

## Bayesian Estimation Using Noninformative Priors in Hierarchical Model

Dal Ho Kim<sup>1)</sup> · Jin Kap Choi<sup>2)</sup> · Hee Jo Choi<sup>3)</sup>

### Abstract

We consider the simultaneous Bayesian estimation for the normal means based on different noninformative type hyperpriors in hierarchical model. We provide numerical example using the famous baseball data in Efron and Morris (1975) for illustration.

**Keywords** : Bayesian estimation, hierarchical model, matching priors, normal means

### 1. Introduction

Bayesian methods have become increasingly popular in the theory and practice of statistics. This is partly due to the fact that even with little or no prior information, one can often employ noninformative priors to draw reliable inference. In practice, empirical and hierarchical Bayes methods are useful, especially in the context of simultaneous estimation of several parameters.

For example, agencies of the federal government have been involved in obtaining estimates of per capita income, unemployment rates, crop yields and so forth simultaneously for several state and local government areas. In such situations, quite often estimates of certain area means, or simultaneous estimates of several area means can be improved by incorporating information from similar neighboring areas. Examples of this type are especially suitable for empirical Bayes (EB) and hierarchical Bayes (HB) analyses.

EB and HB methods are being routinely used whenever there is a need to

---

1) First Author : Associate Professor, Department of Statistics, Kyungpook National University, Taegu, 702-701, Korea  
E-Mail : dalkim@knu.ac.kr

2) Professor, Department of Statistics, Kyungpook National University, Taegu, 702-701, Korea

3) M.S., Department of Statistics, Kyungpook National University, Taegu, 702-701, Korea

"borrow strength" either for inference related to a particular parameter of interest, or simultaneous inference for several parameters. In particular HB methods are gaining increasing popularity in recent years partly due to overcoming the difficulty in calculation even if we use vague or noninformative priors. Thus, not surprisingly, over the years, a wide range of noninformative priors has been proposed and studied. One popular criterion for the development of such priors is to match asymptotically the posterior coverage probability of a Bayesian credible set with the corresponding frequentist coverage probability.

The outline of this paper is as follows. In Section 2, we introduce a matching prior for the hyperparameter in the normal HB model following Datta, Ghosh and Mukerjee (2000). This prior turns out to be different from the one proposed by Morris(1983). In Section 3, we derive expressions for the posterior means and variances with different hyperpriors in the normal HB setup. And in Section 4, we provide numerical example using the famous data of Efron and Morris(1975) for illustration.

## 2. Matching Priors in Hierarchical Model

We consider the following normal HB model

- I.  $Y_i | \theta_i, \mu, \tau^2 \stackrel{ind}{\sim} N(\theta_i, \sigma^2), i = 1, \dots, n;$
- II.  $\theta_i | \mu, \tau^2 \stackrel{iid}{\sim} N(\mu, \tau^2), i = 1, \dots, n;$
- III.  $\pi(\mu, \tau^2) \propto \pi(\tau^2)$

where  $\sigma^2$  is assumed to be known. For the hyperparameter  $(\mu, \tau^2)$ , we assign a hyperprior  $\pi(\mu, \tau^2)$  in step (III). We assume  $\mu$  and  $\tau^2$  are independent with a uniform  $(-\infty, \infty)$  prior for  $\mu$  and a suitable prior  $\pi(\tau^2)$  (to be determined below) on  $\tau^2$ . A uniform prior on  $\mu$  is widely accepted as a reasonable objective prior. We determine the prior  $\pi(\tau^2)$  via posterior and frequentist quantile matching of  $\theta_i$ . To this end we find an asymptotic expansion of the posterior distribution of  $\theta_i$ .

Let  $\mathbf{d} = (y_1, \dots, y_n)^T$  denote the observed value of  $\mathbf{Y} = (Y_1, \dots, Y_n)^T$  and  $\hat{\tau}^2$  denote the residual maximum likelihood estimate of  $\tau^2$  obtained by maximizing the log residual likelihood. This is equivalent to minimizing  $h(\tau^2)$  where

$$nh(\tau^2) = \frac{(n-1)}{2} \log(\tau^2 + \sigma^2) + \frac{1}{2} \sum_{i=1}^n \frac{(y_i - \bar{y})^2}{\tau^2 + \sigma^2}. \quad (2.1)$$

Let  $\pi(\theta_i | \tau^2, \mathbf{d})$  denote the conditional posterior of  $\theta_i$  given  $\tau^2$ , and  $\pi_\rho(\theta_i | \mathbf{d})$  denote the posterior pdf of  $\theta_i$  under the prior  $\pi(\tau^2)$ , where  $\rho(\tau^2) = \log \pi(\tau^2)$ . By the Laplace approximation (see, for example, Kass and Steffey, 1989) it can be shown that

$$\begin{aligned} \pi_\rho(\theta_i | \mathbf{d}) = & \pi(\theta_i | \hat{\tau}^2, \mathbf{d}) + \frac{D\pi(\theta_i | \hat{\tau}^2, \mathbf{d})}{2nh_2} (2\rho_1 - h_3 h_2^{-1}) \\ & + \frac{D^2\pi(\theta_i | \hat{\tau}^2, \mathbf{d})}{2nh_2} + o(n^{-1}) \end{aligned} \quad (2.2)$$

where

$$D = \partial \tau^2 / \partial, \rho_1 = D \log \pi(\hat{\tau}^2), h_k = D^k h(\hat{\tau}^2), k = 2, 3, D^k w(\hat{\tau}^2) = D^k w(\tau^2) |_{\tau^2 = \hat{\tau}^2}.$$

Let

$$G(\theta_i^* | \tau^2, \mathbf{d}) = \int_{-\infty}^{\theta_i^*} \pi(\theta_i | \tau^2, \mathbf{d}) d\theta_i$$

be the conditional posterior cdf of  $\theta_i$ . Also define  $q_i(\tau^2, \alpha; \mathbf{d})$ , the conditional posterior quantile function by

$$G(q_i(\tau^2, \alpha; \mathbf{d}) | \tau^2, \mathbf{d}) = 1 - \alpha. \quad (2.3)$$

Let  $h_i(\pi, \hat{\tau}^2, \alpha; \mathbf{d})$  be such that

$$P^\pi\{\theta_i > h_i(\pi, \hat{\tau}^2, \alpha; \mathbf{d}) | \mathbf{d}\} = \alpha + o(n^{-1}). \quad (2.4)$$

From (2.2)-(2.4) it follows that

$$h_i(\pi, \hat{\tau}^2, \alpha; \mathbf{d}) = q_i(\hat{\tau}^2, \alpha; \mathbf{d}) + \frac{1}{n} u(\pi) \quad (2.5)$$

where  $u(\pi)$  may depend on the prior  $\pi$ , in addition to  $\alpha$  and  $\mathbf{d}$ , and is at most of the order  $O(1)$ . Let  $P_{\tau^2}(\cdot)$  denote the probability measure based on the joint

distribution of  $(Y_i, \theta_i), i = 1, \dots, n$ , as specified by (I) and (II) of the hierarchical model. We will now find an expansion of the probability  $P_{\tau^2}[\theta_i > h_i(\pi, \hat{\tau}^2, \alpha; \mathbf{Y})]$  up to the order  $o(n^{-1})$ . We develop this expansion based on a limiting argument as outlined in Ghosh (1994, p.84) or Datta and Ghosh (1995, p.40). To this end, for an alternative prior  $\bar{\pi}(\tau^2)$  we obtain an expansion of the posterior density  $\pi_{\rho}^{-}(\theta_i | \mathbf{d})$ , where  $\bar{\rho}(\tau^2) = \log \bar{\pi}(\tau^2)$ . As in (2.2), we get

$$\begin{aligned} \pi_{\rho}^{-}(\theta_i | \mathbf{d}) &= \pi(\theta_i | \hat{\tau}^2, \mathbf{d}) + \frac{D\pi(\theta_i | \hat{\tau}^2, \mathbf{d})}{2nh_2} (2\bar{\rho}_1 - h_3h_2^{-1}) \\ &\quad + \frac{D^2\pi(\theta_i | \hat{\tau}^2, \mathbf{d})}{2nh_2} + o(n^{-1}) \end{aligned} \tag{2.6}$$

From (2.2) and (2.6) it follows that

$$\pi_{\rho}^{-}(\theta_i | \mathbf{d}) = \pi_{\rho}(\theta_i | \mathbf{d}) + \frac{D\pi(\theta_i | \hat{\tau}^2, \mathbf{d})}{nh_2} (\bar{\rho}_1 - \rho_1) + o(n^{-1}). \tag{2.7}$$

From (2.4), (2.5), and (2.7) it follows that  $P_{\tau^2}^{-}\{\theta_i > h_i(\pi, \hat{\tau}^2, \alpha; \mathbf{d}) | \mathbf{d}\}$  simplifies to

$$\alpha + \frac{(\bar{\rho}_1 - \rho_1)}{nh_2} \int_{h_i(\pi, \hat{\tau}^2, \alpha, \mathbf{d})}^{\infty} D\pi(\theta_i | \hat{\tau}^2, \mathbf{d}) d\theta_i + o(n^{-1})$$

which is same as

$$\alpha + \frac{(\bar{\rho}_1 - \rho_1)}{nh_2} \int_{q_i(\hat{\tau}^2, \alpha, \mathbf{d})}^{\infty} D\pi(\theta_i | \hat{\tau}^2, \mathbf{d}) d\theta_i + o(n^{-1}).$$

Hence  $E_{\tau^2}^{-}[P_{\tau^2}^{-}\{\theta_i > h_i(\pi, \hat{\tau}^2, \alpha; \mathbf{Y}) | \mathbf{Y}\}]$  reduces to

$$\alpha + \frac{\{\bar{\rho}_1(\tau^2) - \rho_1(\tau^2)\}}{nJ(\tau^2)} E_{\tau^2}^{-}\left[\int_{q_i(\tau^2, \alpha; \mathbf{Y})}^{\infty} D\pi(\theta_i | \tau^2, \mathbf{Y}) d\theta_i\right] + o(n^{-1}) \tag{2.8}$$

since  $h_2 = J(\tau^2) + o(1)$ , where  $J(\tau^2) = (\tau^2 + \alpha^2)^{-2}/2$ .

From (I) and (II), note that  $\pi(\theta_i | \tau^2, \mathbf{d}) = N(\mu_i(\tau^2, \mathbf{d}), V(\tau^2))$ , where  $\mu_i(\tau^2, \mathbf{d}) = y_i - B(\tau^2)(y_i - \bar{y})$ ,  $V(\tau^2) = \sigma^2(1 - B(\tau^2)) + \sigma^2B(\tau^2)/n$  and

$$B(\tau^2) = \sigma^2(\sigma^2 + \tau^2)^{-1}.$$

From this it follows that  $q_i(\tau^2, \alpha, \mathbf{d}) = \mu_i(\tau^2, \mathbf{d}) + z_\alpha \sqrt{V(\tau^2)}$ , where  $z_\alpha$  is the  $1 - \alpha$ th quantile of a standard normal distribution. Hence it can be shown that  $\int_{q_i(\tau^2, \alpha; \mathbf{d})}^{\infty} D\pi(\theta_i | \tau^2, \mathbf{d}) d\theta_i$  simplifies to

$$\left[ \frac{V'(\tau^2)z_\alpha}{2V(\tau^2)} + \frac{\mu_i'(\tau^2)}{\sqrt{V(\tau^2)}} \right] \phi(z_\alpha) + o(1)$$

which in turn is given by

$$- \left[ \frac{B'(\tau^2)z_\alpha}{2\{1 - B(\tau^2)\}} + \frac{B'(\tau^2)(y_i - \bar{y})}{\sigma \sqrt{1 - B(\tau^2)}} \right] \phi(z_\alpha) + o(1), \quad (2.9)$$

where  $\phi(z)$  is the standard normal pdf. Since  $E_{\tau^2}[Y_i - \bar{Y}] = 0$ , by (2.8) and (2.9) we get

$$\begin{aligned} E_{\tau^2} [P^{\bar{\pi}}\{\theta_i > h_i(\pi, \hat{\tau}^2, \alpha; \mathbf{Y} | \mathbf{Y})\}] \\ = \alpha - \frac{\{\bar{\rho}_1(\tau^2) - \rho_1(\tau^2)\} B'(\tau^2) z_\alpha \phi(z_\alpha)}{2nJ(\tau^2)\{1 - B(\tau^2)\}} + o(n^{-1}) \end{aligned} \quad (2.10)$$

It follows from (2.10) that

$$\begin{aligned} \int_0^\infty E_{\tau^2} [P^{\bar{\pi}}\{\theta_i > h_i(\pi, \hat{\tau}^2, \alpha; \mathbf{Y}) | \mathbf{Y}\}] \bar{\pi}(\tau^2) d\tau^2 \\ = \alpha - \frac{z_\alpha \phi(z_\alpha)}{2n} \int_0^\infty \frac{\{\bar{\rho}_1(\tau^2) - \rho_1(\tau^2)\} B'(\tau^2)}{J(\tau^2)\{1 - B(\tau^2)\}} \bar{\pi}(\tau^2) d\tau^2 + o(n^{-1}) \end{aligned} \quad (2.11)$$

Now making  $\bar{\pi}(\tau^2)$  weakly converge to  $\tau^2$  as in Ghosh(1994) or Datta and Ghosh (1995) we get from (2.11),

$$\begin{aligned} P_{\tau^2} [\theta_i > h_i(\pi, \hat{\tau}^2, \alpha; \mathbf{Y})] &= \alpha + \frac{z_\alpha \phi(z_\alpha)}{2n} \left[ \frac{\rho_1(\tau^2) B'(\tau^2)}{J(\tau^2)\{1 - B(\tau^2)\}} \right] \\ &\quad + D \left\{ \frac{B'(\tau^2)}{J(\tau^2)\{1 - B(\tau^2)\}} \right\} + o(n^{-1}) \\ &= \alpha + \frac{z_\alpha \phi(z_\alpha)}{2n\pi(\tau^2)} D \left[ \frac{\pi(\tau^2) B'(\tau^2)}{J(\tau^2)\{1 - B(\tau^2)\}} \right] + o(n^{-1}) \end{aligned} \quad (2.12)$$

The right hand side of (2.12) will be equal to  $\alpha$  to the order  $o(n^{-1})$  if and only if

$$\pi(\tau^2) \propto \frac{J(\tau^2)\{1-B(\tau^2)\}}{B'(\tau^2)}$$

i.e.,

$$\pi(\tau^2) \propto \tau^2 (\sigma^2 + \tau^2)^{-1}$$

**Remark:** The prior given above is the unique matching prior. This prior is different from the uniform prior on  $\tau^2$  (especially for small  $\tau^2$ ) proposed by Morris(1983). Since it is bounded in  $\tau^2$ , the resulting posterior is proper provided  $n > 3$ .

### 3. Estimation of the Multivariate Normal Mean

This section is devoted to the HB procedures for estimating the multivariate normal mean. Now we consider the following HB model

- I.  $Y_i | \theta_i, \mu, \tau^2 \stackrel{iid}{\sim} N(\theta_i, \sigma^2), i = 1, \dots, n;$
- II.  $\theta_i | \mu, \tau^2 \stackrel{iid}{\sim} N(\mu, \tau^2), i = 1, \dots, n;$
- III(a).  $\pi(\mu, \tau^2) \propto 1,$
- III(b).  $\pi(\mu, \tau^2) \propto \tau^2(\sigma^2 + \tau^2)^{-1},$

where  $\sigma^2$  is assumed to be known.

First we consider the hyperprior  $\pi(\mu, \tau^2) \propto 1$  in step III(a) for the hyperparameter  $(\mu, \tau^2)$ . Then the joint pdf of  $\mathbf{y}, \boldsymbol{\theta}, \mu$  and  $\tau^2$  is given by

$$\begin{aligned} f(\mathbf{y}, \boldsymbol{\theta}, \mu, \tau^2) \propto & \exp\left[-\frac{1}{2\sigma^2} (\mathbf{y} - \boldsymbol{\theta})'(\mathbf{y} - \boldsymbol{\theta})\right] \\ & \times (\tau^2)^{-\frac{n}{2}} \exp\left[-\frac{1}{2\tau^2} (\boldsymbol{\theta} - \mu\mathbf{1})'(\boldsymbol{\theta} - \mu\mathbf{1})\right] \end{aligned} \quad (3.1)$$

Now integrating with respect to  $\mu$ , it follows from (3.1) that the joint (improper) pdf of  $\mathbf{y}, \boldsymbol{\theta}$  and  $\tau^2$  is

$$f(\mathbf{y}, \boldsymbol{\theta}, \tau^2) \propto (\tau^2)^{-\frac{n-1}{2}} \exp\left[-\frac{1}{2\sigma^2} \left(\boldsymbol{\theta} - \frac{1}{\sigma^2} \mathbf{E}^{-1} \mathbf{y}\right)' \mathbf{E} \left(\boldsymbol{\theta} - \frac{1}{\sigma^2} \mathbf{E}^{-1} \mathbf{y}\right)\right] \exp\left[\frac{1}{\sigma^2} \mathbf{y}' \mathbf{y} - \frac{1}{\sigma^4} \mathbf{y}' \mathbf{E}^{-1} \mathbf{y}\right] \quad (3.2)$$

where  $\mathbf{E}^{-1} = \sigma^2(1-B)\mathbf{I}_n + \sigma^2 B n^{-1} \mathbf{J}_n$  with  $\mathbf{J}_n = \mathbf{1}\mathbf{1}'$ . Hence, conditional on  $\mathbf{y}$  and  $\tau^2$ ,

$$f(\boldsymbol{\theta}) \propto N\left[(1-B)\mathbf{y} + B\bar{y}\mathbf{1}_n, \sigma^2\left\{(1-B)\mathbf{I}_n + \frac{B}{n}\mathbf{J}_n\right\}\right].$$

Also, integrating with respect to  $\boldsymbol{\theta}$  in (3.2), one gets the joint pdf of  $\mathbf{y}$  and  $\tau^2$  given by

$$f(\mathbf{y}, \tau^2) \propto (\sigma^2 + \tau^2)^{-\frac{n-1}{2}} \exp\left[-\frac{1}{2(\sigma^2 + \tau^2)} \sum_1^n (y_i - \bar{y})^2\right]. \quad (3.3)$$

Since  $B = \sigma^2 / (\sigma^2 + \tau^2)$ , it follows from (3.3) that the joint pdf of  $\mathbf{Y}$  and  $B$  is given by

$$f(\mathbf{y}, B) \propto B^{-\frac{n-5}{2}} \exp\left[-\frac{B}{2\sigma^2} \sum_1^n (y_i - \bar{y})^2\right] \quad (3.4)$$

The HB approaching like the above was first proposed by Strawderman(1971). It follow from (3.4) that

$$E(B|\mathbf{y}) = \int_0^1 B^{-\frac{1}{2}(n-3)} \exp\left[-\frac{B}{2\sigma^2} \sum_1^n (y_i - \bar{y})^2\right] dB \div \int_0^1 B^{-\frac{1}{2}(n-5)} \exp\left[-\frac{B}{2\sigma^2} \sum_1^n (y_i - \bar{y})^2\right] dB \quad (3.5)$$

and

$$E(B^2|\mathbf{y}) = \int_0^1 B^{-\frac{1}{2}(n-1)} \exp\left[-\frac{B}{2\sigma^2} \sum_1^n (y_i - \bar{y})^2\right] dB \div \int_0^1 B^{-\frac{1}{2}(n-5)} \exp\left[-\frac{B}{2\sigma^2} \sum_1^n (y_i - \bar{y})^2\right] dB \quad (3.6)$$

Hence one can obtain

$$E(\boldsymbol{\theta}|\mathbf{y}) = E[\mathbf{y} - B(\mathbf{y} - \bar{y}\mathbf{1})|\mathbf{y}] = \mathbf{y} - E(B|\mathbf{y})(\mathbf{y} - \bar{y}\mathbf{1})$$

and

$$\begin{aligned} V(\boldsymbol{\theta}|\mathbf{y}) &= E[V(\boldsymbol{\theta}|B, \mathbf{y})|\mathbf{y}] + V[E(\boldsymbol{\theta}|B, \mathbf{y})|\mathbf{y}] \\ &= E\left[\sigma^2(1-B)\mathbf{I}_n + \frac{\sigma^2 B}{n} \mathbf{J}_n|\mathbf{y}\right] + V[\mathbf{y} - B(\mathbf{y} - \bar{y}\mathbf{1})|\mathbf{y}] \\ &= \sigma^2 \mathbf{I}_n - \sigma^2 E(B|\mathbf{y})(\mathbf{I}_n - \frac{1}{n} \mathbf{J}_n) + V(B|\mathbf{y})(\mathbf{y} - \bar{y}\mathbf{1})(\mathbf{y} - \bar{y}\mathbf{1})' \end{aligned}$$

Next, we consider the hyperprior  $\pi(\boldsymbol{\mu}, \tau^2) \propto \tau^2(\sigma^2 + \tau^2)^{-1}$  given by III(b), which is the type II Beta density for  $\tau^2$ . It is easy to see that the joint pdf of  $\mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\mu}$  and  $\tau^2$  is given by

$$\begin{aligned} f(\mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\mu}, \tau^2) &\propto (\tau^2)^{-\left(\frac{n}{2}-1\right)} (\sigma^2 + \tau^2)^{-1} \\ &\quad \times \exp\left[-\frac{1}{2\sigma^2} (\mathbf{y} - \boldsymbol{\theta})'(\mathbf{y} - \boldsymbol{\theta}) - \frac{1}{2\tau^2} (\boldsymbol{\theta} - \boldsymbol{\mu}\mathbf{1})'(\boldsymbol{\theta} - \boldsymbol{\mu}\mathbf{1})\right] \end{aligned} \quad (3.7)$$

Now integrating with respect to  $\boldsymbol{\mu}$ , it follows from (3.7) that the joint (improper) pdf of  $\mathbf{y}, \boldsymbol{\theta}$  and  $\tau^2$  is

$$\begin{aligned} f(\mathbf{y}, \boldsymbol{\theta}, \tau^2) &\propto (\tau^2)^{-\frac{n-3}{2}} (\sigma^2 + \tau^2)^{-1} \\ &\quad \exp\left[-\frac{1}{2\sigma^2} (\boldsymbol{\theta} - \frac{1}{\sigma^2} \mathbf{E}^{-1}\mathbf{y})' \mathbf{E} (\boldsymbol{\theta} - \frac{1}{\sigma^2} \mathbf{E}^{-1}\mathbf{y})\right] \\ &\quad \exp\left[-\frac{1}{\sigma^2} \mathbf{y}' \mathbf{y} - \frac{1}{\sigma^4} \mathbf{y}' \mathbf{E}^{-1}\mathbf{y}\right] \end{aligned} \quad (3.8)$$

one gets the joint pdf of  $\mathbf{y}$  and  $\tau^2$  given by

$$f(\mathbf{y}, \tau^2) \propto \sigma^2 \tau^2 (\sigma^2 + \tau^2)^{-\frac{n+1}{2}} \exp\left[-\frac{1}{2(\sigma^2 + \tau^2)} \sum_{i=1}^n (y_i - \bar{y})^2\right] \quad (3.9)$$

The joint pdf of  $\mathbf{y}$  and  $B$  is given by



$$f(\mathbf{y}, B) \propto (1-B)B^{\frac{n-5}{2}} \exp\left[-\frac{B}{2\sigma^2} \sum_{i=1}^n (y_i - \bar{y})^2\right] \quad (3.10)$$

It follows from (3.10) that

$$\begin{aligned} E(B|\mathbf{y}) &= \int_0^1 (1-B)B^{\frac{1}{2}(n-3)} \exp\left[-\frac{B}{2\sigma^2} \sum_{i=1}^n (y_i - \bar{y})^2\right] dB \\ &\div \int_0^1 (1-B)B^{\frac{1}{2}(n-5)} \exp\left[-\frac{B}{2\sigma^2} \sum_{i=1}^n (y_i - \bar{y})^2\right] dB \end{aligned} \quad (3.11)$$

and

$$\begin{aligned} E(B^2|\mathbf{y}) &= \int_0^1 (1-B)B^{\frac{1}{2}(n-1)} \exp\left[-\frac{B}{2\sigma^2} \sum_{i=1}^n (y_i - \bar{y})^2\right] dB \\ &\div \int_0^1 (1-B)B^{\frac{1}{2}(n-5)} \exp\left[-\frac{B}{2\sigma^2} \sum_{i=1}^n (y_i - \bar{y})^2\right] dB \end{aligned} \quad (3.12)$$

One can obtain  $V(B|\mathbf{y})$  from (3.11) and (3.12) to obtain  $E(\boldsymbol{\theta}|\mathbf{y})$  and  $V(\boldsymbol{\theta}|\mathbf{y})$ .

#### 4. Numerical Example

We now revisit the famous baseball data of Efron and Morris(1975). They considered the batting averages of 18 baseball players in 1970 after each had batted 45 times. Based on these batting averages, they estimated the players' batting averages for the remainder of the season. Actually the values  $Y_i$  are minor adjustments to the observed averages after 45 appearances given by  $Y_i = 0.4841 + 0.0659\sqrt{45} \arcsin(2\sqrt{\hat{p}_i} - 1)$ , rounded to three significant figures. The observed average actually is  $\hat{p}_i$ ; for example  $\hat{p}_1 = 18/45 = 0.400$  for player 1 (Roberto Clemente). The arcsin transformation stabilizes variances, and the constants 0.4841 and 0.0659 are chosen so that the  $\{Y_i\}$  and  $\{\hat{p}_i\}$  have the same mean 0.26567 and standard deviation 0.0659. The same transformation  $\theta_i = 0.4841 + 0.0659\sqrt{45} \arcsin(2\sqrt{p_i} - 1)$  was made to the true value  $p_i$ , being the proportion of success during the remainder of the season for batter  $i$ . The name of the players and other information about this problems are contained in Efron and Morris(1975).

We used for formulas (3.5) and (3.6) in case 1 having uniform hyperprior which is III(a). Also we used for formulas (3.16) and (3.17) in case 2 using the hyperprior which is III(b). Then eventually we calculate  $E(\theta_i|\mathbf{y})$  and  $V(\theta_i|\mathbf{y})$ .

The results are summarized in Table 1.

What follows the true values  $\theta_i$ 's refer to the baseball players' actual batting averages for the remainder of the season. Also,  $\hat{\theta}_{i,HB1}$  and  $\hat{\theta}_{i,HB2}$  are denoted by the two different HB estimates of  $\theta_i$  respectively. The stand errors associated with  $\hat{\theta}_{i,HB1}$  and  $\hat{\theta}_{i,HB2}$  are denoted by  $s_{i,HB1}$  and  $s_{i,HB2}$  respectively.

It turns out that

$$(18\sigma^2)^{-1} \sum_{i=1}^{18} (y_i - \theta_i)^2 = 0.9345$$

$$(18\sigma^2)^{-1} \sum_{i=1}^{18} (\hat{\theta}_{i,HB1} - \theta_i)^2 = 0.2937$$

$$(18\sigma^2)^{-1} \sum_{i=1}^{18} (\hat{\theta}_{i,HB2} - \theta_i)^2 = 0.3295$$

The HB estimates serve well as point estimates. The HB1 estimates are slightly better than HB2 estimates overall. But two HB estimates are quite a comparable.

Table 1. The true value, the maximum likelihood estimates and the two HB estimates with standard errors.

$n$	$y_i$	$\theta_i$	$\hat{\theta}_{i,HB1}$	$s_{i,HB1}$	$\hat{\theta}_{i,HB2}$	$s_{i,HB2}$
1	0.395	0.346	0.305	0.049	0.318	0.046
2	0.375	0.300	0.299	0.048	0.310	0.044
3	0.355	0.279	0.292	0.047	0.301	0.043
4	0.334	0.223	0.285	0.046	0.292	0.041
5	0.313	0.276	0.278	0.045	0.284	0.040
6	0.291	0.273	0.271	0.045	0.274	0.040
7	0.269	0.266	0.264	0.044	0.265	0.039
8	0.247	0.211	0.257	0.044	0.256	0.039
9	0.247	0.271	0.257	0.044	0.256	0.039
10	0.247	0.232	0.257	0.044	0.256	0.039
11	0.224	0.266	0.250	0.045	0.246	0.040
12	0.224	0.258	0.250	0.045	0.246	0.040
13	0.224	0.306	0.250	0.045	0.246	0.040
14	0.224	0.267	0.250	0.045	0.246	0.040
15	0.224	0.228	0.250	0.045	0.246	0.040
16	0.200	0.288	0.242	0.045	0.236	0.041
17	0.175	0.318	0.234	0.047	0.225	0.042
18	0.148	0.200	0.226	0.048	0.214	0.044

## References

1. Datta, G. S. and Ghosh, J. K. (1995). On Priors Providing Frequentist Validity for Bayesian Inference. *Biometrika* 82, 37-45.
2. Datta, G. S., Ghosh, M. and Mukerjee, R. (2000). Some New Results on Probability Matching Priors. *Calcutta Statistical Association Bulletin* 50, 179-192.
3. Efron, B. and Morris, C. N. (1975) Data Analysis Using Stein's Estimator and Its Generalizations. *Journal of the American Statistical Association* 70, 311-319.
4. Ghosh, J. K. (1994). Higher Order Asymptotics. Institute of Mathematical Statistics and American Statistical Association, Hayward, California, USA.
5. Kass, R. E. and Steffey, D. (1989). Approximate Bayesian Inference in Conditionally Independent Hierarchical Models (Parametric Empirical Bayes Models). *Journal of the American Statistical Association* 84, 717-726.
6. Morris, C. N. (1983). Parametric Empirical Bayes Confidence Intervals. In *Scientific Inference, Data Analysis and Robustness*, Academic Press, New York, 25-50.

[ received date : Oct. 2004, accepted date : Nov. 2004 ]