ESTIMATING THE FINITE POPULATION MEAN USING EMPIRICAL BAYES METHODS

Balgobin Nandram, Worcester Polytechnic Institute, Worcester, MA
J. Sedransk, National Center for Health Statistics, Washington, DC

KEY WORDS AND PHRASES: Small area estimation; Bayesian credible intervals; Confidence intervals; Uniform integrability; ANOVA estimators.

DISCUSSION:

S. M. Miller
Bureau of Labor Statistics, Washington, D.C.

Summary

Many finite populations which are sampled repeatedly change slowly over time. Then estimation of finite population characteristics for the current occasion, ℓ, may be improved by the use of data from previous surveys. In this paper, we investigate the use of empirical Bayes procedures based on a superpopulation model having two stages: (a) population units on the i-th occasion are a random sample from the normal distribution with mean μ_i and variance σ_i^2, and (b) $\mu_1,\ldots,\mu_\ell$ are a random sample from the normal distribution with mean θ and variance δ^2. Here, the σ_i^2, θ and δ^2 are assumed to be unknown; this generalizes the specification studied by Ghosh and Meeden (1986). We first make large-sample comparisons of the empirical Bayes estimator of the finite population mean on the current occasion with the corresponding Bayes estimator and with several additional "natural" estimators. We also consider empirical Bayes and Bayes credible intervals for the finite population mean.

1. Introduction

Many finite populations which are sampled repeatedly using large scale surveys change slowly over time. Consequently, data from earlier surveys can be used profitably to obtain improved estimates of finite population parameters on the current occasion. In a similar way, estimates may be required for a particular small area, and information is available for other related areas. An example of the former situation is the National Health Interview Survey conducted annually by the National Center for Health Statistics. There is great stability over time in the response to variables such as the presence or absence of color blindness, acute bronchitis and acute digestive systems conditions, number of restricted activity days within the past two weeks, and self-assessment of quality of health.

To address these issues Ghosh and Meeden (1986) used empirical Bayes (EB) methodology. They assumed a normal-theory, two stage linear model with equal sampling variances. Subsequently, Ghosh and Lahiri (1987) relaxed the normality assumption by assuming posterior linearity and the existence of fourth moments in (1.1) and (1.2) below. However, they retained the assumption of equal sampling variances.

Here we generalize this research by considering the important case of unequal, unknown sampling variances. While this problem specification does not fit the EB paradigm exactly, the methods of proof and results are related to those of Ghosh and Meeden (1986).

It is assumed throughout that each of a sequence of ℓ finite populations has been sampled with $\underset{\sim}{Y}_i = (Y_{i1},\ldots,Y_{iN_i})$ denoting the vector of values of the N_i units in the population on the i^{th} occasion $(i = 1,\ldots,\ell)$. Also, given a sample of n_i units $(0 < n_i \le N_i)$ on the i^{th} occasion, let s_i denote the set of units sampled on the i^{th} occasion, $\underset{\sim}{Y}_{s_i} = (Y_{i1},\ldots,Y_{in_i})'$ the vector of values of units sampled on the i^{th} occasion and $\underset{\sim}{Y}_s = (\underset{\sim}{Y}'_{s_1},\underset{\sim}{Y}'_{s_2},\ldots,\underset{\sim}{Y}'_{s_\ell})'$ the vector of values of all sampled units.

As the basis for inference we assume the superpopulation model:

$$Y_{i1},Y_{i2},\ldots,Y_{iN_i}|\mu_i,\sigma_i^2 \overset{i.i.d}{\sim} N(\mu_i,\sigma_i^2) \qquad (1.1)$$

with independence over $i = 1,2,\ldots,\ell$, and

$$\mu_1,\mu_2,\ldots,\mu_\ell|\theta,\delta^2 \overset{i.i.d}{\sim} N(\theta,\delta^2). \qquad (1.2)$$

Our objective is to make inference about the current finite population mean,

$$\gamma(\underset{\sim}{Y}_\ell) = \sum_{j=1}^{N_\ell} Y_{\ell j}/N_\ell$$

where θ, δ^2 and $\underset{\sim}{\sigma}^2 = (\sigma_1^2,\ldots,\sigma_\ell^2)'$ are assumed to be fixed but unknown. We proceed by first finding the Bayes estimator, e_B, of $\gamma(\underset{\sim}{Y}_\ell)$, and then developing an empirical Bayes estimator, e_{EB}, by substituting estimates of θ, δ^2 and $\underset{\sim}{\sigma}^2$ in e_B. The choice of estimators is facilitated by the research of Rao, Kaplan and Cochran (1981), henceforth RKC, who investigated properties of various estimators in the one-way components of variance model defined by (1.1) and (1.2).

Bayes and empirical Bayes point estimators and credible intervals for $\gamma(\underset{\sim}{Y}_\ell)$ are defined in Section 2, together with three alternative point estimators that do not require δ^2 and the σ_i^2 to be estimated. In Section 3 we present the asymptotic properties of the EB estimator and interval. We conclude Section 3 with a brief summary of the results of an extensive numerical investigation of the performance of the EB estimator and interval when sample sizes are small or moderate.

2. Bayes, Empirical Bayes and Alternative Procedures

2.1 Point estimation

Viewing (1.1) as the sampling model and (1.2) as the prior in a Bayesian analysis, it is well known that, given $\underset{\sim}{Y}_s$, θ, $\underset{\sim}{\sigma}^2$ and δ^2, $\mu_1,\mu_2,\ldots,\mu_\ell$ are independently distributed with

$$\mu_i \sim N(\omega_i\theta + (1-\omega_i)\bar{Y}_i, \delta^2\omega_i) \qquad (2.1)$$

where $\bar{Y}_i = \sum_{j=1}^{n_i} Y_{ij}/n_i$ and $\omega_i = (\sigma_i^2/n_i)/\{(\delta^2+(\sigma_i^2/n_i)\}$. Let $f_i = n_i/N_i$ denote the sampling fraction on the i^{th} occasion. Then conditioning on $\underset{\sim}{Y}_s$, it follows from (2.1) that $\gamma(\underset{\sim}{Y}_\ell)$ is univariate normal with mean e_B and variance ν_B^2 where

$$e_B = E\{\gamma(\underset{\sim}{Y}_\ell)|\underset{\sim}{Y}_s\} = \bar{Y}_\ell - (1-f_\ell)\omega_\ell(\bar{Y}_\ell-\theta) \qquad (2.2a)$$

and

$$\nu_B^2 = var(\gamma(\underset{\sim}{Y}_\ell)|\underset{\sim}{Y}_s) = (1-f_\ell)\{f_\ell + (1-f_\ell)(1-\omega_\ell)\}\sigma^2/n_\ell \qquad (2.2b)$$

Thus, under squared error loss, the Bayes estimator of $\gamma(\underset{\sim}{Y}_\ell)$ is $e_B = E(\gamma(\underset{\sim}{Y}_\ell)|\underset{\sim}{Y}_s)$ and the Bayes risk is $\nu_B^2 = var(\gamma(\underset{\sim}{Y}_\ell)|\underset{\sim}{Y}_s)$.

Let $\hat{\omega}_i = (S_i^2/n_i)/\{\delta^2+(S_i^2/n_i)\}$ where $S_i^2 = \sum_{j=1}^{n_i} (Y_{ij}-\bar{Y}_i)^2/(n_i-1)$, $i = 1,2,\ldots,\ell$. Then a pseudo-empirical Bayes estimator of $\gamma(\underset{\sim}{Y}_\ell)$ is e^*_{EB} where

$$e^*_{EB} = \bar{Y}_\ell - (1-f_\ell)\hat{\omega}_\ell(\bar{Y}_\ell - \theta) . \tag{2.3}$$

Defining $\hat{\omega}_i = (S_i^2/n_i)/\{\hat{\delta}^2+(S_i^2/n_i)\}$, the proposed empirical Bayes estimator is

$$e_{EB} = \bar{Y}_\ell - (1-f_\ell)\hat{\omega}_\ell(\bar{Y}_\ell - \hat{\theta}) \tag{2.4}$$

where $\hat{\theta}$ and $\hat{\delta}^2$ are to be determined. Whereas e^*_{EB} only enters the analysis at an intermediate stage, e_{EB} in (2.4) is the estimator whose properties are to be investigated. Note that when the sampling variances are equal, (2.4) reduces to (2.12) in Ghosh and Meeden (1986).

From the results in RKC (1981), we use as the basis for our estimator of δ^2 the unbiased, ANOVA estimator, $\hat{\delta}^2_A$:

$$\hat{\delta}^2_A = \left[\sum_{i=1}^{\ell} n_i\{\bar{Y}_i - (\sum n_i n_.^{-1}\bar{Y}_i)\}^2 - \sum_{i=1}^{\ell}(1-n_i n_.^{-1})S_i^2\right] / \sum_{i=1}^{\ell} n_i(1-n_i n_.^{-1})$$

where $n_. = \sum_{i=1}^{\ell} n_i$.

Proceeding as in Ghosh and Meeden (1986), we modify $\hat{\delta}^2_A$ to

$$\hat{\delta}^2_* = \left[(\ell-1)(\ell-3)^{-1}\sum_{i=1}^{\ell} n_i\{\bar{Y}_i - (\sum n_i n_.^{-1}\bar{Y}_i)\}^2 - \sum_{i=1}^{\ell}(1-n_i n_.^{-1})S_i^2\right] / \sum_{i=1}^{\ell} n_i(1-n_i n_.^{-1})$$

for $\ell \geq 4$ and take

$$\hat{\delta}^2 = \max(0, \hat{\delta}^2_*) . \tag{2.5}$$

Note that for the case of equal variances $\hat{\delta}^2$ in (2.5) is analogous to (2.8) in Ghosh and Meeden (1986).

When $\underset{\sim}{\sigma}^2$ and δ^2 are known, the maximum likelihood estimator of θ is

$$\hat{\theta}_* = \sum_{i=1}^{\ell}(1-\omega_i)\bar{Y}_i / \sum_{i=1}^{\ell}(1-\omega_i) \tag{2.6}$$

where the $\{\bar{Y}_i\}$ are weighted inversely proportional to their variances. Here, we use

$$\hat{\theta} = \begin{cases} \sum_{i=1}^{\ell}(1-\hat{\omega}_i)\bar{Y}_i \,/\, \sum_{i=1}^{\ell}(1-\hat{\omega}_i); & \hat{\delta}^2 > 0 \\ \sum_{i=1}^{\ell}\frac{n_i}{S_i^2}\bar{Y}_i / \sum_{i=1}^{\ell}\frac{n_i}{S_i^2}; & \hat{\delta}^2 = 0 \end{cases} \tag{2.7}$$

where $\hat{\omega}_i = (S_i^2/n_i)/\{\hat{\delta}^2+(S_i^2/n_i)\}$. We need a separate estimator when $\hat{\delta}^2 = 0$ because in this case $\hat{\omega}_i = 1$, $i = 1,2,\ldots,\ell$, and $\hat{\theta}_*$ in (2.6) is indeterminate. Observe, though, that

$$\lim_{\hat{\delta}^2 \to 0} \sum_{i=1}^{\ell}(1-\hat{\omega}_i)\bar{Y}_i / \sum_{i=1}^{\ell}(1-\hat{\omega}_i) = \sum_{i=1}^{\ell}\frac{n_i}{S_i^2}\bar{Y}_i / \sum_{i=1}^{\ell}\frac{n_i}{S_i^2} ;$$

see Rao (1980) for comments about the efficiency of this estimator when $\hat{\delta}^2 = 0$.

As alternatives to e_{EB} in (2.4) we have considered three estimators of $\gamma(\underset{\sim}{Y}_\ell)$ which do not require estimates of δ^2 or the σ_i^2; i.e.,

$$e_1 = \bar{Y}_\ell, \quad e_2 = \sum_{i=1}^{\ell}\bar{Y}_i/\ell \quad \text{and} \quad e_3 = \sum_{i=1}^{\ell} n_i n_.^{-1}\bar{Y}_i . \tag{2.8}$$

Properties of e_1, e_2, and e_3 are easily determined since $\bar{Y}_i$ and S_i^2 are independent with

$\bar{Y}_i \sim N(\theta, \delta^2 + \sigma_i^2 n_i^{-1})$ and $S_i^2 \sim \sigma_i^2\chi^2_{n_i-1}/(n_i-1)$, $i = 1,2,\ldots,\ell$.

2.2. Credible intervals

Given $\underset{\sim}{Y}_s$, θ, δ^2 and the σ_i^2, $\gamma(\underset{\sim}{Y}_\ell)$ has a normal distribution with mean e_B and variance ν_B^2; see (2.2a,b). Thus, an exact $100(1-\alpha)\%$ HPD credible interval for $\gamma(\underset{\sim}{Y}_\ell)$ is

$$e_B \pm z_{\alpha/2}\nu_B \tag{2.9}$$

where $\Phi(z_{\alpha/2}) = 1 - \alpha/2$ and $\Phi(\cdot)$ is the standard normal cumulative distribution function. When θ, δ^2 and the σ_i^2 are unknown, we suggest using

$$e_{EB} \pm z_{\alpha/2}\hat{\nu}_B \tag{2.10}$$

where

$$\hat{\nu}_B = \left[\left\{(1-f_\ell)\left[f_\ell + (1-f_\ell)(1-\hat{\omega}_\ell)\right]S_\ell^2/n_\ell\right\}(n_\ell-1)/(n_\ell-2)\right]^{1/2}; \; n \geq 3.$$

In Section 3 we investigate the quality of the interval, (2.10), as an approximation to the Bayes interval, (2.9).

Other authors have considered EB confidence intervals. Morris (1983 a,b) gave a general definition of an EB confidence interval, but also investigated in greater detail the existence and construction of EB intervals for the μ_i in (1.2) when the σ_i^2 in (1.1) are equal. He provided empirical evidence that the intervals have approximately the correct probability content. Carlin and Gelfand (1990) have recently proposed and studied a potentially useful method to improve the coverage properties of naive EB confidence intervals. Here, the coverage probability may either be conditional on a summary of the data (i.e., quasi-Bayes or Bayes) or averaged over the marginal distribution of the data (i.e., empirical Bayes). Unfortunately, it appears from their examples (2.3, 2.4) that implementation of their methodology for our specification, i.e., unequal unknown σ_i^2, unknown θ and δ^2, and unequal n_i, will be difficult.

3. Asymptotic Properties

Let e be any estimator of $\gamma(\underset{\sim}{Y}_\ell)$; e may be a function of θ, $\underset{\sim}{\sigma}^2$ or δ^2 but not $\underset{\sim}{\mu} = (\mu_1,\ldots,\mu_\ell)$. Then the Bayes risk (risk integrated over both $\underset{\sim}{\mu}$ and $\underset{\sim}{Y}$) under squared error loss, r(e), is

$$r(e) = E_{\underset{\sim}{Y}} E_{\underset{\sim}{\mu}|\underset{\sim}{Y}}(e - \gamma(\underset{\sim}{Y}_\ell))^2 .$$

That is, all expectations are taken over the marginal distribution of $\underset{\sim}{Y}$ obtained from (1.1) and (1.2). Further we have as in Lemma 3 of Ghosh and Meeden (1986) that

$$r(e) - r(e_B) = E(e - e_B)^2 . \tag{3.1}$$

Our principal objective is to prove Theorem 1, which provides an asymptotic upper bound for $r(e_{EB}) - r(e_B)$. To do so, we first state three lemmas and then use (3.1). Our proofs are similar in style to those in Ghosh and Meeden (1986), but the details are usually different. Moreover, our proofs tend to be more cumbersome because each of the σ_i^2 is estimated separately.

All of the results in this section will be established under the following mild conditions analogous to those used by Ghosh and Meeden (1986): $0 < \delta^2 < \infty$, $\sup_{i\geq 1} \sigma_i^2 = \sigma^2 < \infty$, $\inf_{i\geq 1} n_i = 2$, and $\sup_{i\geq 1} n_i = k < \infty$. Lemmas 1 and 2 are easily established. Lemma 3, used in the proof of Theorem 1, is proved in Appendix A.

Lemma 1

$\hat{\delta}_*^2 \xrightarrow{L_2} \delta^2$ as $\ell \to \infty$, where "L_2" means convergence in mean square.

Lemma 2

$\hat{\delta}_*^2 \xrightarrow{a.s.} \delta^2$ as $\ell \to \infty$, where "a.s." means almost surely.

The following two corollaries are immediate consequences of Lemma 2.

Corollary 1

$\hat{\delta}^2 \xrightarrow{a.s.} \delta^2$ as $\ell \to \infty$, where $\hat{\delta}^2$ is defined in (2.5).

Corollary 2

$$\max_{i=1,2,\ldots,\ell} |\hat{\hat{\omega}}_i - \hat{\omega}_i| \xrightarrow{a.s.} 0, \text{ as } \ell \to \infty.$$

See (2.3) and (2.4) for definitions.

Lemma 3

$\hat{\theta} \xrightarrow{p} \theta$ as $\ell \to \infty$, where "p" means convergence in probability.

Proof

Appendix A shows that given $\hat{\delta}^2 > 0$, $\hat{\theta} \xrightarrow{p} \theta$ as $\ell \to \infty$; and by using Lemma 1 (or Lemma 2) the result follows.

We now state and sketch the proof of Theorem 1.

Theorem 1

Under the assumptions

(i) $0 < \delta^2 < \infty$; $0 < \sigma_i^2 \leq \sup_{j\geq 1} \sigma_j^2 = \sigma^2 < \infty$ for $i = 1,\ldots,\ell$

(ii) $\inf_{i\geq 1} n_i = 2$ and $\sup_{i\geq 1} n_i = k < \infty$,

$$\lim_{\ell\to\infty} E(e_{EB}-e_B)^2 \leq \lim_{\ell\to\infty} (1-f_\ell)^2 \omega_\ell \frac{\sigma_\ell^2}{n_\ell} E\left[(1-\hat{\hat{\omega}}_\ell)^2 (S_\ell^2/\sigma_\ell^2 - 1)^2\right]. \quad (3.2)$$

Proof

By using the Cauchy Schwarz inequality,

$$E(e_{EB}-e_B)^2 \leq E(e_{EB}-e_{EB}^*)^2 + 2\{E(e_{EB}-e_{EB}^*)^2\}^{1/2}\{E(e_{EB}^*-e_B)^2\}^{1/2} + E(e_{EB}^*-e_B)^2. \quad (3.3)$$

Now using the definitions of e_{EB}^* and e_{EB} in (2.3) and (2.4) respectively

$$E(e_{EB}-e_{EB}^*)^2 = (1-f_\ell)^2 E[(\hat{\omega}_\ell - \hat{\hat{\omega}}_\ell)(\bar{Y}_\ell - \theta) - \hat{\omega}_\ell(\hat{\theta}-\theta)]^2. \quad (3.4)$$

Also by the distributional properties of $\bar{Y}_i$ and S_i^2 it follows that

$$E(e_{EB}^*-e_B)^2 = (1-f_\ell)^2 \omega_\ell \frac{\sigma_\ell^2}{n_\ell} E\left[(1-\hat{\hat{\omega}}_\ell)^2 (S_\ell^2/\sigma_\ell^2 - 1)^2\right]$$
$$\leq \delta^2 + \frac{\sigma^2}{2} < \infty. \quad (3.5)$$

Appendix B shows that $E(e_{EB}-e_{EB}^*)^2 \longrightarrow 0$ as $\ell \to \infty$. The result (3.2) follows by applying (3.3), (3.4) and (3.5).

The bound given by (3.2) can be replaced by other, more useful, bounds. First, since $1 - \hat{\hat{\omega}}_\ell \leq n_\ell \delta^2/S_\ell^2$

$$\lim_{\ell\to\infty} E(e_{EB}-e_B)^2 \leq \delta^2 \lim_{\ell\to\infty} (1-f_\ell)^2 (1-\omega_\ell). \quad (3.6)$$

Second, since $E\left[(1-\hat{\hat{\omega}}_\ell)^2 (S_\ell^2/\sigma_\ell^2 - 1)^2\right] \leq 2/(n_\ell - 1)$,

$$\lim_{\ell\to\infty} E(e_{EB}-e_B)^2 \leq 2 \lim_{\ell\to\infty} (1-f_\ell)^2 \omega_\ell \sigma_\ell^2/n_\ell(n_\ell - 1). \quad (3.7)$$

In practical situations, the bounds on $\lim_{\ell\to\infty}\{r(e_{EB}) - r(e_B)\} = \lim_{\ell\to\infty} E(e_{EB} - e_B)^2$ may be very small. If δ^2 is small, the bound in (3.6) will be small since $(1-f_\ell)^2(1-\omega_\ell) \leq 1$ for any ℓ. Second, writing $\lim_{\ell\to\infty} n_\ell = n_*$, the right side of (3.7) is $O(n_*^{-3})$, which should be small in many applications. Moreover, one may easily imagine a sequence of surveys or experiments with improving precision of measurement so that $\sigma_\ell^2/n_\ell \to 0$ as $\ell \to \infty$. In the latter case, using (3.7), $\lim_{\ell\to\infty} E(e_{EB} - e_B)^2 = 0$ and the EB estimator is asymptotically optimal in the sense of Robbins (1955).

To compare e_{EB} with e_1, e_2 and e_3 first note that

$$E(e_1 - e_B)^2 = (1-f_\ell)^2 \omega_\ell^2 \{\delta^2 + (\sigma_\ell^2/n_\ell)\} \quad (3.8)$$

and

$$E(e_i - e_B)^2 = \sum_{j=1}^{\ell} a_j^2 \{\delta^2 + (\sigma_j^2/n_j)\}$$
$$+ \{1-(1-f_\ell)\omega_\ell - 2a_\ell\}\{1-(1-f_\ell)\omega_\ell\}\{\delta^2 + (\sigma_\ell^2/n_\ell)\} \quad (3.9)$$

where $a_j = \ell^{-1}$ for $i = 2$ and $a_j = n_j n_.^{-1}$ for $i = 3$.

Using (3.7), and (3.8), $\lim_{\ell\to\infty} E(e_{EB} - e_B)^2 \leq \lim_{\ell\to\infty} E(e_1 - e_B)^2$ provided that $\lim_{\ell\to\infty} n_\ell \geq 3$. Moreover, using (3.6) and (3.9), it can be shown that

$\lim_{\ell\to\infty} E(e_{EB} - e_B)^2 \leq \lim_{\ell\to\infty} E(e_i - e_B)^2$ for $i = 2,3$.

We next present asymptotic results which give conditions when the EB credible interval in (2.10) will be a good approximation for the Bayes interval in (2.9).

Theorem 2

Under the conditions of Theorem 1,

$$\lim_{\ell\to\infty} E(\hat{\nu}_B - \nu_B)^2 \leq 6\sqrt{2} \lim_{\ell\to\infty}\{(1-f_\ell)\sigma_\ell^2/n_\ell(n_\ell - 1)^{1/2}\}. \quad (3.10)$$

The proof of (3.10) is sketched in Appendix C.

Using Theorem 2 and the Liapunov inequality,

$$\lim_{\ell\to\infty} E|\hat{\nu}_B - \nu_B| \leq 6^{1/2} 2^{1/4} \lim_{\ell\to\infty} (1-f_\ell)^{1/2} \sigma_\ell / n_\ell^{1/2} (n_\ell - 1)^{1/4}. \quad (3.11)$$

Consequently, the width of the estimated H.P.D. credible interval may be close to the width of the true H.P.D. credible interval for $\gamma(\underset{\sim}{Y}_\ell)$.

Finally, we state and prove Corollary 3 which follows by an application of both Theorems 1 and 2.

Corollary 3

Under the conditions of Theorem 1,

$$\lim_{\ell\to\infty} E|e_{EB} - e_B \pm z_{\alpha/2}(\hat{\nu}_B - \nu_B)|$$
$$\leq 2^{1/2} \lim_{\ell\to\infty} (1-f_\ell)^{1/2} \{(1-f_\ell)^{1/2} \omega_\ell^{1/2} (n_\ell - 1)^{-1/4}$$
$$+ 3^{1/2} 2^{1/4} z_{\alpha/2}\} \sigma_\ell n_\ell^{-1/2} (n_\ell - 1)^{-1/4}. \quad (3.12)$$

Proof

By the Liapunov and Minkowski inequalities,

$$\lim_{\ell\to\infty} E|e_{EB} - e_B \pm z_{\alpha/2}(\hat{\nu}_B - \nu_B)|$$
$$\leq \{\lim_{\ell\to\infty} E(e_{EB} - e_B)^2\}^{1/2} + z_{\alpha/2}\{\lim_{\ell\to\infty} E(\hat{\nu} - \nu_B)^2\}^{1/2} \quad (3.13)$$

Corollary 3 follows by substituting (3.7) and (3.10) in (3.13).

From (3.10), (3.11) and (3.12) it is clear that for large ℓ the EB interval, (2.10), will provide a reasonable approximation for the Bayes H.P.D. interval, (2.9), when n_ℓ is large or σ_ℓ^2/n_ℓ is small.

In concluding we summarize the results of our numerical investigation but we omit the details. We perform a sequence of numerical examples which indicate that when $\ell \geq 20$ and $n \geq 20$ ($n_i = n$, $i = 1,2,\ldots,\ell$), the empirical Bayes point estimator and interval (of the finite population mean) are reasonable approximations for the Bayes estimator and HPD interval respectively. Moreover, e_{EB} is always better than e_1; and except for three cases (with $n = 10$) e_{EB} is better than e_2. Our results also show that increasing n is more profitable than increasing ℓ. For example, for the credible interval, it is preferable to have ($\ell = 10$, $n = 20$) rather than ($\ell = 20$, $n = 10$) and there are substantial gains by having ($\ell = 10$, $n = 30$) rather than ($\ell = 30$, $n = 10$).

APPENDIX A: COMPLETION OF PROOF OF LEMMA 3

$P\{|\hat{\theta}-\theta| > \epsilon \mid \hat{\delta}^2 > 0\} \longrightarrow 0$ as $\ell \longrightarrow \infty$, $\forall \epsilon > 0$

Proof

Noting that all arguments apply for $\hat{\delta}^2 > 0$, we first establish the bound in (A.2) for $|\hat{\theta} - \theta|$.

It is easy to show that

$$\left\{\sum_{i=1}^{\ell} (1-\hat{\omega}_i)\right\}^{-1} \leq \frac{1}{2\ell}\left[k + (\hat{\delta}^2)^{-1} \max_{j=1,2,\ldots,\ell} S_j^2\right]$$

and

$$|\sum_{i=1}^{\ell} (1-\hat{\omega}_i)(\bar{Y}_i-\theta)| \leq |\sum_{i=1}^{\ell} (1-\hat{\hat{\omega}}_i)(\bar{Y}_i-\theta)|$$

$$+ \max_{i=1,2,\ldots,\ell} |\hat{\omega}_i-\hat{\hat{\omega}}_i| \cdot \frac{1}{\ell}\sum_{i=1}^{\ell} |\bar{Y}_i-\theta|. \quad (A.1)$$

Now, using (A.1), $|\hat{\theta}-\theta| \leq \frac{1}{2}\left[k + (\hat{\delta}^2)^{-1} \max_{j=1,2,\ldots,\ell} S_j^2\right]$

$$\times\left\{\frac{1}{\ell}\left|\sum_{i=1}^{\ell} (1-\hat{\hat{\omega}}_i)(\bar{Y}_i-\theta)\right|\right.$$

$$\left. + \max_{i=1,2,\ldots,\ell}\left|\hat{\omega}_i-\hat{\hat{\omega}}_i\right| \cdot \frac{1}{\ell}\sum_{i=1}^{\ell} |\bar{Y}_i-\theta|\right\}. \quad (A.2)$$

Second we show that both $\frac{1}{\ell}\left|\sum_{i=1}^{\ell} (1-\hat{\hat{\omega}}_i)(\bar{Y}_i-\theta)\right|$ and $\max_{i=1,2,\ldots,\ell}\left|\hat{\omega}_i-\hat{\hat{\omega}}_i\right| \cdot \frac{1}{\ell}\sum_{i=1}^{\ell} |\bar{Y}_i-\theta|$ converge to zero in probability as $\ell \to \infty$. By an argument similar to Ghosh and Meeden (1986), $\frac{1}{\ell}\sum_{i=1}^{\ell} |\bar{Y}_i-\theta|$ has finite expectation, so it is bounded in probability as $\ell \longrightarrow \infty$. It follows by Corollary 2 that

$$\max_{i=1,2,\ldots\ell} |\hat{\omega}_i-\hat{\hat{\omega}}_i| \cdot \frac{1}{\ell}\sum_{i=1}^{\ell} |\bar{Y}_i-\theta| \xrightarrow{p} 0 \text{ as } \ell \longrightarrow \infty. \quad (A.3)$$

Since $\frac{1}{\ell}\sum_{i=1}^{\ell} (1-\hat{\hat{\omega}}_i)(\bar{Y}_i-\theta)$ is an unbiased estimator of 0 and $\mathrm{var}\{\frac{1}{\ell}\sum_{i=1}^{\ell} (1-\hat{\hat{\omega}}_i)(\bar{Y}_i-\theta)\} \longrightarrow 0$ as $\ell \to \infty$, it follows that $\frac{1}{\ell}\sum_{i=1}^{\ell} (1-\hat{\hat{\omega}}_i)(\bar{Y}_i-\theta) \xrightarrow{p} 0$ as $\ell \longrightarrow \infty$. Consequently, by (A.2) and (A.3) $\frac{1}{\ell}\sum_{i=1}^{\ell} (1-\hat{\omega}_i)(\bar{Y}_i-\theta) \xrightarrow{p} 0$ as $\ell \longrightarrow \infty$.

We complete the proof by using the distributional properties of the S_j^2 to show that $\max_{j=1,2,\ldots,\ell} S_j^2$ is bounded in probability as $\ell \to \infty$.

APPENDIX B: COMPLETION OF PROOF OF THEOREM 1

$E(e_{EB}-e^*_{EB})^2 \longrightarrow 0$ as $\ell \longrightarrow \infty$.

Proof

Like Ghosh and Meeden (1986) we show that

$$(\hat{\omega}_\ell-\hat{\hat{\omega}}_\ell)(\bar{Y}_\ell-\theta) - \hat{\omega}_\ell(\hat{\theta}-\theta) \xrightarrow{p} 0 \text{ as } \ell \longrightarrow \infty$$

and the sequence,

$$\{[(\hat{\omega}_\ell-\hat{\hat{\omega}}_\ell)(\bar{Y}_\ell-\theta) - \hat{\omega}_\ell(\hat{\theta}-\theta)]^2\},$$

is uniformly integrable (u.i.); see also Serfling (1980, p. 15).

Now

$$|(\hat{\omega}_\ell-\hat{\hat{\omega}}_\ell)(\bar{Y}_\ell-\theta) - \hat{\omega}_\ell(\hat{\theta}-\theta)|$$

$$\leq \left\{\max_{i=1,2,\ldots,\ell}|\hat{\omega}_i-\hat{\hat{\omega}}_i|\right\}|\bar{Y}_\ell-\theta| + |\hat{\theta}-\theta|. \quad (B.1)$$

Since $|\bar{Y}_\ell-\theta|$ is bounded in probability as $\ell \longrightarrow \infty$, it follows by Corollary 2, Lemma 3 and (B.1) that

$$(\hat{\omega}_\ell-\hat{\hat{\omega}}_\ell)(\bar{Y}_\ell-\theta)-\hat{\omega}_\ell(\hat{\theta}-\theta) \xrightarrow{p} 0 \text{ as } \ell \longrightarrow \infty.$$

In the rest of Appendix B we establish that $\{[(\hat{\omega}_\ell-\hat{\hat{\omega}}_\ell)(\bar{Y}_\ell-\theta) - \hat{\omega}_\ell(\hat{\theta}-\theta)]^2\}$ is u.i. Now,

$$[(\hat{\omega}_\ell-\hat{\hat{\omega}}_\ell)(\bar{Y}_\ell-\theta) - \hat{\omega}_\ell(\hat{\theta}-\theta)]^2 \leq 2\{(\bar{Y}_\ell-\theta)^2+(\hat{\theta}-\theta)^2\}. \quad (B.2)$$

From Ghosh and Meeden (1986) $\{(\bar{Y}_\ell-\theta)^2\}$ is u.i., so we need only show that $(\hat{\theta}-\theta)^2$ is u.i. There are two cases.

First, suppose that $\hat{\delta}^2 = 0$. Then

$(\hat{\theta}-\theta)^2 = \left\{\sum_{i=1}^{\ell} \frac{n_i}{S_i^2}(\bar{Y}_i-\theta)/\sum_{i=1}^{\ell}\frac{n_i}{S_i^2}\right\}^2$. It suffices to show that

$$E\left\{\sum_{i=1}^{\ell} \frac{n_i}{S_i^2}(\bar{Y}_i-\theta)/\sum_{i=1}^{\ell}\frac{n_i}{S_i^2}\right\}^4 < \infty; \quad (B.3)$$

see Serfling (1980; p. 13). Denoting the vector of estimators of σ^2 by $\underset{\sim}{S}^2 = (S_1^2,S_2^2,\ldots,S_\ell^2)'$,

$$E\left\{\sum_{i=1}^{\ell} \frac{n_i}{S_i^2}(\bar{Y}_i-\theta)/\sum_{i=1}^{\ell}\frac{n_i}{S_i^2}\right\}^4$$

$$= E_{\underset{\sim}{S}^2}\left\{\mathrm{var}\left[\left\{\sum_{i=1}^{\ell} \frac{n_i}{S_i^2}(\bar{Y}_i-\theta)\right\}^2 \middle| \underset{\sim}{S}^2\right]\middle/\left\{\sum_{i=1}^{\ell}\frac{n_i}{S_i^2}\right\}^4\right\}$$

$$+ E_{\underset{\sim}{S}^2}\left\{\mathrm{var}\left\{\sum_{i=1}^{\ell} \frac{n_i}{S_i^2}(\bar{Y}_i-\theta) \mid \underset{\sim}{S}^2\right\}\middle/\left\{\sum_{i=1}^{\ell}\frac{n_i}{S_i^2}\right\}^2\right\}^2. \quad (B.4)$$

Now by the distributional properties of $\{\bar{Y}_i,S_i^2\}$, given $\underset{\sim}{S}^2$,

$$\sum_{i=1}^{\ell} \frac{n_i}{S_i^2}(\bar{Y}_i-\theta) \sim N\left[0, \sum_{i=1}^{\ell}\left[\frac{n_i}{S_i^2}\right]^2\left[\delta^2+(\sigma_i^2/n_i)\right]\right]. \quad (B.5)$$

Using (B.5),

$$\mathrm{var}\left\{\left[\sum_{i=1}^{\ell} \frac{n_i}{S_i^2}(\bar{Y}_i-\theta)\right]^2 \middle| \underset{\sim}{S}^2\right\}\middle/\left\{\sum_{i=1}^{\ell}\frac{n_i}{S_i^2}\right\}^4$$

$$= 2\left\{\sum_{i=1}^{\ell}\left[\frac{n_i}{S_i^2}\right]^2\left[\delta^2+(\sigma_i^2/n_i)\right]\right\}^2 \Big/ \left\{\sum_{i=1}^{\ell}\frac{n_i}{S_i^2}\right\}^4 \le 2\left[\delta^2+\frac{\sigma^2}{2}\right]^2 < \infty.$$

Thus the first term in (B.4) is bounded. Since

$$\mathrm{var}\left\{\sum_{i=1}^{\ell}\frac{n_i}{S_i^2}(\bar{Y}_i-\theta)\,\Big|\,\underline{S}^2\right\}\Big/\left\{\sum_{i=1}^{\ell}\frac{n_i}{S_i^2}\right\}^2 \le \left[\delta^2+\frac{\sigma^2}{2}\right] < \infty,$$

the second term in (B.4) is bounded.

In the second case $\hat{\delta}^2 > 0$. Then $(\hat{\theta}-\theta)^2 = \left\{\sum_{i=1}^{\ell}(1-\hat{w}_i)(\bar{Y}_i-\theta)/\sum_{i=1}^{\ell}(1-\hat{w}_i)\right\}^2$. After some algebra it follows that

$$\left\{\sum_{i=1}^{\ell}(1-\hat{w}_i)(\bar{Y}_i-\theta)/\sum_{i=1}^{\ell}(1-\hat{w}_i)\right\}^2$$

$$\le\left[1+(\hat{\delta}_*^2)^{-1}\max_{i=1,2,..,\ell}\left[\frac{S_i^2}{n_i}\right]\right]^2\frac{1}{\ell}\sum_{i=1}^{\ell}(\bar{Y}_i-\theta)^2. \qquad (B.6)$$

Also, $\max_{i=1,2,\ldots,\ell}\frac{S_i^2}{n_i}\overset{st}{\le}\frac{\sigma^2}{2}V$, where $V\sim\chi^2_{k-1}$, and "st" means stochastically. Thus, by (B.6) letting $\bar{U}_\ell = \frac{1}{\ell}\sum_{i=1}^{\ell}U_i$, where $U_1,U_2,\ldots,U_\ell$ are i.i.d χ_1^2,

$$\left\{\sum_{i=1}^{\ell}(1-\hat{w}_i)(\bar{Y}_i-\theta)/\sum_{i=1}^{\ell}(1-\hat{w}_i)\right\}^2\overset{st}{\le}\left[\delta^2+\frac{\sigma^2}{2}\right]\left[1+\sigma^2V/2\hat{\delta}_*^2\right]^2\bar{U}_\ell. \qquad (B.7)$$

By SLLN, $\bar{U}_\ell \xrightarrow{a.s.} 1$ as $\ell \longrightarrow \infty$. It follows by Lemma 2 that

$$\left[1+\sigma^2V/2\hat{\delta}_*^2\right]^2\bar{U}_\ell \xrightarrow{a.s.} \left[1+\sigma^2V/2\delta^2\right]^2 \text{ as } \ell\longrightarrow\infty.$$

But since $V\sim\chi^2_{k-1}$, $k<\infty$, $E\left[1+\sigma^2V/2\delta^2\right]^2<\infty$. It suffices to show that

$$\lim_{\ell\to\infty}E\left\{(1+\sigma^2V/2\hat{\delta}_*^2)^2\bar{U}_\ell-\left[1+\sigma^2V/2\delta^2\right]^2\right\}=0; \qquad (B.8)$$

see Serfling (1980; p. 15).

Now it is easy to show that

$$E\left\{\left[1+\frac{\sigma^2V}{2\hat{\delta}_*^2}\right]^2\bar{U}_\ell-\left[1+\frac{\sigma^2}{2\delta^2}V\right]^2\right\}$$

$$=(\sigma^2/\delta^2)E\{V[(\delta^2/\hat{\delta}_*^2)\bar{U}_\ell-1]\}+\frac{\sigma^4}{4\delta^4}E\{V^2[(\delta^2/\hat{\delta}_*^2)^2\bar{U}_\ell-1]\}. \qquad (B.9)$$

Applying Holder's inequality to each term on the right–hand side of (B.9),

$$\left|E\left\{V\left[\left[\frac{\delta^2}{\hat{\delta}_*^2}\right]\bar{U}_\ell-1\right]\right\}\right|\le\left[E(V^2)\right]^{1/2}\left\{E\left[\left[\frac{\delta^2}{\hat{\delta}_*^2}\right]\bar{U}_\ell-1\right]^2\right\}^{1/2}$$

and

$$\left|E\left\{V^2\left[\left[\frac{\delta^2}{\hat{\delta}_*^2}\right]^2\bar{U}_\ell-1\right]\right\}\right|\le\{E(V^4)\}^{1/2}\left\{E\left[\left[\frac{\delta^2}{\hat{\delta}_*^2}\right]^2\bar{U}_\ell-1\right]^2\right\}^{1/2}. \qquad (B.10)$$

By Minkowski's inequality and using $\ell\bar{U}_\ell\sim\chi^2_\ell$,

$$E\left[\left[\frac{\delta^2}{\hat{\delta}^2}\right]\bar{U}_\ell-1\right]^2\le\left\{E\left[\left[\frac{\delta^2}{\hat{\delta}_*^2}-1\right]\bar{U}_\ell\right]^2\right\}^{1/2}. \qquad (B.11)$$

Since $\bar{U}_\ell\xrightarrow{a.s.}1$ as $\ell\to\infty$, there exists a finite real number A s.t. $\sup_{\ell\ge1}\bar{U}_\ell\le A$ a.s. Thus by (B.11),

$$\lim_{\ell\to\infty}E\left[\left[\frac{\delta^2}{\hat{\delta}_*^2}\right]\bar{U}_\ell-1\right]^2\le A\left\{\lim_{\ell\to\infty}E\left[\frac{\delta^2}{\hat{\delta}_*^2}-1\right]^2\right\}^{1/2}$$

and by Lemmas 1 and 2

$$\lim_{\ell\to\infty}E\left[\left[\frac{\delta^2}{\hat{\delta}_*^2}\right]\bar{U}_\ell-1\right]^2=0. \qquad (B.12)$$

A similar argument shows that

$$\lim_{\ell\to\infty}E\left[\left[\frac{\delta^2}{\hat{\delta}_*^2}\right]^2\bar{U}_\ell-1\right]^2=0. \qquad (B.13)$$

But since $V\sim\chi^2_{k-1}$, $E(V^r)<\infty$ for every finite $r>0$, and (B.8) follows from (B.9)–(B.13).

APPENDIX C: PROOF OF THEOREM 2

$$\lim_{\ell\to\infty}E(\hat{\nu}_B-\nu_B)^2\le 6\sqrt{2}\lim_{\ell\to\infty}\left\{(1-f_\ell)\sigma_\ell^2/n_\ell(n_\ell-1)^{1/2}\right\}$$

Proof

First we show that

$$\lim_{\ell\to\infty}E(\hat{\nu}_B^2-\nu_B^2)^2\le 8\lim_{\ell\to\infty}\left\{(1-f_\ell)\sigma_\ell^2/n_\ell(n_\ell-1)^{1/2}\right\}^2. \qquad (C.1)$$

By Minkowski's inequality,

$$E(\hat{\nu}_B^2-\nu_B^2)^2\le(1-f_\ell)^2$$

$$\left[\frac{1}{n_\ell}\left\{E(S_\ell^2-\sigma_\ell^2)^2\right\}^{1/2}+\left\{E(\hat{w}_\ell\hat{\delta}^2-w_\ell\delta^2)^2\right\}^{1/2}\right]^2. \qquad (C.2)$$

Now,

$$\lim_{\ell\to\infty}\frac{1}{n_\ell^2}E(S_\ell^2-\sigma_\ell^2)^2=2\lim_{\ell\to\infty}(\sigma_\ell^2/n_\ell)^2/(n_\ell-1). \qquad (C.3)$$

To establish (C.1) we only need to show that

$$\lim_{\ell\to\infty}E(\hat{w}_\ell\hat{\delta}^2-w_\ell\delta^2)^2\le 2\lim_{\ell\to\infty}(\sigma_\ell^2/n_\ell)^2/(n_\ell-1). \qquad (C.4)$$

Now,

$$[\hat{w}_\ell\hat{\delta}^2-w_\ell\delta^2]^2\le(\hat{w}_\ell\hat{\delta}^2-\hat{\hat{w}}_\ell\delta^2)^2+2\delta^2|\hat{w}_\ell\hat{\delta}^2-\hat{\hat{w}}_\ell\delta^2|$$

$$+\delta^4(\hat{\hat{w}}_\ell-w_\ell)^2, \qquad (C.5)$$

and

$$[\hat{w}_\ell\hat{\delta}^2-\hat{\hat{w}}_\ell\delta^2]^2\le(\hat{w}_\ell-\hat{\hat{w}}_\ell)^2\hat{\delta}^4+2\hat{\delta}^2|\hat{\delta}^2-\delta^2|+(\hat{\delta}^2-\delta^2)^2. \qquad (C.6)$$

It is easy to show that

$$E\left[\hat{w}_\ell-\hat{\hat{w}}_\ell\right]^2\le\frac{1}{\delta^4}E(\hat{\delta}_*^2-\delta^2)^2+P(\hat{\delta}^2=0),\quad \delta^2>0 \qquad (C.7)$$

and

$$\lim_{\ell\to\infty}\delta^4E|\hat{\hat{w}}_\ell-w_\ell)^2\le 2\lim_{\ell\to\infty}(\sigma_\ell^2/n_\ell)^2(n_\ell-1)^{-1}. \qquad (C.8)$$

Using (C.7) and Lemmas 1 and 2 it follows that $\lim_{\ell\to\infty}E\left[\hat{w}_\ell-\hat{\hat{w}}_\ell\right]^2=0$. Thus, by using (C.6) and Lemmas 1 and 2 again,

$$\lim_{\ell\to\infty}E\left[\hat{w}_\ell\hat{\delta}^2-\hat{\hat{w}}_\ell\delta^2\right]^2=0. \qquad (C.9)$$

Thus (C.4) follows from (C.5), (C.8) and (C.9).

Next, we prove that

$$E(\hat{\nu}_B-\nu_B)^2\le 3\{E(\hat{\nu}_B^2-\nu_B^2)^2\}^{1/2} \qquad (C.10)$$

by first showing that

$$(\hat{\nu}_B-\nu_B)^2\le 3|\hat{\nu}_B^2-\nu_B^2|. \qquad (C.11)$$

By an application of the Liapunov inequality (C.10) follows from (C.11).

Finally, by a second application of the Liapunov inequality, (C.1) and (C.10), Theorem 2 is proved.

REFERENCES

Carlin, B.P. and Gelfand, A.E. (1990), "Approaches for Empirical Bayes Confidence Intervals," Journal of the American Statistical Association, 85, 105–114.

Ghosh, M. and Meeden, G. (1986), "Empirical Bayes Estimation in Finite Population Sampling," Journal of the American Statistical Association, 81, 1058–1062.

Ghosh, M. and Lahiri, P. (1987), "Robust Empirical Bayes Estimation of Means from Stratified Samples," Journal of the American Statistical Association, 82, 1153–1162.

Morris, C. (1983a), "Parametric Empirical Bayes Inference: Theory and Applications," Journal of the American Statistical Association, 78, 47–59.

Morris, C. (1983b), "Parametric Empirical Bayes Confidence Intervals," In Scientific Inference, Data Analysis, and Robustness, G.E.P. Box, T. Leonard, and C. F. Wu (Eds.), Academic Press, New York, 25–49.

Rao, J.N.K. (1980), "Estimating the Common Mean of Possibly Different Normal Populations: A Simulation Study," Journal of the American Statistical Association, 75, 447–453.

Rao, P.S.R.S., Kaplan, J. and Cochran, W.G. (1981), "Estimators for the One-Way Random Effects Model with Unequal Error Variances," Journal of the American Statistical Association, 76, 89–97.

Robbins, H. (1955), "An Empirical Bayes Approach to Statistics," in Proceedings of the 3rd Berkeley Symposium on Mathematical Statistics and Probability (Vol. 1), Berkeley, CA: University of California Press, 159–163.

Serfling, R.J. (1980), Approximation Theorems of Mathematical Statistics, John Wiley, New York.