The Korean Communications
in Statistics Vol. 6, No. 2, 1999
pp. 501-509

A Comparative Study on Bayes Estimators for the Multivariate Normal Mean

Dal Ho Kim[1], In Suk Lee[2] and Hyun Sook Kim[3]

Abstract

In this paper, we consider a comparable study on three Bayes procedures for the multivariate normal mean estimation problem. In specific, we consider hierarchical Bayes, empirical Bayes, and robust Bayes estimators for the normal means. Then three procedures are compared in terms of the four comparison criteria (i.e., Average Relative Bias (ARB), Average Squared Relative Bias (ASRB), Average Absolute Bias (AAB), Average Squared Deviation (ASD)) using the real data set.

1. Introduction

It is apparent that both the empirical Bayes (EB) and the hierarchical Bayes (HB) procedures recognize the uncertainty in the prior information. Whereas the HB procedure models the uncertainty in the prior information by assigning a distribution (often noninformative or improper) to the hyperparameters, the EB procedure attempts to estimate the unknown hyperparameters, typically by some classical method such as method of moments, method of maximum likelihood, etc., and use the resulting estimated priors of the parameters for inferential purposes.

In the context of point estimation, both methods often lead to comparable results. However, when it comes to the question of measuring the standard errors associated with those estimators, the HB method has a clear edge over a naive EB method. EB theory by itself does not indicate how to incorporate the hyperparameter estimation error in the analysis. The HB analysis incorporates such errors automatically and hence is generally more reasonable of the approaches. Also, there are no clear cut measures of standard errors associated with EB point estimators. But the same is not true with HB estimators. To be specific, if one estimates the parameter of interest by its posterior mean, then a very natural estimate of the risk associated with this estimator is its posterior variance. Estimates of the standard errors associated with EB point estimators usually need an ingenious approximation (see, e.g., Morris (1983a,b)), whereas the posterior variances associated with the HB estimators, though often complicated, can be found exactly. For more detailed discussion, one may refer to Lindley and

1) Assistant Professor, Department of Statistics, Kyungpook National University, Taegu, 702-701, Korea.
2) Professor, Department of Statistics, Kyungpook National University, Taegu 702-701, Korea.
3) Department of Statistics, Kyungpook National University, Taegu, 702-701, Korea.

Smith (1972), Casella (1985), Berger (1985) and Ghosh (1992).

The subjective Bayesian approach has been frequently criticized on the ground that it presumes an ability to completely and accurately quantify subjective information in terms of a single prior distribution. Given the common and unavoidable practical limitations on factors such as available prior elicitation techniques and time, it is rather unrealistic to be able to quantify prior information in terms of one distribution with complete accuracy. In view of this difficulty in prior elicitation, there has long been a robust Bayesian (RB) view point that assumes only that subjective information can be quantified only in terms of a class Γ of prior distributions. These general ideas can be found in Good (1965), and more recently in Berger (1984, 1985, 1990).

In this paper, we consider the comparable HB, EB, and RB estimators for the normal means. (See, e.g., Casella and Berger (1990), p.508). Then three procedures are compared in terms of the four comparison criteria (i.e., Average Relative Bias (ARB), Average Squared Relative Bias (ASRB), Average Absolute Bias (AAB), Average Squared Deviation (ASD)) using the famous baseball data of Efron and Morris (1975).

The outline of the remaining sections is as follows. In Section 2, we review the HB and EB estimators for the multivariate normal mean.

In Section 3, we provide comparable RB estimators for the normal means using the ε -contamination class of priors where the contamination class includes all unimodal distributions..

In Section 4, we illustrate the three procedures obtained from Section 2 and Section 3 with baseball data, and we compare these estimators in terms of the four comparison criteria.

2. HB and EB Estimators

In order to derive HB estimators for the normal means estimation problem, we consider the following hierarchical model.

I. Conditional on $\theta_1,\cdots,\theta_k,$ μ and τ^2, let $Y_1,\cdots,Y_k$ be independently distributed with $Y_i \sim \mathrm{N}(\theta_i, \sigma^2)$, $i=1,\cdots,k$, where the σ^2 is known positive constant;

II. Conditional on μ and τ^2, $\theta_1,\cdots,\theta_k$ are independently distributed with $\theta_i \sim \mathrm{N}(\mu,\tau^2)$, $i=1,\cdots,k$;

III. Marginally μ and τ^2 are independent with $\mu \sim$ uniform$(-\infty,\infty)$ and $\tau^2 \sim$ uniform$(0,\infty)$.

We shall use the notations $\boldsymbol{y}=(y_1,\cdots,y_k)^T$ and $\boldsymbol{\theta}=(\theta_1,\cdots,\theta_k)^T$.

Then the joint pdf of $\boldsymbol{y}$, $\boldsymbol{\theta}$, μ and τ^2 is given by

$$p(\boldsymbol{y},\boldsymbol{\theta},\mu,\tau^2) \propto (\tau^2)^{-\frac{k}{2}} \exp\left[-\frac{1}{2}\left\{\frac{1}{\sigma^2}\sum_{i=1}^{k}(y_i-\theta_i)^2+\frac{1}{\tau^2}\sum_{i=1}^{k}(\theta_i-\mu)^2\right\}\right]. \tag{2.1}$$

Now, integrating with respect to μ, it follows from (2.1) that the joint pdf of $\boldsymbol{y}$, $\boldsymbol{\theta}$ and τ^2 is given by

$$p(\boldsymbol{y},\boldsymbol{\theta},\tau^2) \propto (\tau^2)^{-\frac{(k-1)}{2}} \exp\left[-\frac{1}{2\sigma^2}(\boldsymbol{\theta}-\boldsymbol{D}^{-1}\boldsymbol{y})^T\boldsymbol{D}(\boldsymbol{\theta}-\boldsymbol{D}^{-1}\boldsymbol{y})\right] \times \exp\left[-\frac{B}{2\sigma^2}\sum_{i=1}^{k}(y_i-\bar{y})^2\right] \tag{2.2}$$

where $D=\boldsymbol{I}_k+\frac{\sigma^2}{\tau^2}(\boldsymbol{I}_k-\frac{\boldsymbol{J}_k}{k})$, $\boldsymbol{D}^{-1}=(1-B)\boldsymbol{I}_k+Bk^{-1}\boldsymbol{J}_k$, and $B=\sigma^2/(\sigma^2+\tau^2)$. Here we use usual notations $\mathbf{1}_k=(1,\cdots,1)^T$, $\boldsymbol{I}_k=$ Diag$\{1,\cdots,1\}$ and $\boldsymbol{J}_k=\mathbf{1}_k\mathbf{1}_k^T$. Then the posterior distribution of $\boldsymbol{\theta}$ given $\boldsymbol{y}$ and τ^2 is a k-variate normal distribution

$$\boldsymbol{N}_k[(1-B)y+B\bar{y}\mathbf{1}_k,\ \sigma^2(1-B)\boldsymbol{I}_k+\sigma^2Bk^{-1}\boldsymbol{J}_k)]. \tag{2.3}$$

Also, integrating with respect to $\boldsymbol{\theta}$ in (2.2), one gets the joint pdf of $\boldsymbol{y}$ and τ^2 given by

$$p(\boldsymbol{y},\tau^2) \propto B^{\frac{k-1}{2}}\exp\left(-\frac{B}{2}\sigma^2\sum_{i=1}^{k}(y_i-\bar{y})^2\right). \tag{2.4}$$

Since $B=\sigma^2/(\sigma^2+\tau^2)$, it follows from (2.4) that

$$p(B|\boldsymbol{y}) \propto B^{\frac{k-5}{2}}\exp\left(-\frac{B}{2\sigma^2}\sum_{i=1}^{k}(y_i-\bar{y})^2\right).$$

Hence, the posterior mean and variance are given by

$$E(\boldsymbol{\theta}|\boldsymbol{y})=\boldsymbol{y}-E(B|\boldsymbol{y})(\boldsymbol{y}-\bar{y}\mathbf{1}_k) \tag{2.5}$$

and

$$Var(\boldsymbol{\theta}|\boldsymbol{y})=\sigma^2\boldsymbol{I}_k-\sigma^2E(B|\boldsymbol{y})(\boldsymbol{I}_k-k^{-1}\boldsymbol{J}_k)+Var(B|\boldsymbol{y})(\boldsymbol{y}-\bar{y}\mathbf{1}_k)(\boldsymbol{y}-\bar{y}\mathbf{1}_k)^T \tag{2.6}$$

where

$$E(B|\boldsymbol{y})=\frac{\int_0^1 B^{\frac{k-3}{2}}\exp\left(-\frac{B}{2\sigma^2}\sum_{i=1}^{k}(y_i-\bar{y})^2\right)dB}{\int_0^1 B^{\frac{k-5}{2}}\exp\left(-\frac{B}{2\sigma^2}\sum_{i=1}^{k}(y_i-\bar{y})^2\right)dB}, \tag{2.7}$$

$$E(B^2|\boldsymbol{y})=\frac{\int_0^1 B^{\frac{k-1}{2}}\exp\left(-\frac{B}{2\sigma^2}\sum_{i=1}^{k}(y_i-\bar{y})^2\right)dB}{\int_0^1 B^{\frac{k-5}{2}}\exp\left(-\frac{B}{2\sigma^2}\sum_{i=1}^{k}(y_i-\bar{y})^2\right)dB}, \tag{2.8}$$

and

$$Var(B|\boldsymbol{y})=E(B^2|\boldsymbol{y})-(E(B|\boldsymbol{y}))^2. \tag{2.9}$$

In order to develop a closed form solution to HB point estimates and posterior variances given by (2.5) and (2.6), Morris (1983b) suggested approximations to (2.7) and (2.8) involving replacement of the integral over B in (0, 1) by the integral over B in $(0,\infty)$. But we will compute (2.7) and (2.8) using Gauss-Legendre quadrature for an illustration in Section 3.

For the EB analysis, we consider the same model, except that the hyperprior distributions are not considered. Instead, we estimate the hyperparameters μ and τ^2 based on the marginal distribution of $\boldsymbol{y}$.

To derive the posterior distribution of θ for given $\boldsymbol{y}$, we start with the joint pdf of $\boldsymbol{y}$ and $\boldsymbol{\theta}$

given by

$$p(\mathbf{y},\boldsymbol{\theta}) \propto \prod_{i=1}^{k}\left\{\exp\left[\frac{-(y_i-\theta_i)^2}{2\sigma^2}\right]\frac{1}{\sqrt{\tau^2}}\exp\left[\frac{-(\theta_i-\mu)^2}{2\tau^2}\right]\right\}. \tag{2.10}$$

Using $B=\sigma^2/(\sigma^2+\tau^2)$, the usual square completion technique on the sum of the two exponents in (2.10) leads to

$$p(\mathbf{y},\boldsymbol{\theta}) \propto \prod_{i=1}^{k}\frac{1}{\sqrt{\sigma^2(1-B)}\,,\sqrt{\sigma^2+\tau^2}}\exp\left[-\frac{(\theta_i-(1-B)\,y_i-B\mu)^2}{2\sigma^2(1-B)}\right] \times \exp\left[\frac{-(y_i-\mu)^2}{2(\sigma^2+\tau^{2)}}\right]. \tag{2.11}$$

From (2.11) the posterior distribution of $\boldsymbol{\theta}$ given $\mathbf{y}$ is then

$$\mathbf{N}_k[\,(1-B)\,\mathbf{y}+B\mu\;\mathbf{1}_k,\;\sigma^2(1-B)\;\mathbf{I}_k)]. \tag{2.12}$$

Also, the marginal density of y_i is derived by

$$m(y_i)=\frac{1}{\sqrt{2\pi(\sigma^2+\tau^2)}}\exp\left[\frac{-(y_i-\mu)^2}{2(\sigma^2+\tau^2)}\right].$$

Thus the joint marginal density of $\mathbf{y}$ is given by $\mathbf{N}_k(\mu\;\mathbf{1}_k,\;\sigma^2B^{-1}\;\mathbf{I}_k)$. Hence, we can estimate μ and τ^2 based on $m(\,\mathbf{y}|\mu,\tau^2)$.

One could then pretend that the θ_i are independently $N(\widehat{\mu},\widehat{\tau^2})$, and proceed with a Bayesian analysis. This indeed works well k is large. For small or moderate k, however, such an analysis leaves something to be desired, since it ignores the fact that μ and τ^2 were estimated. The errors undoubtedly introduced in the hyperparameter estimation will not be reflected in any of the conclusions. This is indeed a general problem with the EB Approach, and will leads us to recommend the HB approach when k is small or moderate.

However, Morris (1983a,b) has developed EB approximations to the HB answers which do take into account the uncertainty in $\widehat{\mu}$ and $\widehat{\tau^2}$. These approximations can best be described in EB terms, as providing an estimated posterior distribution.

The estimates Morris (1983a) suggests for $E(\theta_i|y)$ and $Var(\theta_i|y)$ are (when $k\geq 4$)

$$E(\theta_i|\,\mathbf{y})=(1-\widehat{B})y_i+\widehat{B}\,\widehat{\mu} \tag{2.13}$$

and

$$Var(\theta_i|\,\mathbf{y})=\sigma^2(1-\frac{(k-1)}{k}\,\widehat{B})+\frac{2}{(k-3)}\,\widehat{B}^2\,(y_i-\bar{y})^2 \tag{2.14}$$

where

$$\widehat{\mu}=\bar{y}$$

$$\widehat{B}=(\frac{k-3}{k-1})\frac{\sigma^2}{\sigma^2+\widehat{\tau^2}} \tag{2.15}$$

$$\widehat{\tau^2}=max\left\{0,\;\frac{1}{(k-1)}\,S^2-\sigma^2\right\} \tag{2.16}$$

and

$$S^2 = \sum_{i=1}^{k}(y_i - \bar{y})^2.$$

In way of explanation, $\hat{\tau}^2$ is just a slight modification of τ^2 in (2.16) and the factor $(k-3)/(k-1)$ in (2.15) has to do with adjusting for the error in the estimation of τ^2, and $(k-1)/k$ in (2.14) and the last term in (2.14) have to do with the error in estimating μ. We will compute (2.13) and (2.14) for the illustration in Section 3.

3. RB Estimators

The hierarchical model here is similar to the one in Section 2, except that now the distribution of τ^2 has the ε-contamination class of priors (c.f. Berger and Berliner (1986)), namely

$$\Gamma = \{h : h = (1-\varepsilon)h_0 + \varepsilon q, \ q \in Q\} \tag{3.1}$$

where, $0 \le \varepsilon \le 1$ is given, h_0 is the inverse gamma distribution with pdf

$$h_0(\tau^2) = \frac{\alpha_0^{\beta_0}}{\Gamma(\beta_0)} \frac{1}{(\tau^2)^{\beta_0+1}} e^{-\frac{\alpha_0}{\tau^2}} I_{(0,\infty)}(\tau^2),$$

denoted by $\mathrm{IG}(\alpha_0, \beta_0)$, and Q is the class of all unimodal distributions with the same mode τ_0^2 as that of h_0.

Let $f(\mathbf{y}|\tau^2) = \int\int p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta}|\mu, \tau^2)p(\mu)d\mu d\boldsymbol{\theta}$. We denote by $m(\mathbf{y}|h)$ the marginal distribution of $\mathbf{y}$ with respect to the prior h, namely

$$m(\mathbf{y}|h) = \int f(\mathbf{y}|\tau^2)h(d\tau^2).$$

For $h \in \Gamma$, we get

$$m(\mathbf{y}|h) = (1-\varepsilon)m(\mathbf{y}|h_0) + \varepsilon m(\mathbf{y}|q).$$

Our objective is to choose the ML-II prior $\hat{h}$ which maximizes $m(\mathbf{y}|h)$ over Γ. This amounts to maximization of $m(\mathbf{y}|q)$ over $q \in Q$. Using the representation of each $q \in Q$ as a mixture of uniform densities, the ML-II prior is given by

$$\hat{h}(\tau^2) = (1-\varepsilon)h_0(\tau^2) + \varepsilon\,\hat{q}(\tau^2) \tag{3.2}$$

where $\hat{q}$ is uniform $(\tau_0^2,\ \tau_0^2 + \hat{z})$, $\hat{z}$ being the solution of the equation

$$f(\mathbf{y}|z) = \frac{1}{z}\int_{\tau_0^2}^{\tau_0^2+z} f(\mathbf{y}|\tau^2)d\tau^2 \tag{3.3}$$

and τ_0^2 is the unique mode of $h_0(\tau^2)$ $(i.e., \tau_0^2 = \alpha_0/(\beta_0+1))$. In fact,

$$f(\mathbf{y}|\tau^2) \propto (\tau^2 + \sigma^2)^{-\frac{k-1}{2}} \exp\left(-\frac{1}{2}\frac{1}{\tau^2+\sigma^2}\sum_{i=1}^{k}(y_i - \bar{y})^2\right). \tag{3.4}$$

Now, we find the solution $\hat{z}$ to the equation

$$f(\boldsymbol{y}|z) = \frac{1}{z}\int_{\tau_0^2}^{\tau_0^2+z} f(\boldsymbol{y}|\tau^2)d\tau^2,$$

and so

$$f(\boldsymbol{y}|z) + z\frac{d}{dz}f(\boldsymbol{y}|z) = f(\boldsymbol{y}|\tau_0^2+z). \tag{3.5}$$

Here,

$$\frac{d}{dz}f(\boldsymbol{y}|z) = (z+\sigma^2)^{-\frac{k-1}{2}}\left(-\frac{k-1}{2z} + \sigma^2 + \frac{\sum_{i=1}^{k}(y_i-\bar{y})^2}{2(z+\sigma^2)^2}\right)\exp\left(-\frac{1}{z+\sigma^2}\frac{\sum_{i=1}^{k}(y_i-\bar{y})^2}{2(z+\sigma^2)}\right).$$

Hence, from (3.5) one gets

$$(z+\sigma^2)^{-\frac{k-1}{2}}\left(1-\frac{k-1}{2}\frac{z}{z+\sigma^2} + \frac{z\sum_{i=1}^{k}(y_i-\bar{y})^2}{2(z+\sigma^2)^2}\right)\exp\left(-\frac{1}{z+\sigma^2}\frac{\sum_{i=1}^{k}(y_i-\bar{y})^2}{2(z+\sigma^2)}\right)$$

$$= (z+\tau_0^2+\sigma^2)^{-\frac{k-1}{2}}\exp\left(-\frac{1}{z+\sigma^2}\frac{\sum_{i=1}^{k}(y_i-\bar{y})^2}{2(z+\tau_0^2+\sigma^2)}\right). \tag{3.6}$$

So we can find the solution $\hat{z}$ to the equation (3.6).

Then we have the following HB model based on ML-II prior:

I. Conditional on $\theta_1, \cdots, \theta_k$, μ and τ^2

$$y_i \overset{ind}{\sim} \mathrm{N}(\theta_i, \sigma^2) \qquad (i=1,\cdots,k)$$

II. Conditional on μ and τ^2,

$$\theta_i \overset{ind}{\sim} \mathrm{N}(\mu, \tau^2) \qquad (i=1,\cdots,k)$$

III. Marginally μ and τ^2 are independent with $\mu \sim$ uniform$(-\infty,\infty)$ and $\hat{h}(\tau^2) = (1-\varepsilon)h_0(\tau^2) + \varepsilon\hat{q}(\tau^2)$. where $\hat{q}(\tau^2) \sim$ uniform$(\tau_0^2, \tau_0^2+\hat{z})$.

From Section 2, recall that the posterior distribution of $\boldsymbol{\theta}$ given $\boldsymbol{y}$ and τ^2 is

$$\boldsymbol{N}_k((1-B)\,\boldsymbol{y}+B\bar{y}\,\boldsymbol{1}_k,\ \sigma^2(1-B)\,\boldsymbol{I}_k+\sigma^2Bk^{-1}\,\boldsymbol{J}_k).$$

Also, recall that the posterior mean and the posterior variance are given by

$$E(\boldsymbol{\theta}|\,\boldsymbol{y}) = \boldsymbol{y} - E(B|\,\boldsymbol{y})(\boldsymbol{y}-\bar{y}\,\boldsymbol{1}_k)$$

and

$$Var(\boldsymbol{\theta}|\,\boldsymbol{y}) = \sigma^2(\,\boldsymbol{I}_k - E(B|\,\boldsymbol{y})(\,\boldsymbol{I}_k-k^{-1}\,\boldsymbol{J}_k)) + Var(B|\,\boldsymbol{y})(\boldsymbol{y}-\bar{y}\,\boldsymbol{1}_k)(\boldsymbol{y}-\bar{y}\,\boldsymbol{1}_k)^T.$$

But in the RB setting we need

$$\begin{aligned} h(\tau^2|\,\boldsymbol{y}) &\propto (\tau^2+\sigma^2)^{-\frac{k-1}{2}}\exp\left(-\frac{1}{2}\frac{1}{\tau^2+\sigma^2}\sum_{i=1}^{k}(y_i-\bar{y})^2\right)[(1-\varepsilon)h_0(\tau^2)+\varepsilon\hat{q}(\tau^2)] \\ &\propto (1-\varepsilon)\frac{\alpha_0^{\beta_0}}{\Gamma(\beta_0)}\frac{1}{(\tau^2)^{\beta_0+1}}(\tau^2+\sigma^2)^{-\frac{k-1}{2}}\exp\left(-\frac{1}{2}\frac{1}{\tau^2+\sigma^2}\sum_{i=1}^{k}(y_i-\bar{y})^2-\frac{\alpha_0}{\tau^2}\right) \\ &\quad + \frac{\varepsilon}{\hat{z}}(\tau^2+\sigma^2)^{-\frac{k-1}{2}}\exp\left(-\frac{1}{2}\frac{1}{\tau^2+\sigma^2}\sum_{i=1}^{k}(y_i-\bar{y})^2\right) \end{aligned} \tag{3.7}$$

in order to calculate the above posterior mean and variance. Since $B = \sigma^2/(\sigma^2+\tau^2)$ $(0 < B < 1)$, it follows from (3.7) that

$$
\begin{aligned}
h(B|\mathbf{y}) \propto\ & B^{-2}(1-\varepsilon)\frac{\alpha_0^{\beta_0}}{\Gamma(\beta_0)}\ \frac{1}{(\sigma^2/B-\sigma^2)^{\beta_0+1}}\ (\sigma^2/B)^{-\frac{k-1}{2}} \\
& \times \exp\left(-\frac{1}{2}\ \frac{1}{\sigma^{2}/B}\ \sum_{i=1}^{k}(y_i-\bar{y})^2 - \frac{\alpha_0}{\sigma^2/B}\right) \\
& + B^{-2}\frac{\varepsilon}{\hat{z}}\ (\sigma^2/B)^{-\frac{k-1}{2}}\exp\left(-\frac{1}{2}\ \frac{1}{\sigma^2/B}\ \sum_{i=1}^{k}(y_i-\bar{y})^2\right).
\end{aligned}
\tag{3.8}
$$

Then the above posterior mean and variance can be obtained by computing $E(B|\mathbf{y})$ and $Var(B|\mathbf{y})$ based on (3.8). This is also one-dimensional integration problem and so we will use Gauss-Legendre quadrature for an illustration in Section 3 with baseball data.

4. An Example

In this section, we will compare all the estimators discussed in Section 2 and Section 3 using the famous baseball data of Efron and Morris (1975). The data consist of the batting averages of 18 major league players through their first 45 official at bats of the 1970 season. Here the goal is to find which procedure gives closest estimates to the 'true' probability of a hit, θ_i, for each player's batting average over the remainder of the season using the data.

To begin, the binomial model is fitted to the 45 at-bats for each player i, but, to simplify the computations, the binomial likelihood is approximated by a normal likelihood under the arcsine transformation. The arcsin transformation is used to create normal random variables of roughly unit variance for all values of the parameter θ_i that are not close to 0 or 1. That is, the normal approximations for the binomial proportion under arcsin transformation give

$$f(\mathbf{y}|\boldsymbol{\theta}) = \prod_{i=1}^{18} \mathrm{N}(y_i|\theta_i, 1).$$

The HB, EB and RB estimates and their posterior standard deviations are obtained under the normal model. In finding RB estimates, we have tried several cases for α_0, β_0 and ε and have reported the result using $\alpha_0=10$, $\beta_0=1$, $\varepsilon=.1$ for our convinience. Table 3.1 provides the "truth" and three Bayes estimates and the standard errors associated with three Bayes estimates.

Table 3.1 Estimates of θ_i (standard error)

i	θ_i	$\hat{\theta}_{i,\mathrm{HB}}(s_{i,\mathrm{HB}})$	$\hat{\theta}_{i,\mathrm{EB}}(s_{i,\mathrm{EB}})$	$\hat{\theta}_{i,\mathrm{RB}}(s_{i,\mathrm{RB}})$
1	.346	.306(.0733)	.290(.0737)	.309(.0735)
2	.298	.299(.0731)	.286(.0732)	.302(.0733)
3	.276	.292(.0729)	.282(.0728)	.295(.0731)
4	.222	.285(.0728)	.277(.0725)	.287(.0729)
5	.273	.278(.0727)	.273(.0723)	.279(.0728)
6	.270	.278(.0727)	.273(.0723)	.279(.0728)
7	.263	.271(.0726)	.268(.0721)	.272(.0728)
8	.210	.264(.0726)	.264(.0720)	.264(.0727)
9	.269	.257(.0726)	.259(.0721)	.256(.0727)
10	.230	.257(.0726)	.259(.0721)	.256(.0727)
11	.264	.249(.0727)	.254(.0722)	.248(.0728)
12	.256	.249(.0727)	.254(.0722)	.248(.0728)
13	.303	.249(.0727)	.254(.0722)	.248(.0728)
14	.264	.249(.0727)	.254(.0722)	.248(.0728)
15	.226	.249(.0727)	.254(.0722)	.248(.0728)
16	.285	.242(.0728)	.249(.0725)	.240(.0729)
17	.316	.234(.0729)	.244(.0728)	.232(.0731)
18	.200	.226(.0731)	.239(.0734)	.223(.0734)

Now, we use the following four criteria to compare the different estimates. Let θ_i denote the baseball players' actual batting average for the remainder of the season. For any estimate $e=(e_1,\cdots,e_{18})^T$, we compute the following four criteria:

$$\text{average relative bias} = (18)^{-1}\sum_{i=1}^{18}|\theta_i - e_i|/\theta_i$$

$$\text{average squared relative bias} = (18)^{-1}\sum_{i=1}^{18}|\theta_i - e_i|^2/\ \theta_i^2$$

$$\text{average absolute bias} = (18)^{-1}\sum_{i=1}^{18}|\theta_i - e_i|$$

$$\text{average squared deviation} = (18)^{-1}\sum_{i=1}^{18}(\theta_i - e_i)^2$$

From Table 3.1, we can see that three Bayes estimates are quite similar to each other. But Table 3.2 indicates that the EB estimates seems to be the best, but there is no second best under all criteria for this data set. On the whole the differences are quite small. So we can say that all three procedures are quite comparable.

Table 3.2 A Comparison of Estimates For θ_i Under Four Different Criteria

Estimate	Average relative bias	Average squared relative bias	Average absolute bias	Average squared deviation
HB	.1067394	.0187807	.0277576	.0012784
EB	.1019850	.0178852	.0263850	.0011957
RB	.1088814	.0192212	.0283835	.0013140

References

[1] Berger, J. O. (1984). The robust Bayesian viewpoint (with discussion), in *Robustness of Bayesian Analysis*, Ed. J. Kadane. North-Holland, Amsterdam, 63-124.

[2] Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*. (2nd edn.). Springer-Verlag, New York.

[3] Berger, J. O. (1990). Robust Bayesian analysis: Sensitivity to the prior. *Journal of Statistical Planning and Inference*, 25, 303-328.

[4] Berger, J. O. and Berliner, L. M. (1986). Robust Bayes and empirical Bayes analysis with ε-contaminated priors. *The Annals of Statistics*, 14, 461-486.

[5] Casella, G. (1985). An introduction to empirical Bayes data analysis. *American Statistician*, 39, 83-87.

[6] Casella, G. and Berger, R. L. (1990). *Statistical Inference*. Wadsworth, Belmont.

[7] Efron, B. and Morris, C. (1975). Data analysis using Stein's estimator and its generalizations. *Journal of American Statistical Association*, 70, 311-319.

[8] Ghosh, M. (1992). Hierarchical and empirical Bayes multivariate estimation, in *Current Issues in Statistical Inference: Essays in Honor of D. Basu*, Eds. M. Ghosh and P.K. Pathak. IMS Lecture Notes-Monographic Series, 151-177.

[9] Good, I. J. (1965). *The Estimation of Probabilities: An Essay on Modern Bayesian Methods*. Cambridge, Massachusetts: M.I.T press.

[10] Lindley, D. and Smith, A. F. M. (1972). Bayes estimates for the linear model. Journal of Royal Statistical Society, Ser. B, 34, 1-41.

[11] Morris, C. N. (1983a). Parametric empirical Bayes inference: Theory and applications (with comments). *Journal of the American Statistical Association*, 78, 47-65.

[12] Morris, C. N. (1983b). Parametric empirical Bayes confidence intervals, in *Scientific Inference, Data Analysis, and Robustness*, Academic Press, New York, 25-50.