot} - \bar{\theta}_{\cdot j} + \bar{\theta}_{\cdot\cdot}) + (\bar{\theta}_{i\cdot} - \bar{\theta}_{\cdot\cdot}) + (\bar{\theta}_{\cdot j} - \bar{\theta}_{\cdot\cdot}) + \theta_{\cdot\cdot}$,

$$\mathrm{MSE}(\boldsymbol{\delta}, \boldsymbol{\theta}) = E\left[\sum_{i=1}^{r}\sum_{j=1}^{s}(\delta_{ij} - \theta_{ij})^2\right] = E_1 + E_2 + E_2 + E_4,$$

where

$$E_1 = E\left[\sum_{i=1}^{r}\sum_{j=1}^{s}\left[(1-\rho_{AB})\left(\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot j\cdot} + \bar{y}_{\cdot\cdot\cdot}\right) - \left(\theta_{ij} - \bar{\theta}_{i\cdot} - \bar{\theta}_{\cdot j} + \bar{\theta}_{\cdot\cdot}\right)\right]^2\right],$$

$$E_2 = E\left[s\sum_{i=1}^{r}\left[(1-\rho_A)(\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot\cdot\cdot}) - \left(\bar{\theta}_{i\cdot} - \bar{\theta}_{\cdot\cdot}\right)\right]^2\right],$$

$$E_3 = E\left[r\sum_{j=1}^{s}\left[(1-\rho_B)(\bar{y}_{\cdot j\cdot} - \bar{y}_{\cdot\cdot\cdot}) - \left(\bar{\theta}_{\cdot j} - \bar{\theta}_{\cdot\cdot}\right)\right]^2\right],$$

$$E_4 = E\left[rs(\bar{y}_{\cdot\cdot\cdot} - \theta_{\cdot\cdot})^2\right] = \tau^2/n.$$

To evaluate E_1, we have

$$\begin{aligned} E_1 &= E\left[(\rho_{AB}^2 - 2\rho_{AB})T_{AB}^2/n\right] \\ &\quad + 2E\left[\rho_{AB}\sum_{i=1}^{r}\sum_{j=1}^{s}\left(\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot j\cdot} + \bar{y}_{\cdot\cdot\cdot}\right)\left(\theta_{ij} - \bar{\theta}_{i\cdot} - \bar{\theta}_{\cdot j} + \bar{\theta}_{\cdot\cdot}\right)\right] \\ &\quad + (r-1)(s-1)\tau^2/n, \end{aligned}$$

since $E[T_{AB}^2/n] = v_1^2 + (r-1)(s-1)\tau^2/n$ from (2.4). Now using the results of Theorem 1 leads to

$$\begin{aligned} E_1 &= \frac{\tau^2}{n}E\left[\left(\rho_{AB}(\underset{\sim}{K})\right)^2\chi^2_{L_1} - 2\rho_{AB}(\underset{\sim}{K})\left(\chi^2_{L_1} - 2K_1\right)\right] \\ &\quad + (r-1)(s-1)\frac{\tau^2}{n}. \end{aligned}$$

Similarly,

$$E_2 = \frac{\tau^2}{n}E\left[\left(\rho_A(\underset{\sim}{K})\right)^2\chi^2_{L_2} - 2\rho_A(\underset{\sim}{K})\left(\chi^2_{L_2} - 2K_2\right)\right] + (r-1)\frac{\tau^2}{n}$$

and

$$E_3 = \frac{\tau^2}{n}E\left[\left(\rho_B(\underset{\sim}{K})\right)^2\chi^2_{L_3} - 2\rho_B(\underset{\sim}{K})\left(\chi^2_{L_3} - 2K_3\right)\right] + (s-1)\frac{\tau^2}{n}.$$

Summing up E_1, E_2, E_3 and $E_4 = \tau^2/n$, we obtain the MSE as shown in (3.3).

By completing the quadratic forms of $\rho_{AB}(\underset{\sim}{K})$, $\rho_A(\underset{\sim}{K})$ and $\rho_B(\underset{\sim}{K})$ from (3.3), respectively, we have

$$\begin{aligned}\operatorname{MSE}(\boldsymbol{\delta},\boldsymbol{\theta}) &= \frac{\tau^2}{n}\left\{E\left[\chi^2_{L_1}\left(\rho_{AB}(\underset{\sim}{K}) - 1 + \frac{2K_1}{\chi^2_{L_1}}\right)^2\right]\right. \\ &\qquad + E\left[\chi^2_{L_2}\left(\rho_A(\underset{\sim}{K}) - 1 + \frac{2K_2}{\chi^2_{L_2}}\right)^2\right] \\ &\qquad \left. + E\left[\chi^2_{L_3}\left(\rho_B(\underset{\sim}{K}) - 1 + \frac{2K_3}{\chi^2_{L_3}}\right)^2\right]\right\} + B_{rs},\end{aligned} \tag{3.6}$$

where

$$\begin{aligned}B_{rs} = \frac{rs\tau^2}{n} - \frac{\tau^2}{n}&\left\{E\left[\frac{\left(\chi^2_{L_1} - 2K_1\right)^2}{\chi^2_{L_1}}\right]\right. \\ &\quad + E\left[\frac{\left(\chi^2_{L_2} - 2K_2\right)^2}{\chi^2_{L_2}}\right] \\ &\quad \left. + E\left[\frac{\left(\chi^2_{L_3} - 2K_3\right)^2}{\chi^2_{L_3}}\right]\right\}.\end{aligned}$$

By noting that, for any constants $k \geq 0$ and $a > 2$,

$$E\left[\frac{\left(\chi^2_{2k+a} - 2k\right)^2}{\chi^2_{2k+a}}\right] = 2k + a - 4k + \frac{4k^2}{2k + a - 2} = 2 + \frac{(a-2)^2}{2k + a - 2},$$

we have

$$B_{rs} = \frac{\tau^2}{n}(rs - 6 - U),$$

where U is given in (3.5). This completes the proof. □

Let us consider a subset of the estimators of (1.8) such that

$$\rho_{AB} = \frac{c_1 S^2}{T^2_{AB}}, \qquad \rho_A = \frac{c_2 S^2}{T^2_A}, \qquad \rho_B = \frac{c_3 S^2}{T^2_B}, \tag{3.7}$$

where c_1, c_2, c_3 are constants. Denote such estimators by $\boldsymbol{\delta}^c$. Employing (3.3), the MSE of $\boldsymbol{\delta}^c$ is given by

$$\begin{aligned} \mathrm{MSE}(\boldsymbol{\delta}^c, \boldsymbol{\theta}) &= \frac{\tau^2}{n} E\left[\frac{c_1^2 N(N+2) - 2c_1 N[(r-1)(s-1)-2]}{L_1}\right] \\ &\quad + \frac{\tau^2}{n} E\left[\frac{c_2^2 N(N+2) - 2c_2 N(r-3)}{L_2 - 2}\right] \\ &\quad + \frac{\tau^2}{n} E\left[\frac{c_3^2 N(N+2) - 2c_3 N(s-3)}{L_3 - 2}\right] + rs\frac{\tau^2}{n}, \end{aligned} \tag{3.8}$$

where $N = (n-1)rs$ and K_1, K_2, K_3 and L_1, L_2, L_3 given in Theorem 1. For any choices

$$\begin{aligned} &0 < c_1 < 2\left(\frac{(r-1)(s-1)-2}{N+2}\right), \qquad 0 < c_2 < 2\left(\frac{r-3}{N+2}\right), \\ &\qquad\qquad 0 < c_3 < 2\left(\frac{s-3}{N+2}\right), \end{aligned} \tag{3.9}$$

the $\mathrm{MSE}(\boldsymbol{\delta}^c, \boldsymbol{\theta})$ is smaller than $rs\tau^2/n$, which is the MSE of the maximum likelihood estimates ($\bar{y}_{ij\cdot}$, $i = 1, \dots, r$, $j = 1, \dots, s$), for any $\boldsymbol{\theta}$ and $\tau^2 > 0$. Corresponding to such choices, therefore, $\boldsymbol{\delta}^c$ is minimax since the MLE is minimax under the squared loss.

Note also that the MSE in (3.8) is minimized at

$$c_1 = \frac{(r-1)(s-1)-2}{N+2}, \qquad c_2 = \frac{r-3}{N+2}, \qquad c_3 = \frac{s-3}{N+2}. \tag{3.10}$$

This yields the so-called James–Stein estimator for the two-way multivariate normal means:

$$\begin{aligned} \boldsymbol{\delta}^{\mathrm{JS}} &= \left(1 - \frac{(r-1)(s-1)-2}{N+2}\frac{S^2}{T_{AB}^2}\right)\left(\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot j\cdot} + \bar{y}_{\cdot\cdot\cdot}\right) \\ &\quad + \left(1 - \frac{r-3}{N+2}\frac{S^2}{T_A^2}\right)\left(\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot\cdot\cdot}\right) \\ &\quad + \left(1 - \frac{s-3}{N+2}\frac{S^2}{T_B^2}\right)\left(\bar{y}_{\cdot j\cdot} - \bar{y}_{\cdot\cdot\cdot}\right) + \bar{y}_{\cdot\cdot\cdot} \end{aligned} \tag{3.11}$$

and

$$\mathrm{MSE}(\boldsymbol{\delta}^{\mathrm{JS}}, \boldsymbol{\theta}) = \frac{\tau^2}{n}\left(rs - \frac{N}{N+2}U\right), \tag{3.12}$$

where U is given in (3.5) and $N = (n-1)rs$.

A class of dominating estimators of the form (1.8) is given in the next theorem. Define

$$\rho_{AB} = C_{AB} S^2 / T_{AB}^2, \qquad \rho_A = C_A S^2 T_A^2, \qquad \rho_B = C_B S^2 T_B^2, \tag{3.13}$$

where $C_* = C_*(S^2, T_{AB}^2, T_A^2, T_B^2)$ for $* = AB, A, B$.

THEOREM 3. *For an estimator* $\boldsymbol{\delta}$ *determined by* (1.8) *and* (3.13), *if* C_* $(* = AB, A, B)$ *satisfy the following*:

(a) C_{AB} *is nonincreasing in* S^2 *and nondecreasing in* T_{AB}^2 *and*

$$0 \le C_{AB} \le 2\left(\frac{(r-1)(s-1)-2}{N+2}\right);$$

(b) C_A *is nonincreasing in* S^2 *and nondecreasing in* T_A^2 *and*

$$0 \le C_A \le 2\left(\frac{r-3}{N+2}\right);$$

(c) C_B *is nonincreasing in* S^2 *and nondecreasing in* T_B^2 *and*

$$0 \le C_B \le 2\left(\frac{s-3}{N+2}\right),$$

then $\mathrm{MSE}(\boldsymbol{\delta}, \boldsymbol{\theta}) \le rs\tau^2/n$, *provided that* $r \ge 3$, $s \ge 3$.

An example of the dominating estimators is the positive version of the James–Stein estimator:

$$\begin{aligned} \boldsymbol{\delta}^+ = {} & \left(1 - \frac{(r-1)(s-1)-2}{N+2}\frac{S^2}{T_{AB}^2}\right)_+ \left(\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot j\cdot} + \bar{y}_{\cdots}\right) \\ & + \left(1 - \frac{r-3}{N+2}\frac{S^2}{T_A^2}\right)_+ \left(\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdots}\right) \\ & + \left(1 - \frac{s-3}{N+2}\frac{S^2}{T_B^2}\right)_+ \left(\bar{y}_{\cdot j\cdot} - \bar{y}_{\cdots}\right) + \bar{y}_{\cdots}, \end{aligned} \tag{3.14}$$

where $(a)_+ = \max\{0, a\}$. It can also be shown that $\mathrm{MSE}(\boldsymbol{\delta}^+, \boldsymbol{\theta}) < \mathrm{MSE}(\boldsymbol{\delta}^{\mathrm{JS}}, \boldsymbol{\theta})$; that is, the James–Stein estimator is dominated by its positive version. See also Baranchik [1].

4. The MSE as $r, s \to \infty$. It is obvious that $\mathrm{MSE}(\boldsymbol{\delta}, \boldsymbol{\theta})$ of Theorem 2 goes to 0 when $n \to \infty$ while (r, s) and $\boldsymbol{\theta}, \tau^2$ are fixed, provided that ρ_{AB}, ρ_A and ρ_B are bounded in terms of n. Of more interest are the asymptotic MSE values when $r, s \to \infty$ and n is fixed. To investigate such asymptotic properties of the MSE, we assume

$$Q_{AB} = \lim_{r,s\to\infty} \frac{v_1^2}{rs} = \lim_{r,s\to\infty} \frac{1}{rs} \sum_{i=1}^{r} \sum_{j=1}^{s} \left(\theta_{ij} - \bar{\theta}_{i\cdot} - \bar{\theta}_{\cdot j} + \bar{\theta}_{\cdot\cdot}\right)^2,$$

$$Q_A = \lim_{r,s\to\infty} \frac{v_2^2}{rs} = \lim_{r,s\to\infty} \frac{1}{r} \sum_{i=1}^{r} \left(\bar{\theta}_{i\cdot} - \bar{\theta}_{\cdot\cdot}\right)^2,$$

$$Q_B = \lim_{r,s\to\infty} \frac{v_3^2}{rs} = \lim_{r,s\to\infty} \frac{1}{s} \sum_{j=1}^{s} \left(\bar{\theta}_{\cdot j} - \bar{\theta}_{\cdot\cdot}\right)^2$$

exist and are finite. We first compute the asymptotic value of the lower bound B_{rs}, when $r, s \to \infty$.

THEOREM 4. *When n is fixed, the lower bound B_{rs} satisfies*

$$\lim_{r,s\to\infty} \frac{B_{rs}}{rs} = \frac{Q_{AB}\tau^2/n}{Q_{AB} + \tau^2/n}. \tag{4.1}$$

PROOF. Using the fact

$$\frac{1}{2\lambda + a} \le E\left[\frac{1}{2K + a}\right] \le \frac{1}{2\lambda + a - 2} \quad \text{if } K \sim \text{Poisson}(\lambda),\ a > 0 \tag{4.2}$$

[see, e.g., Green and Strawderman (1991)], we know U of (3.5) is such that

$$\lim_{r,s\to\infty} (rs)^{-1} U = \left(nQ_{AB}/\tau^2 + 1\right)^{-1}.$$

Therefore the lower bound B_{rs} of (3.4) satisfies

$$\lim_{r,s\to\infty} \frac{B_{rs}}{rs} = \frac{\tau^2}{n}\left(1 - \frac{1}{nQ_{AB}/\tau^2 + 1}\right) = \frac{Q_{AB}\tau^2/n}{Q_{AB} + \tau^2/n},$$

when Q_{AB}, Q_A, Q_B exist and are finite. This shows the result. □

It follows from (3.12) that $\mathrm{MSE}(\boldsymbol{\delta}^{\mathrm{JS}}, \boldsymbol{\theta})/rs$ has the same limit as in (4.1). Since $\mathrm{MSE}(\boldsymbol{\delta}^{+}, \boldsymbol{\theta}) < \mathrm{MSE}(\boldsymbol{\delta}^{\mathrm{JS}}, \boldsymbol{\theta})$, the positive James–Stein estimator $\boldsymbol{\delta}^{+}$ also possesses the limit.

Note that $\mathrm{MSE}(\boldsymbol{\delta}^0, \boldsymbol{\theta}) = rs\tau^2/n$, where $\boldsymbol{\delta}^0$ is the MLE of $\boldsymbol{\theta}$, so that, from Theorems 2 and 4, the lower bound of the limiting MSE ratio $\mathrm{MSE}(\boldsymbol{\delta}, \boldsymbol{\theta})/\mathrm{MSE}(\boldsymbol{\delta}^0, \boldsymbol{\theta})$ is given by $Q_{AB}/(Q_{AB} + \tau^2/n)$. In the following, we will give the conditions under which this limiting ratio achieves the minimum value.

We say the estimator $\boldsymbol{\delta}$ has *the minimum limiting MSE ratio* if

$$\lim_{r,s\to\infty} (rs)^{-1} \mathrm{MSE}(\boldsymbol{\delta}, \boldsymbol{\theta}) = \left(Q_{AB}\tau^2/n\right)/\left(Q_{AB} + \tau^2/n\right). \tag{4.3}$$

THEOREM 5. *An estimator* $\boldsymbol{\delta}$ *determined by* (1.8) *and* (3.13) *has the minimum limiting MSE ratio if*

$$C_{AB} \to (n-1)^{-1}, \qquad C_A \to 0, \qquad C_B \to 0 \tag{4.4}$$

in probability as $r, s \to \infty$.

PROOF. For $\boldsymbol{\delta}^c$ that is determined by (3.7) and (1.8) with $c_1 = (n-1)^{-1}$ and $c_2 = c_3 = 0$, its MSE obviously satisfies (4.3), by using the result of (3.8). The proof then follows from the fact that $C_{AB} \to c_1$, $C_A \to c_2$, $C_B \to c_3$ in probability and $(rs)^{-1}[\mathrm{MSE}(\boldsymbol{\delta}, \boldsymbol{\theta}) - \mathrm{MSE}(\boldsymbol{\delta}^c, \boldsymbol{\theta})] \to 0$. □

COROLLARY 1. *An estimator* $\boldsymbol{\delta}$ *determined by* (1.8) *has the minimum limiting MSE ratio if*

$$\rho_{AB} \to \frac{\tau^2}{nQ_{AB} + \tau^2}, \qquad \rho_A \to 0, \qquad \rho_B \to 0 \tag{4.5}$$

in probability as $r, s \to \infty$, *provided that* $Q_A \neq 0$, $Q_B \neq 0$.

PROOF. The result follows from Theorem 4 and the facts that

$$\frac{S^2}{T_{AB}^2} \to \frac{(n-1)\tau^2}{nQ_{AB} + \tau^2}, \qquad \frac{S^2}{T_A^2} \to \frac{(n-1)\tau^2}{nQ_A}, \qquad \frac{S^2}{T_B^2} \to \frac{(n-1)\tau^2}{nQ_B}$$

in probability as $r, s \to \infty$, recalling that $\rho_{AB} = C_{AB}S^2/T_{AB}^2$, $\rho_A = C_A S^2/T_A^2$ and $\rho_B = C_B S^2/T_B^2$. □

5. Modal estimator. Common hierarchical Bayesian analysis of the two-way model chooses, as the prior distribution for the hyperparameter $\boldsymbol{\phi} = (\mu, \tau^2, \sigma_{AB}^2, \sigma_A^2, \sigma_B^2)$,

$$p(\boldsymbol{\phi}) = p(\mu, \tau^2, \sigma_{AB}^2, \sigma_A^2, \sigma_B^2) = p(\mu)p(\tau^2)p(\sigma_{AB}^2, \sigma_A^2, \sigma_B^2), \tag{5.1}$$

where

$$\begin{aligned} p(\mu) &\propto 1, \\ p(\tau^2) &\propto (\tau^2)^{-(\nu_0+2)/2} \exp\left\{-\frac{\nu_0\lambda_0}{2\tau^2}\right\}, \\ p(\sigma_{AB}^2, \sigma_A^2, \sigma_B^2) &\propto 1. \end{aligned} \tag{5.2}$$

A similar prior structure for one-way hierarchical models was proposed by Leonard and Ord (1976). The posterior $p(\boldsymbol{\phi}|\mathbf{y})$ is proportional to $p(\mathbf{y}|\boldsymbol{\phi})p(\boldsymbol{\phi})$, of course, where $p(\mathbf{y}|\boldsymbol{\phi})$ is given by (1.6). Moreover, the posterior of the variance components $(\tau^2, \sigma_{AB}^2, \sigma_A^2, \sigma_B^2)$ has the form

$$\begin{aligned} &p(\tau^2, \sigma_{AB}^2, \sigma_A^2, \sigma_B^2|\mathbf{y}) \\ &\quad \propto (\tau^2)^{-(N+\nu_0+2)/2}(\tau^2 + n\sigma_{AB}^2)^{-(r-1)(s-1)/2} \\ &\qquad \times(\tau^2 + n\sigma_{AB}^2 + ns\sigma_A^2)^{-(r-1)/2}(\tau^2 + n\sigma_{AB}^2 + nr\sigma_B^2)^{-(s-1)/2} \\ &\qquad \times \exp\left\{-\frac{1}{2}\left[\frac{S^2 + \nu_0\lambda_0}{\tau^2} + \frac{T_{AB}^2}{\tau^2 + n\sigma_{AB}^2}\right.\right. \\ &\qquad\qquad \left.\left. + \frac{T_A^2}{\tau^2 + n\sigma_{AB}^2 + ns\sigma_A^2} + \frac{T_B^2}{\tau^2 + n\sigma_{AB}^2 + nr\sigma_B^2}\right]\right\} \end{aligned} \tag{5.3}$$

[after integrating out μ from $p(\boldsymbol{\phi}|\mathbf{y})$], where $N = (n-1)rs$.

The modal estimator δ_{ij}^M has the form of (1.5) with the replacements of μ by $\hat{\mu} = \bar{y}_{...}$ and $(\tau^2, \sigma_{AB}^2, \sigma_A^2, \sigma_B^2)$ by the mode of the posterior $p(\tau^2, \sigma_{AB}^2, \sigma_A^2, \sigma_B^2|\mathbf{y})$:

$$\begin{aligned} &\hat{\tau}^2 = \frac{S^2 + \nu_0\lambda_0}{N + \nu_0 + 2}, \qquad \hat{\sigma}_{AB}^2 = \frac{1}{n}\left(\frac{T_{AB}^2}{(r-1)(s-1)} - \hat{\tau}^2\right)_+, \\ &\hat{\sigma}_A^2 = \frac{1}{ns}\left(\frac{T_A^2}{r-1} - \hat{\tau}^2 - n\hat{\sigma}_{AB}^2\right)_+, \\ &\hat{\sigma}_B^2 = \frac{1}{nr}\left(\frac{T_B^2}{s-1} - \hat{\tau}^2 - n\hat{\sigma}_{AB}^2\right)_+. \end{aligned} \tag{5.4}$$

THEOREM 6. *The modal estimator* $\boldsymbol{\delta}^M$ *is minimax when* $\lambda_0 = 0$ *and*

$$\nu_0 \geq -[(n-1)rs + 2]\min\left\{\frac{(r-1)(s-1) - 4}{(r-1)(s-1) - 2}, \frac{r-5}{r-3}, \frac{s-5}{s-3}\right\}.$$

If $n > 1$, $r \geq 5$ *and* $s \geq 5$, *then the estimator is minimax when* $\nu_0 \geq 0$.

PROOF. The proof is straightforward application of Theorem 3. □

Asymptotic results about $\boldsymbol{\delta}^M$ can also be obtained. Noting that $N = (n-1)rs$ and $T_{AB}^2/S^2 \to (nQ_{AB} + \tau^2)/[(n-1)\tau^2]$ in probability as $r, s \to \infty$, we have

$$C_{AB}^M \to (n-1)^{-1}$$

in probability. It is also noticed that, as $r, s \to \infty$,

$$C_A^M \leq \frac{(r-1)(S^2 + \nu_0 \lambda_0)}{(N + \nu_0 + 2)S^2} \to 0 \quad \text{so that } C_A^M \to 0.$$

Similarly, $C_B^M \to 0$. From Theorem 4, the modal estimator $\boldsymbol{\delta}^M$ has the minimum limiting MSE ratio.

Acknowledgments. I would like to thank an Associate Editor and the referees for their constructive suggestions and comments and for providing some useful references. I am grateful to Professor J. V. Zidek for his helpful comments and to Professors I. Guttman and T. Leonard for their help with many aspects of the paper.

REFERENCES

BARANCHIK, A. J. (1970). A family of minimax estimators of the mean of a multivariate normal distribution. *Ann. Math. Statist.* **41** 642–645.

BERGER, J. (1980). *Statistical Decision Theory: Foundations, Concepts, and Methods*. Springer, New York.

BERGER, J. (1982). Selecting a minimax estimator of a multivariate normal mean. *Ann. Statist.* **10** 81–92.

BERGER, J. and WOLPERT, R. (1983). Estimating the mean function of a Gaussian process and the Stein effect. *J. Multivariate Anal.* **13** 401–424.

BOX, G. E. P. and TIAO, G. C. (1968). Bayesian estimation of means for the random effect model. *J. Amer. Statist. Assoc.* **63** 174–181.

CASELLA, G. and HWANG, J. T. (1982). Limit expressions for the risk of James–Stein estimators. *Canad. J. Statist.* **10** 305–309.

EFRON, B. and MORRIS, C. (1973). Stein's estimation rule and its competitors—an empirical Bayes approach. *J. Amer. Statist. Assoc.* **68** 117–130.

GIRI, N. (1970). Bayesian estimation of means of a two-way classification with random effect model. *Statist. Hefte* (N.F.) **2** 254–260.

GREEN, E. J. and STRAWDERMAN, W. E. (1991). A James–Stein type estimator for combining unbiased and possibly biased estimators. *J. Amer. Statist. Assoc.* **86** 1001–1006.

JAMES, W. and STEIN, C. (1961). Estimation with quadratic loss. *Proc. Fourth Berkeley Symp. Math. Statist. Probab.* **1** 361–380. Univ. California Press, Berkeley.

LEONARD, T. (1976). Some alternative approaches to multiparameter estimation. *Biometrika* **63** 69–75.

LEONARD, T. and ORD, K. (1976). An investigation of the *F*-test procedure as an estimation short-cut. *J. Roy. Statist. Soc. Ser. B* **38** 95–98.

LINDLEY, D. V. (1972). A Bayesian solution for two-way analysis of variance. *Colloq. Math. Soc. János Bolyai* **9** 475–496.

LINDLEY, D. V. and SMITH, A. F. M. (1972). Bayes estimates for the linear model (with discussion). *J. Roy. Statist. Soc. Ser. B* **34** 1–41.

MORRIS, C. N. (1983). Parametric empirical Bayes inference: theory and applications (with discussion). *J. Amer. Statist. Assoc.* **78** 47–65.

SMITH, A. F. M. (1973). Bayes estimates in one-way and two-way models. *Biometrika* **60** 319–329.

STEIN, C. (1956). Inadmissibility of the usual estimator of the mean of a multivariate normal distribution. *Proc. Third Berkeley Symp. Math. Statist. Probab.* **1** 197–206. Univ. California Press, Berkeley.

STEIN, C. (1966). An approach to the recovery of interblock information in balanced incomplete designs. In *Festschrift for J. Neyman* (F. N. David, ed.). Wiley, New York.

STEIN, C. (1981). Estimation of the mean of a multivariate normal distribution. *Ann. Statist.* **9** 1135–1151.

STRAWDERMAN, W. E. (1971). Proper Bayes minimax estimators of the multivariate normal mean. *Ann. Math. Statist.* **42** 385–388.

SUN, L. (1992). The Bayes estimation procedures for one and two-way hierarchical models. Ph.D. dissertation, Dept. Statistics, Univ. Toronto.

DEPARTMENT OF MATHEMATICS AND STATISTICS
SIMON FRASER UNIVERSITY
BURNABY, BRITISH COLUMBIA
V5A 1S6 CANADA