The Korean Communications
in Statistics Vol. 9, No. 2, 2002
pp. 315-324

Hierarchical Bayes Estimators of the Error Variance in Two-Way ANOVA Models[1)]

In Hong Chang[2)] and Byung Hwee Kim[3)]

Abstract

For estimating the error variance under the relative squared error loss in two-way analysis of variance models, we provide a class of hierarchical Bayes estimators and then derive a subclass of the hierarchical Bayes estimators, each member of which dominates the best multiple of the error sum of squares which is known to be minimax. We also identify a subclass of non-minimax hierarchical Bayes estimators.

Keywords : Balanced two-way ANOVA model, Error variance, Hierarchical Bayes estimation, Minimaxity, Relative squared error loss.

1. Introduction

Consider the following two-way analysis of variance(ANOVA) models :

$$y_{ijk}=\theta_{ij}+\varepsilon_{ijk}, \quad i=1,2,\cdots,p(>1)\,;\; j=1,2,\cdots,q(>1),\; k=1,2,\cdots,n(>1),$$

where the ε_{ijk}'s are independently and identically distributed(i.i.d.) as $N(0,\sigma^2)$. Then the minimal sufficient statistic for $(\theta_{11},\cdots,\theta_{pq},\sigma^2)$ is $(\overline{y_{11.}},\cdots,\overline{y_{pq.}},S)$, where $\overline{y_{ij.}}=\frac{1}{n}\sum_{k=1}^{n}y_{ijk}$ and $S=\sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{n}(y_{ijk}-\overline{y_{ij.}})^2$. Here we use the notations

$$\overline{y_{ij.}}=y_{ij},\quad \boldsymbol{y}=(y_{11},\cdots,y_{pq})^T \text{ and } \boldsymbol{\theta}=(\theta_{11},\cdots,\theta_{pq})^T,$$

where "T" denotes the transpose. Then $\boldsymbol{y}$ and S are independent with

$$\boldsymbol{y}\sim N(\boldsymbol{\theta},\frac{\sigma^2}{n}I_{pq})$$

and

$$S\sim\sigma^2\chi^2_{pq(n-1)},$$

---

1) Research by the financial support of Hanyang University, Korea, made in the program year of 1999

2) Lecturer, Department of Mathematics, Hanyang University, Seoul 133-791, Korea
3) Professor, Department of Mathematics, Hanyang University, Seoul 133-791, Korea
E-mail : bkim@hanyang.ac.kr

where I_{pq} is the $p\times q$ identity matrix.

Consider the problem of estimating for the error variance σ^2 under the relative squared error loss

$$L(\sigma^2,\delta) = (\delta\sigma^{-2}-1)^2 = \sigma^{-4}(\delta-\sigma^2)^2. \tag{1.1}$$

The best affine equivariant estimator is $\delta_0(S)=(pq(n-1)+2S)^{-1}$ which is a minimax estimator with constant risk $2(pq(n-1)+2)^{-1}$. Stein(1964) showed that δ_0 can be improved by considering a class of scale equivariant estimators $\delta(W,S)=\phi(W)S$ for $W=n\, y^T(I_{pq}-X(X^TX)^{-1}X^T)y/S$, where X is a $pq\times k$ matrix with rank $k(<pq)$. He found a specific better estimator $\delta(W,S)=\phi^{ST}(W)S$,

where $\phi^{ST}(W)=\min\left\{\frac{1}{pq(n-1)+2},\ \frac{1}{pqn-k+2}(1+W)\right\}$. Brewster and Zidek(1974) derived an improved generalized Bayes estimator

$$\delta^{BZ}(W,S)=\phi^{BZ}(W)S, \tag{1.2}$$

where

$$\phi^{BZ}(W)=\frac{1}{pqn-k+2}\frac{\int_0^1\lambda^{\frac{pq-k}{2}-1}(1+\lambda W)^{-\frac{pqn-k}{2}-1}d\lambda}{\int_0^1\lambda^{\frac{pq-k}{2}-1}(1+\lambda W)^{-\frac{pqn-k}{2}-2}d\lambda}.$$

Though Stein's estimator δ^{ST} improves on δ_0 at $\eta=\frac{n\theta^T(I_{pq}-X(X^TX)^{-1}X^T)\theta}{2\sigma^2}=0$, it is not an analytic function and hence is inadimissible. On the other hand, Brewster-Zidek's estimator δ^{BZ} does not improve on δ_0 at $\eta=0$ though it is admissible as shown in Proskin(1985). Therefore it is desirable to get better analytic estimators dominating δ_0, especially at $\eta=0$.

To do this, we consider the hierarchical Bayes estimation for σ^2 under the loss(1.1). In Section 2, we develop a class of hierarchical Bayes estimators of σ^2. In Section 3, we provide a subclass, each member of which dominates δ_0 and also identify a subclass non-minimax hierarchical Bayes estimators. All the results of this paper can be regarded as a two-way extension of the results given by Datta and Ghosh(1995). In Section 4, we compare numerically risks of the hierarchical Bayes estimators with risk of δ_0 for some special cases.

2. Hierarchical Bayes Estimators

Consider the following hierarchical Bayesian model:

(Ⅰ) Conditionally on $\boldsymbol{\theta}$, $\boldsymbol{\alpha}=(\alpha_1,\cdots,\alpha_k)^T$, σ^2, and σ_1^2, $\boldsymbol{y}$ and S are mutually independent with $\boldsymbol{y}\sim N_{pq}(\boldsymbol{\theta}, \frac{\sigma^2}{n}I_{pq})$ and $S\sim\sigma^2\chi^2_{pq(n-1)}$;

(Ⅱ) Conditionally on $\boldsymbol{\alpha}$, σ^2, and σ_1^2, $\boldsymbol{\theta}\sim N_{pq}(X\boldsymbol{\alpha}, \sigma_1^2 I_{pq})$, where X is a $pq\times k$ matrix of rank $k(<pq)$;

(Ⅲ) $(\boldsymbol{\alpha}, \sigma^2, \sigma_1^2)\sim(\sigma^2)^{\frac{a-b}{2}-1}(\sigma^2+n\sigma_1^2)^{-\frac{4-b}{2}}$, $0\le a<pqn-k+2$, $0\le b<pq-k+2$.

Remark 2.1 Note that (Ⅰ) and (Ⅱ) represent a two-fold nested classification mixed effect model if we let $\theta_{ij}=\mu_i+\gamma_{ij}$ where μ_i's are fixed, γ_{ij}'s are i.i.d. as $N(0,\sigma_1^2)$, and γ_{ij}'s and ε_{ijk}'s are independent. In this case, $k=p$, $X=I_p\otimes\mathbf{1}_q$, and $\boldsymbol{\alpha}=(\mu_1,\cdots,\mu_p)^T$. Also, (Ⅰ) and (Ⅱ) represent a two-way crossed classification mixed effect model with intersections if we let $\theta_{ij}=\mu_i+\beta_j+\gamma_{ij}$, where μ_i's and β_j's are fixed with $\sum_{j=0}^{q}\beta_j=0$, γ_{ij}'s are i.i.d. as $N(0,\sigma_1^2)$, and γ_{ij}'s and ε_{ijk}'s are independent. In this case, $k=p+q-1$, $X=\left(I_p\otimes\mathbf{1}_q,\ \mathbf{1}_p\otimes\begin{pmatrix}I_{q-1}\\ -\mathbf{1}^T_{q-1}\end{pmatrix}\right)$, and $\boldsymbol{\alpha}=(\mu_1,\cdots,\mu_p,\beta_1,\cdots,\beta_{q-1})$. Here, $\mathbf{1}_q$ denotes the $q\times 1$ matrix with 1 as all elements and $\otimes$ the usual kronecker's product.

Remark 2.2 (Ⅲ) is equivalent to say that $\pi(\boldsymbol{\alpha})\propto 1$, $\pi(\sigma^2,\sigma_1^2)\propto(\sigma^2)^{\frac{a-b}{2}-1}(\sigma^2+n\sigma_1^2)^{-\frac{4-b}{2}}$, and $\boldsymbol{\alpha}$ and (σ^2,σ_1^2) are independent. This prior for (σ^2,σ_1^2) contains Jeffreys' noninformative prior with $a=b=2$ as a special case and was first used in Portnoy(1971) with slightly different notations who treated the problem of estimating σ_1^2 under scale invariant loss in the one-way random effect model $y_{ij}=\mu+e_i+e_{ij}$, $i=1,2,\cdots,p$, $j=1,2,\cdots,q$ where μ is fixed, e_i's are i.i.d. as $N(0,\sigma_1^2)$, e_{ij}'s are i.i.d. as $N(0,\sigma^2)$, and e_i's and e_{ij}'s are independent.

From (Ⅰ) and (Ⅱ), the conditional density of $\boldsymbol{y}$, S, and $\boldsymbol{\theta}$ given $\boldsymbol{\alpha}$, σ^2, and σ_1^2 is

$$\begin{aligned}
&f(\boldsymbol{y}, s, \boldsymbol{\theta}\mid\boldsymbol{\alpha},\sigma^2,\sigma_1^2)\\
&= f(\boldsymbol{y}, s\mid\boldsymbol{\theta},\boldsymbol{\alpha},,\sigma^2,\sigma_1^2)\cdot\pi(\boldsymbol{\theta}\mid\boldsymbol{\alpha},\sigma_1^2,\sigma^2)\\
&\propto (\sigma^2)^{-\frac{pq}{2}}(\sigma_1^2)^{-\frac{pq}{2}}e^{-\frac{1}{2}[\frac{n}{\sigma^2}(\boldsymbol{y}-\boldsymbol{\theta})^T(\boldsymbol{y}-\boldsymbol{\theta})+\frac{1}{\sigma_1^2}(\boldsymbol{\theta}-X\boldsymbol{\alpha})^T(\boldsymbol{\theta}-X\boldsymbol{\alpha})]}(\sigma^2)^{-\frac{pq(n-1)}{2}}s^{-\frac{pq(n-1)}{2}-1}e^{-\frac{s}{2\sigma^2}}.
\end{aligned}\tag{2.1}$$

Integrating out $\boldsymbol{\theta}$ in (2.1) gives the conditional density of $\boldsymbol{y}$ and S given $\boldsymbol{\alpha}$, σ^2, and σ_1^2

$$\begin{aligned}
&f(\boldsymbol{y}, s\mid\boldsymbol{\alpha},\sigma^2,\sigma_1^2)\\
&= (\sigma^2+n\sigma_1^2)^{-\frac{pq}{2}}(\sigma^2)^{-\frac{pq(n-1)}{2}}s^{-\frac{pq(n-1)}{2}}\\
&\quad e^{-\frac{1}{2}[\frac{n}{\sigma^2+n\sigma_1^2}(\boldsymbol{\alpha}-(X^TX)^{-1}X^T\boldsymbol{y})^T(X^TX)(\boldsymbol{\alpha}-(X^TX)^{-1}X^T\boldsymbol{y})+\frac{n}{\sigma^2+n\sigma_1^2}\boldsymbol{y}^T(I_{pq}-X(X^TX)^{-1}X^T)\boldsymbol{y}+\frac{s}{\sigma^2}]}.
\end{aligned}$$

Then the posterior density of $\boldsymbol{\alpha}$, σ^2, and σ_1^2 is given by

$$\pi(\boldsymbol{\alpha}, \sigma^2, \sigma_1^2 \mid \boldsymbol{y}, s)$$
$$\propto f(\boldsymbol{y}, s \mid \boldsymbol{\alpha}, \sigma^2, \sigma_1^2) \cdot \pi(\boldsymbol{\alpha}, \sigma^2, \sigma_1^2)$$
$$\propto (\sigma^2 + n\sigma_1^2)^{-\frac{pq-b+4}{2}} (\sigma^2)^{-\frac{pq(n-1)-a+b+2}{2}}$$
$$e^{-\frac{1}{2}[\frac{n}{\sigma^2+n\sigma_1^2}(\boldsymbol{\alpha}-(X^TX)^{-1}X^T\boldsymbol{y})^T(X^TX)(\boldsymbol{\alpha}-(X^TX)^{-1}X^T\boldsymbol{y})+\frac{n}{\sigma^2+n\sigma_1^2}\boldsymbol{y}^T(I_{pq}-X(X^TX)^{-1}X^T)\boldsymbol{y}+\frac{s}{\sigma^2}]}. \quad (2.2)$$

Next, integrating with respect to $\boldsymbol{\alpha}$ in (2.2), it follows that the posterior density of σ^2 and σ_1^2 is

$$\pi(\sigma^2, \sigma_1^2 \mid \boldsymbol{y}, s)$$
$$\propto (\sigma^2 + n\sigma_1^2)^{-\frac{pq-k-b+4}{2}} (\sigma^2)^{-\frac{pq(n-1)-a+b+2}{2}} e^{-\frac{n}{2(\sigma^2+n\sigma_1^2)}\boldsymbol{y}^T(I_{pq}-X(X^TX)^{-1}X^T)\boldsymbol{y}-\frac{s}{2\sigma^2}}. \quad (2.3)$$

Now we use the transformation $r = \frac{1}{\sigma^2}$ and $\lambda = \frac{\sigma^2}{n\sigma_1^2+\sigma^2}$. Then, from (2.3), we get the posterior density of r and λ given by

$$\pi(r, \lambda \mid \boldsymbol{y}, s) \propto r^{\frac{pqn-k-a}{2}} \lambda^{\frac{pq-k-b}{2}} e^{-\frac{r}{2}(n\lambda \boldsymbol{y}^T(I_{pq}-X(X^TX)^{-1}X^T)\boldsymbol{y}+s)}. \quad (2.4)$$

From (2.4) we have

$$\pi(r \mid \lambda, \boldsymbol{y}, s) \propto r^{\frac{pqn-k-a}{2}} e^{-\frac{r}{2}(n\lambda \boldsymbol{y}^T(I_{pq}-X(X^TX)^{-1}X^T)\boldsymbol{y}+s)}$$

and

$$\pi(\lambda \mid \boldsymbol{y}, s) \propto \lambda^{\frac{pq-k-b}{2}} \cdot [n\lambda \boldsymbol{y}^T(I_{pq}-X(X^TX)^{-1}X^T)\boldsymbol{y}+s]^{-\frac{pqn-k-a+2}{2}}.$$

Theorem 2.1 Under the loss (1.1) the hierarchical Bayes estimator of $\sigma^2 = \frac{1}{r}$ is given by

$$\delta_{a,b}^{HB}(\boldsymbol{y}, S) = \phi_{a,b}^{HB}(W, S) = \delta_{a,b}^{HB}(W, S), \quad (2.5)$$

where $$\phi_{a,b}^{HB}(w) = \frac{1}{pqn-k-a+4} \frac{\int_0^1 \lambda^{\frac{pq-k-b}{2}-1}(1+\lambda w)^{-\frac{pqn-k-a}{2}-2} d\lambda}{\int_0^1 \lambda^{\frac{pq-k-b}{2}}(1+\lambda w)^{-\frac{pqn-k-a}{2}-3} d\lambda}$$

with $w = \frac{n\boldsymbol{y}^T(I_{pq}-X(X^TX^T)^{-1}X^T)\boldsymbol{y}}{s}$.

Proof. Under the loss (1.1) the posterior risk of an estimator $\delta(\boldsymbol{y})$ of $\sigma^2 = \frac{1}{r}$ is given by

$$E[\sigma^{-4}(\delta(\boldsymbol{y},s)-\sigma^2)^2 \mid \boldsymbol{y},s] = E\left[r^2\left(\delta(\boldsymbol{y},s)-\frac{1}{r}\right)^2 \mid \boldsymbol{y},s\right]$$
$$= E(r^2 \mid \boldsymbol{y},s)\left[\delta(\boldsymbol{y},s) - \frac{E(r \mid \boldsymbol{y},s)}{E(r^2 \mid \boldsymbol{y},s)}\right]^2 + 1 - \frac{[E(r \mid \boldsymbol{y},s)]^2}{E(r^2 \mid \boldsymbol{y},s)}.$$

Hence the posterior risk is minimized when

$$\delta(\mathbf{y},s)=\delta^{HB}_{a,b}(\mathbf{y},s)=\frac{E(r\,|\,\mathbf{y},s)}{E(r^2\,|\,\mathbf{y},s)}=\frac{E\{E(r\,|\,\lambda,\ \mathbf{y},s)\,|\,\mathbf{y},s\}}{E\{E(r^2\,|\,\lambda,\ \mathbf{y},s)\,|\,\mathbf{y},s\}}. \tag{2.6}$$

Now,

$$E(r\,|\,\lambda,\ \mathbf{y},s)=(pqn-k-a+2)[n\lambda\mathbf{y}^T(I_{pq}-X(X^TX)^{-1}X^T)\mathbf{y}+s]^{-1} \tag{2.7}$$

and

$$E(r^2\,|\,\lambda,\ \mathbf{y},s)=(pqn-k-a+4)(pqn-k-a+2)[n\lambda\mathbf{y}^T(I_{pq}-X(X^TX)^{-1}X^T)\mathbf{y}+s]^{-2}. \tag{2.8}$$

From (2.6), (2.7) and (2.8), we get

$$\begin{aligned}\frac{E(r\,|\,\mathbf{y},s)}{E(r^2\,|\,\mathbf{y},s)}&=\frac{1}{pqn-k-a+4}\,\frac{E[\,(n\lambda\mathbf{y}^T(I_{pq}-X(X^TX)^{-1}X^T)\mathbf{y}+s)^{-1}\,|\,\mathbf{y},s]}{E[\,(n\lambda\mathbf{y}^T(I_{pq}-X(X^TX)^{-1}X^T)\mathbf{y}+s)^{-2}\,|\,\mathbf{y},s]}\\&=\frac{S}{pqn-k-a+4}\,\frac{\int_0^1\lambda^{\frac{pq-k-b}{2}}(1+\lambda w)^{-\frac{pqn-k-a}{2}-2}d\lambda}{\int_0^1\lambda^{\frac{pq-k-b}{2}}(1+\lambda w)^{-\frac{pqn-k-a}{2}-3}d\lambda}.\end{aligned}$$

It is noted that $\delta^{HB}_{2,2}$ coincides with Brewster-Zidek's estimator δ^{BZ} in (1.2).

3. Minimaxity and Non-minimaxity

Making a transformation $\frac{\lambda w}{1+\lambda w}=z$, $\phi^{HB}_{a,a}(w)$ in (2.5) becomes

$$\phi^{HB}_{a,b}(W)=\frac{S}{pqn-k-a+4}\,\frac{\int_0^{\frac{W}{1+W}}z^{\frac{pq-k-b}{2}}(1-z)^{-\frac{pq(n-1)-a+b}{2}}dz}{\int_0^{\frac{W}{1+W}}z^{\frac{pq-k-b}{2}}(1-z)^{-\frac{pq(n-1)-a+b+2}{2}}dz}. \tag{3.1}$$

Next $\phi^{HB}_{a,b}(w)$ in (3.1) is represented through the hypergeometric function

$$F(c_1,c_2,c_3,x)=1+\sum_{l=1}^{\infty}\frac{(c_1)_l\,(c_2)_l}{(c_3)_l}\frac{x^l}{l!}\quad\text{for}\quad (c)_l=c(c+1)\cdots(c+l-1).$$

The following facts about $F(c_1,c_2,c_3,x)$, from Abramobitz and Stegun(1964), are need;

$$\int_0^x t^{c_1}-1(1-t)^{c_2-1}dt=\frac{x^{c_1}}{c_1}F(c_1,1-c_2,c_1+1,x)\ \text{for}\ c_1,c_2>1, \tag{3.2}$$

$$F(c_1,c_2,c_3,x)=(1-x)^{c_3-c_1-c_2}F(c_3-c_1,c_3-c_2,c_3,x), \tag{3.3}$$

$$\begin{aligned}(c_3-c_1-c_2)\ &F(c_1,c_2,c_3,x)-(c_3-c_1)F(c_1-1,c_2,c_3,x)\\&+c_2(1-x)F(c_1,c_2+1,c_3,x)=0,\end{aligned} \tag{3.4}$$

$$\begin{aligned}(c_2-c_1)(1-x)\ &F(c_1,c_2,c_3,x)-(c_3-c_1)F(c_1-1,c_2,c_2,x)\\&+(c_3-c_2)F(c_1,c_2-1,c_3,x)=0,\end{aligned} \tag{3.5}$$

$$F(c_1, c_2, c_3, 1) = \infty \quad \text{when} \quad c_3 - c_1 - c_2 \le -1. \tag{3.6}$$

Using (3.2) and (3.3), (3.1) becomes

$$\frac{\int_0^{\frac{w}{1+w}} z^{\frac{pq-k-b}{2}} (1-z)^{-\frac{pq(n-1)-a+b}{2}} dz}{\int_0^{\frac{w}{1+w}} z^{\frac{pq-k-b}{2}} (1-z)^{-\frac{pq(n-1)-a+b+2}{2}} dz}$$

$$= (1+w) \frac{F(1, \frac{pqn-k-a+4}{2}, \frac{pq-k-b+4}{2}, \frac{w}{1+w})}{F(1, \frac{pqn-k-a+6}{2}, \frac{pq-k-b+4}{2}, \frac{w}{1+w})}.$$

Moreover by (3.4) and (3.5), $\phi^{HB}_{a,b}(w)$ is expressed as

$$\phi^{HB}_{a,b}(w) = \frac{1}{pq(n-1)-a+b+2} \times \left(1 - \frac{pq-k-b+2}{(pq(n-1)-a+b+2) F(1, \frac{pqn-k-a+4}{2}, \frac{pq-k-b+4}{2}, \frac{w}{1+w}) + (pq-k-b+2)}\right). \tag{3.7}$$

Considering the case of $a = b$ and making use of (3.7), we prove the following theorem :

Theorem 3.1 The estimator $\delta^{HB}_{a,a}$ with $2 \le a < pq-k+2$ is a minimax estimator of σ^2 under the loss (1.1).

Proof. We verify that $\phi^{HB}_{a,a}(w)$ with $2 \le a < pq-k+2$ satisfies the condition for minimaxity proposed by Brewster and Zidek(1974) : $\phi^{HB}_{a,a}(w)$ is nondecreasing in w for $2 \le a < pq-k+2$ and $\phi^{BZ}(w) = \phi^{HB}_{2,2}(w) \le \phi^{HB}_{a,a}(w) \le \frac{1}{pq(n-1)+2}$ for all w and $2 \le a < pq-k+2$. Now,

$$\phi^{HB}_{a,a}(w) = \frac{1}{pq(n-1)+2} \times \left(1 - \frac{pq-k-a+2}{(pq(n-1)+2) F(1, \frac{pqn-k-a+4}{2}, \frac{pq-k-a+4}{2}, \frac{w}{1+w}) + (pq-k-a+2)}\right).$$

Since $F(1, \frac{pqn-k-a+4}{2}, \frac{pq-k-a+4}{2}, \frac{w}{1+w})$ is increasing in w for $2 \le a < pq-k+2$, $\phi^{HB}_{a,a}(w)$ is increasing in w. And since

$$\frac{pq-k-a+4}{2} - \frac{pqn-k-a+4}{2} - 1 = -\frac{pq(n-1)+2}{2} \le -1,$$

it is clear from (3.6) that $\lim_{w \to \infty} \phi^{HB}_{a,a}(w) = \frac{1}{pq(n-1)+2}$. Also, for $2 \le a < pq-k+2$, $\phi^{HB}_{a,a}(w)$ is increasing in a since

$$F(1, \frac{pqn-k-a+4}{2}, \frac{pq-k-a+4}{2}, \frac{w}{1+w})$$
$$= 1 + \frac{pqn-k-a+4}{pq-k-a+4}\left(\frac{w}{1+w}\right) + \frac{(pqn-k-a+4)(pqn-k-a+6)}{(pq-k-a+4)(pq-k-a+6)}\left(\frac{w}{1+w}\right)^2 + \cdots$$
$$= 1 + \left(1 + \frac{pq(n-1)}{pq-k-a+4}\right)\left(\frac{w}{1+w}\right) + \left(1 + \frac{pq(n-1)}{pq-k-a+4} + \frac{pq(n-1)}{pq-k-a+6}\right)\left(\frac{w}{1+w}\right)^2 + \cdots$$

is increasing in a. Hence $\delta^{HB}_{a,a}=\phi^{HB}_{a,a}(W)S$ is minimax for $2\le a< pq-k+2$.

Theorem 3.2 The estimator $\delta^{HB}_{a,a}$ with $0\le a< 2$ is not minimax.

Proof. Without loss of generality we assume $\sigma^2=1$. Let $U=n\boldsymbol{y}^T(I_{pq}-X(X^TX)^{-1}X^T)\boldsymbol{y}$. Then $U\sim\chi^2_{pq-k}(\eta)$ with $\eta=\dfrac{n\boldsymbol{\theta}^T(I_{pq}-X(X^TX)^{-1}X^T)\boldsymbol{\theta}}{2}$. When $\eta=0$, $U\sim\chi^2_{pq-k}$ with p.d.f. $f_{pq-k}(\lambda)$, and c.d.f. $F_{pq-k}(u)$, $u>0$. Let $W=\dfrac{U}{S}$ and $V=S\sim\chi^2_{pq(n-1)}$. Then $U=WV$ and $S=V$. Also U and S are independent. Then the joint p.d.f. $f(u,s)$ of U and S is given by

$$f(u,s)=f_{pq-k}(u)f_{pq(n-1)}(s).$$

And the joint p.d.f. $g(w,v)$ of W and V is $g(w,v)=f_{pq-k}(wv)f_{pq(n-1)}(v)\cdot v$. Now by using Kubokawa's(1994) method we first investigate the nature of the risks of $\delta^{HB}_{a,a}$ and δ_0 at $\eta=0$. The risk of $\delta_{a,a}$ at $\eta=0$ is written as

$$R(0,\delta^{HB}_{a,a})=\int_0^\infty\int_0^\infty(\phi^{HB}_{a,a}(w)v-1)^2vf_{pq-k}(wv)f_{pq(n-1)}(v)\,dwdv.$$

Since $\lim_{w\to\infty}\phi^{HB}_{a,a}(w)=\dfrac{1}{pq(n-1)+2}$, the risk of δ_0 at $\eta=0$ is given by

$$\begin{aligned}
R(0,\delta_0) &= \int_0^\infty\Big(\frac{1}{pq(n-1)+2}v-1\Big)^2 f_{pq(n-1)}(v)\,dv\\
&= \Big[\int_0^\infty(\phi^{HB}_{a,a}(w)v-1)^2 f_{pq(n-1)}(v)F_{pq-k}(wv)\,dv\ \Big]^\infty_{w=0}\\
&= \int_0^\infty\frac{d}{dw}\Big\{\int_0^\infty[\phi^{HB}_{a,a}(w)v-1]^2 f_{pq(n-1)}(v)F_{pq-k}(wv)\,dv\Big\}dw\\
&= \int_0^\infty\int_0^\infty\frac{d}{dw}\{(\phi^{HB}_{a,a}(w)v-1)^2 f_{pq(n-1)}(v)F_{pq-k}(wv)\}\,dvdw\\
&= \int_0^\infty\int_0^\infty\Big\{2(\phi^{HB}_{a,a}(w)v-1)\Big(v\frac{d\phi^{HB}_{a,a}(w)}{dw}\Big)f_{pq(n-1)}(v)F_{pq-k}(wv)\Big\}\,dvdw\\
&\quad+\int_0^\infty\int_0^\infty(\phi^{HB}_{a,a}(w)v-1)^2 f_{pq(n-1)}(v)vf_{pq-k}(wv)\,dvdw.
\end{aligned}$$

Therefore the risk difference δ_0 and $\delta_{a,a}$ at $\eta=0$ is written as

$$\begin{aligned}
R(0,\delta_0)-R(0,\delta^{HB}_{a,a}) &= 2\int_0^\infty\int_0^\infty\Big(v\frac{d\phi^{HB}_{a,a}(w)}{dw}\Big)(\phi^{HB}_{a,a}(w)v-1)f_{pq(n-1)}(v)F_{pq-k}(wv)\,dvdw\\
&= 2\int_0^\infty\int_0^\infty\Big(\frac{d\phi^{HB}_{a,a}(w)}{dw}\Big)\phi^{HB}_{a,a}(w)v^2 f_{pq(n-1)}(v)F_{pq-k}(wv)\,dvdw\\
&\quad-2\int_0^\infty\Big(\frac{d\phi^{HB}_{a,a}(w)}{dw}\Big)\Big\{\int_0^\infty vf_{pq(n-1)}(v)F_{pq-k}(wv)\,dv\Big\}dw.
\end{aligned}$$

Since $\phi_{2,2}^{HB}(w)=\dfrac{\int_0^{\infty} v f_{pq(n-1)}(v)F_{pq-k}(wv)\,dv}{\int_0^{\infty} v^2 f_{pq(n-1)}(v)F_{pq-k}(wv)\,dv}$,

$$\begin{aligned} &R(0,\ \delta_0)-R(0,\ \delta_{a,a}^{HB}) \\ &= 2\int_0^{\infty}\left(\frac{d\phi_{a,a}^{HB}(w)}{dw}\right)(\phi_{a,a}^{HB}(w)-\phi_{2,2}^{HB}(w))\left\{\int_0^{\infty} v^2 f_{pq(n-1)}(v)F_{pq-k}(wv)\,dv\right\}dw. \end{aligned}$$

We see that Brewster-Zidek's estimator $\delta^{BZ}(W,S)=\delta_{2,2}^{HB}(W,S)=\phi_{2,2}^{HB}(W)S$ does not improve on the best equivariant estimator $\delta_0=\dfrac{1}{pq(n-1)+2}S$ at $\eta=0$. See also Rukhin(1992). On the other hand, since $\phi_{a,a}^{HB}(w)$ is strictly increasing in a, $\delta_{a,a}^{HB}$ with $2<a<pq-k+2$ improves on δ_0, especially at $\eta=0$. $\delta_{a,a}^{HB}$ with $0\le a<2$ does not improve on δ_0 at $\eta=0$, i.e., $R(0,\delta_0)-R(0,\delta_{a,a})<0$ for $0\le a<2$ which implies that $\delta_{a,a}^{HB}$ with $0\le a<2$ is not minimax.

Remark 3.1 Maruyama(1998) proposed the minimax generalized Bayes estimator $\delta_a^M(W,S)=\phi_a^M(W)S$ with $a\ge 1$, where

$$\phi_a^M(w)=\frac{1}{pqn-k+2}\,\frac{\int_0^1 \lambda^{\frac{a(pq-k)}{2}-1}(1+\lambda w)^{-a[\frac{pqn-k}{2}+1]}d\lambda}{\int_0^1 \lambda^{\frac{a(pq-k)}{2}-1}(1+\lambda w)^{-a[\frac{pqn-k}{2}+1]-1}d\lambda},$$

Cleary δ_1^M coincides with $\delta^{BZ}=\delta_{2,2}^{HB}$. He showed that δ_a^M with $a>1$ improves on δ_0 at $\eta=0$. δ_a^M with $a>1$ differs from $\delta_{a,a}^{HB}$ with $2<a<pq-k+2$.

Remark 3.2 Following Kim, Chang, and Choi(2001), we may verify the admissibility of $\delta_{a,a}^{HB}$ with $0\le a<pq-k+2$. But this will not be given here.

4. Numerical Comparison of Risks

In this section, we compare numerically the risks of best equivariant estimator δ_0 and $\delta_{a,a}^{HB}$ with $0\le a<pq-k+2$.

First note that the risk of δ_0 is $R(\sigma^2,\delta_0)=\dfrac{2}{pq(n-1)+2}$. Next following Datta and Ghosh(1995), we can get, after lengthy calculations, an exact expression for the risk difference of δ_0 and $\delta_{a,a}^{HB}$ under the loss (1.1) as follows:

$$R(\sigma^2,\delta_0)-R(\sigma^2,\delta_{a,a}^{HB})=E[(pqn-k+2L)(pqn-k+2L+2)A_{L,a}], \tag{4.1}$$

where
$$A_{L,a} = E\Big[\frac{(1-V)\psi_a(V)}{(pq(n-1)+2)(pqn-k-a+4)} \times (2L+a-2)\Big(\frac{2}{pqn-k+2L+2} - \frac{1-V}{pq(n-1)+2}\Big)\Big|L\Big]$$

with $\psi_a(V) = \frac{2}{pqn-k-a+4}\frac{V^{\frac{pq-k-b}{2}}(1-V)^{\frac{pq(n-1)+2}{2}}}{\int_0^V z^{\frac{pq-k-a}{2}}(1-z)^{\frac{pq(n-1)+2}{2}}dz}$

and $V = \frac{W}{1+W}$ for $W = \frac{n\,\mathbf{y}^T(I_{pq}-X(X^TX)^{-1}X^T)\mathbf{y}}{S}$. Here $L \sim Poisson(\eta)$ with $\eta = \frac{n\boldsymbol{\theta}^T(I_{pq}-X(X^TX)^{-1}X^T)\boldsymbol{\theta}}{2\sigma^2}$ and $V|L \sim$ Beta $(\frac{pq-k+2L}{2}, \frac{pq(n-1)}{2})$. The details for deriving (4.1) will not given here.

Based on (4.1) Figure 1 gives a numerical comparison of the risks of δ_0 and $\delta_{a,a}^{HB}$ with $a=1,2,3,4,5$ for $p=4$, $q=3$, $n=2$ when we take $X=(I_p \otimes \mathbf{1}_q,\ \mathbf{1}_p \otimes (I_{q-1}, -\mathbf{1}_{q-1})^T)$.

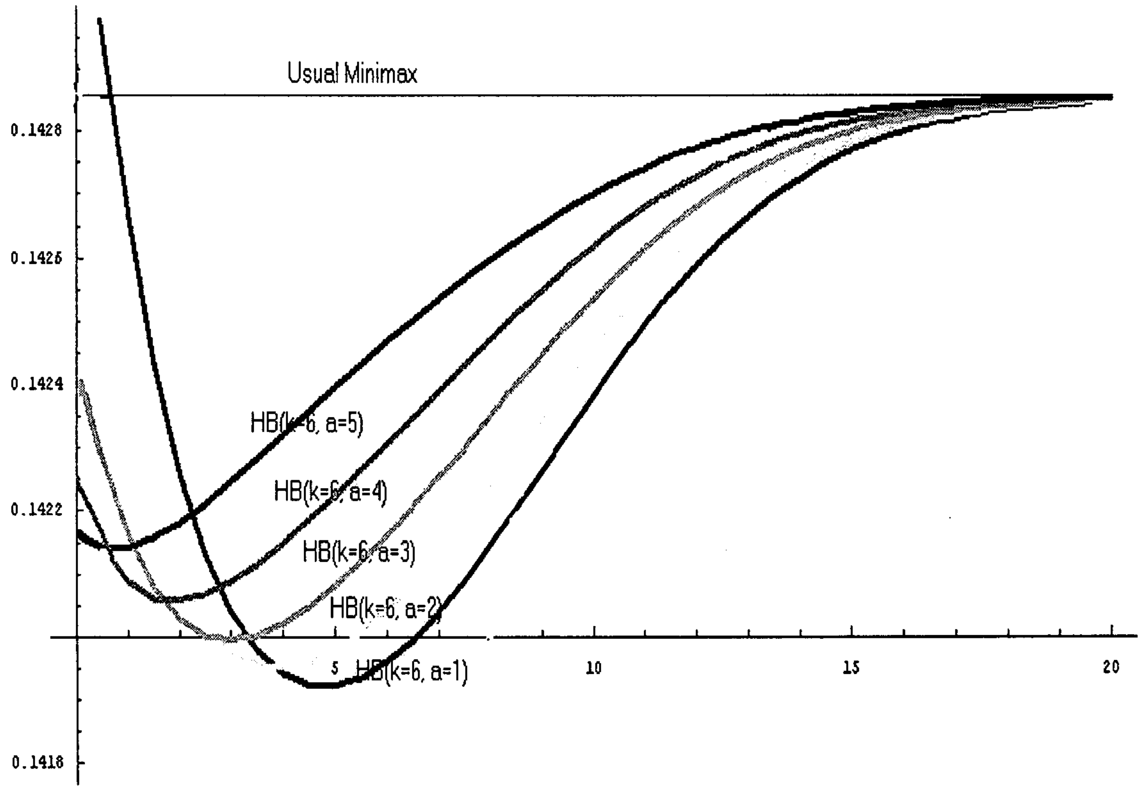

Figure 1. Comparison of risks of δ_0 and $\delta_{a,a}^{HB}$ with $a=1,2,3,4,5$

In this case, $k=p+q-1=6$, $W=2\sum_{i=1}^{4}\sum_{j=1}^{3}(y_{ij}-\overline{y_{i.}}-\overline{y_{.j}}+\overline{y_{..}})^2/S$ and $\eta=\sum_{i=1}^{4}\sum_{j=1}^{3}(\theta_{ij}-\overline{\theta_{i.}}-\overline{\theta_{.j}}+\overline{\theta_{..}})^2/\sigma^2$, where $\overline{y_{i.}}=\frac{1}{3}\sum_{j=1}^{3}y_{ij}$, $\overline{y_{.j}}=\frac{1}{4}\sum_{i=1}^{4}y_{ij}$, $\overline{y_{..}}=\frac{1}{12}\sum_{i=1}^{4}\sum_{j=1}^{3}y_{ij}$, $\overline{\theta_{i.}}=\frac{1}{3}\sum_{j=1}^{3}\theta_{ij}$, $\overline{\theta_{.j}}=\frac{1}{4}\sum_{i=1}^{4}\theta_{ij}$, and $\overline{\theta_{..}}=\frac{1}{12}\sum_{i=1}^{4}\sum_{j=1}^{3}\theta_{ij}$. As found in (4.1), the risk difference depends only on η. Hence the plot uses η in the horizontal axis.

As proved in Theorem 3.2, $\delta_{1,1}^{HB}$ is not minimax, i.e., it is possible to have negative risk improvement over δ_0 (since $a<2$). However, in spite of $\delta_{1,1}^{HB}$, it can perform much better than the minimax estimators($a=2,3,4,5$) for a wide range of values of η. Also Figure 1 confirms the fact that $\delta_{2,2}^{HB}=\delta^{BZ}$ does not improve on δ_0 at $\eta=0$. Finally, when η goes to infinity, all the hierarchical Bayes estimators seem to be very close.

References

[1] Abramowitz M, Stegun I. A.(1964). *Handbook of Mathematical Functions.* Dover Publications, New York.

[2] Brewster, J. F. and Zidek, J. V.(1973). Improving on Equivariant Estimators. *The Annals of Statistics,* 2, No. 1, 21-38.

[3] Datta, G. S. and Ghosh, M.(1995). Hierarchical Bayes Estimators of the Error Variance in One-Way ANOVA Models. *Journal of Statistical Planning and Inference,* 45, 399-411.

[4] Kim, B. H. Chang, I. H., and Choi, K. J.(2001). On the Admissibility of Hierarchical Bayes Estimators. *Submitted to Statistics.*

[5] Kubokawa, T.(1994). A Unified Approach to Improving Equivariant Estimators. *The Annals of Statistics,* 22, No. 1, 290-299.

[6] Maruyama, Y.(1998). Minimax Estimators of a Normal Variance. *Metrika* 48, 209-214

[7] Portony, S.(1971). Formal Bayes Estimation with Application to a Random Effects Model. *The Annals of Mathematical Statistics,* 42, 1379-1402.

[8] Proskin, H. M.(1985). An Admissibility Theorem with Applications to the Estimation of the Variance of the Normal Distribution. *Ph. D. dissertation, Department of Statistics,* Rutgers University.

[9] Rukhin, A. L.(1992). Asymptotic Risk Behavior of Mean Vector and Variance Estimators and the Problem of Positive Mean. *Annals of the Institute of Statistical Mathematics*, 44, 299-311.

[10] Stein, C.(1964). Inadmissibility of the Usual Estimator for the Variance of a Normal Distribution with Unknown Mean. *Annals of the Institute of Statistical Mathematics,* 16, 155-160.

[2002년 4월 접수, 2002년 7월 채택]