EMPIRICAL BAYES CONFIDENCE INTERVAL FOR THE FINITE POPULATION MEAN OF A SMALL AREA

B. Nandram, Worcester Polytechnic Institute, 100 Institute Road, Worcester, MA 01609

Key Words: Asymptotic, Bayes risk, HPD interval

ABSTRACT

An estimator of a finite population quantity of a small area based only on data from this area is likely to have an unacceptably large standard error which can be reduced by using data from other areas. Thus given data from ℓ similar areas we obtain an interval estimator for the finite population mean of a small area. We assume that the individuals of the population of the i^{th} area have values which are a random sample from a normal distribution with mean μ_i and variance σ_i^2. Then given σ_i^2, the μ_i are a random sample from a normal distribution with mean θ and variance $\sigma_i^2\tau$ while the σ_i^2 are a random sample from an inverse gamma distribution with index η and scale $(\eta-1)\delta$. The parameters θ,τ,δ and η are all assumed fixed and unknown. We construct an empirical Bayes confidence interval for the finite population mean and investigate its asymptotic properties (as $\ell \to \infty$) by comparing the center, width and probability content of the proposed empirical Bayes interval with the highest posterior density interval.

1. INTRODUCTION

Many Federal government agencies are required to obtain estimates of population counts, unemployment rates, per capita income, health needs, crop yields, and livestock numbers for states and local government areas. For example, the National Health Planning and Resources Development Act of 1974 has created a need for accurate small area estimates. The Health Systems Agencies, mandated by the Planning Act, are required to collect and analyze data relating to the health status of the residents and to the health delivery systems in their health service areas. Consequently, there is a growing demand for reliable statistics for small areas.

Typically there is great variability among the sample sizes for the small areas with small sample sizes dominating. Since estimators based only on statistics from each area are likely to yield unacceptably large standard errors, alternative estimators which "borrow strength" from other small areas are normally used. In particular, empirical Bayes methods have been proposed for use in such situations (see, e.g., Ghosh and Meeden (1986)).

Fay and Herriot (1979), Dempster, Rubin, and Tsutakawa (1981), and Battese, Harter, and Fuller (1988) have proposed three small area models when there are covariates. Prasad and Rao discussed best linear unbiased (BLU) estimator (or predictor) for the finite population mean and described how to estimate the mean squared error of the estimator of a small area. They also described asymptotic properties of their estimators. Hulting and Harville (1991) described frequentist and Bayesian methods for constructing approximate prediction intervals for a small area population mean when possibly a mixed linear model holds. We share their concern that interval estimation has received relatively little attention in the small area literature.

As a natural extension of the model proposed by Ghosh and Meeden (1986), and a basis for inference, we assume for the N_i individuals in the population in area i the superpopulation model

$$Y_{i1},Y_{i2},\dots,Y_{iN_i} \mid \mu_i,\sigma_i^2 \overset{i.i.d.}{\sim} N(\mu_i,\sigma_i^2) \quad (1.1)$$

with independence over $i = 1,2,\dots,\ell$. Next, we specify

$$\mu_i \mid \sigma_i^2 \sim N(\theta,\tau\sigma_i^2) \quad (1.2)$$

$i = 1,2,\dots,\ell$ with independence across areas. For small area estimation, since the sample sizes are usually small, it is not entirely unreasonable to assume that the variances share an effect (i.e., they have a common distribution). In addition, because of small sample sizes it is difficult to assess whether the variability across small areas is homoscedastic. Thus finally we specify

$$\sigma_1^2,\sigma_2^2,\dots,\sigma_\ell^2 \overset{i.i.d.}{\sim} IG(\eta,(\eta-1)\delta) \quad (1.3)$$

where the inverse gamma density in (1.3) is given by

$$f(\sigma_i^2) = ((\eta-1)\delta)^\eta\left[1/\sigma_i^2\right]^{\eta+1} e^{-(\eta-1)\delta/\sigma_i^2}/\Gamma(\eta),\ \sigma_i^2 > 0$$

and $f(\sigma_i^2) = 0$ otherwise. We assume θ,τ,δ and η are fixed but unknown parameters. Since $E(\sigma_i^2) = \delta$ and for $\eta > 2$ $\mathrm{var}(\sigma_i^2) = \delta^2/(\eta-2)$, $\eta > 2$, for fixed δ, small values of η express a belief that the variances are very different whereas large values express a belief that they are very similar.

While (1.1), (1.2) and (1.3) provide a simple specification, the results might provide insight about the appropriate methodology for small area estimation. However, the model specifications are expected to hold within strata (or clusters) of the entire population of small areas. Nandram and Sedransk (1992) used a model of the form (1.1) and (1.2) with σ_i^2 fixed but unknown to construct an interval estimator of the finite population mean on the current occasion. They studied the asymptotic properties of the interval and using an extensive sampling experiment they showed that the EB interval has reasonable coverage properties for moderate sample sizes.

Letting $\underset{\sim}{Y}_i = (Y_{i1},\dots,Y_{iN_i})'$ be the vector of all values from the i^{th} area, $i = 1,2,\dots,\ell$, and $\underset{\sim}{Y} = (\underset{\sim}{Y}_1', \underset{\sim}{Y}_2',\dots,\underset{\sim}{Y}_\ell')$ be the vector of all values, we want a $100(1-\alpha)\%$ interval estimator for $\gamma(\underset{\sim}{Y}_\ell) = \sum_{j=1}^{N_\ell} Y_{\ell j}/N_\ell$. Letting s_i denote the set of n_i individuals sampled from the i^{th} area, $\overline{Y}_i = \sum_{j\epsilon s_i} Y_{ij}/n_i$,

$S_i^2 = \sum_{j\epsilon s_i} (Y_{ij}-\overline{Y}_i)^2/(n_i-1)$ and $\omega_i = (1+n_i\tau)^{-1}$,

$i = 1,2,...,\ell$. Also letting $\underset{\sim}{\rho} = (\theta,\tau,\delta,\eta)'$ be fixed, a $100(1-\alpha)\%$ highest posterior density (HPD) interval for $\gamma(\underset{\sim}{Y}_\ell)$ is

$$e_B \pm \nu_B t_{2\eta,\alpha/2} \tag{1.4}$$

where

$$e_B = \overline{Y}_\ell - (1-f_\ell)\omega_\ell(\overline{Y}_\ell-\theta),$$

$$\nu_B^2 = (1-\eta^{-1})\delta(1-f_\ell)\{f_\ell+(1-f_\ell)(1-\omega_\ell)\}/n_\ell,\ \eta>1$$

and $t_{2\eta,\alpha/2}$ is the $100(1-\alpha/2)$ percentile point of the Student t distribution with 2η degrees of freedom. Under squared error loss, e_B is the Bayes estimator of $\gamma(\underset{\sim}{Y}_\ell)$. However, since $\underset{\sim}{\rho}$ is unknown, our objective is to construct a $100(1-\alpha)\%$ EB confidence interval for $\gamma(\underset{\sim}{Y}_\ell)$ by "substituting" point estimators of the components of $\underset{\sim}{\rho}$ based on $\overline{Y}_i$ and S_i^2 into (1.4). (The interval is evaluated over the marginal distribution of the Y_{ij} in (1.1), (1.2) and (1.3).)

The literature about empirical Bayes confidence intervals provides little guidance about the present problem. For example, Morris (1983 a,b) gave a general definition of an EB confidence interval, but only investigated in detail relatively simple cases such as the existence and construction of EB intervals for the μ_i in (1.2) when the σ_i^2 in (1.1) are fixed, unknown and equal. He provided empirical evidence that these intervals have approximately the correct probability content. Laird and Louis (1987) described a sampling based method and showed how the bootstrap can be used to adjust an empirical Bayes confidence interval for uncertainty in the estimated prior distribution. They discussed a special case of our model with equal σ_i^2. It seems difficult to implement the approach of Laird and Louis (1987); see Morris' comments and the rejoinder on the unequal σ_i^2 case. Carlin and Gelfand (1990) proposed and studied a method to improve the coverage properties of naive EB confidence intervals. Unfortunately, while the method of Carlin and Gelfand (1990) is potentially useful, it is also difficult to implement their methodology for our specification (i.e., unknown $\underset{\sim}{\rho} = (\theta,\tau,\delta,\eta)'$ and unequal n_i). Laird and Louis (1989) compared the empirical performance of classical, Morris–type and bootstrap intervals on a random sample of bioassays from the National Cancer Institute data base on potential chemical carcinogens. We prefer a Morris–type approach.

Henceforth, maintaining the EB spirit, all analyses are based on the marginal distribution of the Y_{ij}. That is, the parameters μ_i and σ_i^2, $i = 1,2,...,\ell$ are eliminated (by integration) in the model given by (1.1), (1.2) and (1.3); $\underset{\sim}{\rho}$ is fixed but unknown. In section 2 point estimators for $\underset{\sim}{\rho} = (\theta,\tau,\eta,\delta)'$ are obtained and their asymptotic properties are described. In section 3 we develop a two stage empirical Bayes confidence interval for $\gamma(\underset{\sim}{Y}_\ell)$ and, using asymptotic theory, we compare it with the HPD interval in (1.4). Section 4 has concluding remarks.

2. PARAMETRIC POINT ESTIMATORS

In this section we construct point estimators for $\underset{\sim}{\rho}$ and investigate their asymptotic properties. Like Ghosh and Meeden (1986) we assume throughout that $\inf_{i\geq 1} n_i = 2$ and $\sup_{i\geq} n_i = k < \infty$. Both assumptions are realistic in many applications including small area estimation. The assumption $\inf_{i\geq 1} n_i = 2$ can be relaxed since a few n_i could be unity but the presentation is more difficult. Using the marginal distribution of the Y_{ij} we construct point estimators of $\underset{\sim}{\rho}$ and provide relevant asymptotic properties.

2.1 Point Estimators

We note that while for each i $\overline{Y}_i$ and S_i^2 are not independently distributed, $\overline{Y}_i$ and S_i^2 are respectively independently distributed over i, $i = 1,2,...,\ell$. Moreover,

$$\{(1-\omega_i)/(1-\eta^{-1})\delta\tau\}^{1/2}(\overline{Y}_i-\theta) \sim t_{2\eta} \tag{2.1}$$

$i = 1,2,...,\ell$. Although the distribution of S_i^2 can be written down, it is relatively complicated.

First, assuming that τ,δ and η are known, we construct an estimator of θ. Using (2.1) it is easy to show that the best linear unbiased estimator of θ is $\hat{\theta}_*$ where

$$\hat{\theta}_* = \sum_{i=1}^{\ell} (1-\omega_i)\overline{Y}_i / \sum_{i=1}^{\ell} (1-\omega_i). \tag{2.2}$$

Note that $\hat{\theta}_*$ depends only on τ. Let $n_T = \sum_{i=1}^{\ell} n_i$,

$$\overline{\overline{Y}} = n_T^{-1} \sum_{i=1}^{\ell} n_i\overline{Y}_i.$$

$$BMS = (\ell-1)^{-1} \sum_{i=1}^{\ell} n_i(\overline{Y}_i-\overline{\overline{Y}})^2$$

and

$$\hat{\delta} = WMS = (n_T-\ell)^{-1} \sum_{i=1}^{\ell} (n_i-1)S_i^2. \tag{2.3}$$

We note that in (2.3) $\hat{\delta}$ is an unbiased estimator of δ. Using arguments similar to ones in Ghosh and Meeden (1986), an estimator of τ is

$$\hat{\tau} = \max(0,\hat{\tau}_*) \tag{2.4}$$

where

$$\hat{\tau}_* = (\ell-1)\left[n_T - n_T^{-1}\sum_{i=1}^{\ell} n_i^2\right]^{-1}\left\{\frac{(\ell-1)\,BMS}{(\ell-3)\,WMS} - 1\right\}, \quad \ell > 3$$

Formula (2.4) is exactly the same as (2.8) in Ghosh and Meeden (1986). Thus, we use the estimator

$$\hat{\theta} = \sum_{i=1}^{\ell}(1-\hat{\omega}_i)\overline{Y}_i \Big/ \sum_{i=1}^{\ell}(1-\hat{\omega}_i), \quad \hat{\tau} > 0$$

$$= \overline{\overline{Y}}, \quad \hat{\tau} = 0$$

where $\hat{\omega}_i = (1+n_i\hat{\tau})^{-1}$. We need a separate estimator when $\hat{\tau} = 0$ because in this case $\hat{\omega}_i = 1$, $i = 1,2,\ldots,\ell$, and $\sum_{i=1}^{\ell}(1-\hat{\omega}_i)\overline{Y}_i / \sum_{i=1}^{\ell}(1-\hat{\omega}_i)$ is indeterminate. The second estimator is sensible because $\lim_{\hat{\tau}\to 0} \sum_{i=1}^{\ell}(1-\hat{\omega}_i)\overline{Y}_i / \sum_{i=1}^{\ell}(1-\hat{\omega}_i) = \overline{\overline{Y}}$.

Next we construct an estimator for η. Consider $\sum_{i=1}^{\ell} h_i(S_i^2-\hat{\delta})^2$ where $h_i = (n_i-1)(n_T-\ell)^{-1}$, $i = 1,2,\ldots,\ell$ and $\hat{\delta}$ is defined in (2.3). Then it is easy to show that

$$E\left[\sum_{i=1}^{\ell} h_i(S_i^2 - \hat{\delta})^2\right] = \delta^2(n_T-\ell)^{-1}\sum_{i=1}^{\ell}(1-h_i)\{2+(n_i+1)/(\eta-2)\} \quad \eta > 2.$$

Thus as an estimator of η we consider

$$\hat{\eta} = 2 + \{\max(\ell^{-c}, \hat{\eta}_*^{-1})\}^{-1} \tag{2.6}$$

where $c > 0$ and

$$\hat{\eta}_*^{-1} = \left\{2+(\ell-1)^{-1}(n_T-\ell)\left[1 - \sum_{i=1}^{\ell} h_i^2\right]\right\}^{-1}\left[\left\{(\ell-1)^{-1}(n_T-\ell)\sum_{i=1}^{\ell} h_i(S_i^2\hat{\delta}^{-1}-1)^2\right\} - 2\right].$$

It is necessary to have a truncated estimator of the form $\hat{\eta}$ in (2.6) because $\hat{\eta}^*$ could be negative and indeterminate. A suitable truncation point is a sequence in ℓ which vanishes as $\ell \to \infty$. A simple choice is ℓ^{-c}, $c > 0$ (e.g., $c = 1$ or $1/2$). However, we prefer to choose c in (2.6) in such a way that $\hat{\eta}_*^{-1}$ dominates ℓ^{-c} when $\hat{\eta}_*^{-1}$ is positive. A suitable choice of c can be obtained by a sampling experiment.

2.2 Asymptotic Properties

We present in Lemma 1 asymptotic properties of the point estimators of $\underset{\sim}{\rho} = (\theta,\delta,\tau,\eta)'$.

Lemma 1. Assume $\inf_{i\geq 1} n_i = 2$ and $\sup_{i\geq 1} n_i = k < \infty$. Then as $\ell \to \infty$

(a) $\hat{\delta} \xrightarrow{a.s.} \delta$ and $E(\hat{\delta}-\delta)^2 \longrightarrow 0$

(b) $\hat{\tau} \xrightarrow{a.s.} \tau$ and $\max_{i=1,2,\ldots,\ell} |\hat{\omega}_i-\omega_i| \xrightarrow{a.s.} 0$

(c) $\hat{\eta} \xrightarrow{a.s.} \eta$

(d) $\hat{\theta} \xrightarrow{a.s.} \theta$.

Proof. (a) Since $2 \leq \inf_{i\geq 1} n_i \leq \sup_{i\geq 1} n_i \leq k < \infty$ and $\mathrm{var}(S_i^2) = [(2+(n_i+1)/(\eta-2)]\delta^2/(n_i-1) \leq [(k+1)/(\eta-2) + 2]\delta^2 = A$, by the Kolmogorov strong law of large numbers (SLLN), $\hat{\delta} \xrightarrow{a.s.} \delta$; see Serfling (1980, pg. 27). Also since $E(\hat{\delta}-\delta)^2 \leq A\ell^{-1}$, $E(\hat{\delta}-\delta)^2 \longrightarrow 0$ as $\ell \longrightarrow \infty$.

(b) By using (a) and applying the SLLN to each term in $BMS = (\ell-1)^{-1}\{\sum_{i=1}^{\ell} n_i\overline{Y}_i^2 - n_T\overline{\overline{Y}}^2\}$, we have $\hat{\tau}^* \xrightarrow{a.s.} \tau$. Thus, by continuity, $\max(0,\hat{\tau}_*) \xrightarrow{a.s.} \tau$ as $\ell \to \infty$. Also since $|\hat{\omega}_i-\omega_i| \leq |1-\hat{\tau}\tau^{-1}|$ $i = 1,2,\ldots,\ell$, $\max_{i=1,2,\ldots,\ell} |\hat{\omega}_i-\omega_i| \xrightarrow{a.s.} 0$ as $\ell \to \infty$.

(c) Appendix A shows that $\hat{\eta}_*^{-1} \xrightarrow{a.s.} (\eta-2)^{-1}$ as $\ell\to\infty$. Since $\hat{\eta} = 2 + \{\max(\ell^{-c},\hat{\eta}_*^{-1})\}^{-1}$ by continuity, $\hat{\eta} \xrightarrow{a.s.} \eta$ as $\ell \to \infty$.

(d) If $\hat{\tau} = 0$, $\hat{\theta} - \theta = n_T^{-1}\sum_{i=1}^{\ell} n_i(\overline{Y}_i-\theta)$, and by SLLN, $\hat{\theta} - \theta \xrightarrow{a.s.} 0$ as $\ell\to\infty$.

If $\hat{\tau} > 0$, $|\hat{\theta}-\theta| = |\sum_{i=1}^{\ell}(1-\hat{\omega}_i)(\overline{Y}_i-\theta)| / \sum_{i=1}^{\ell}(1-\hat{\omega}_i)$

$$\leq (k+\hat{\tau}^{-1})\left\{\left|\ell^{-1}\sum_{i=1}^{\ell}(1-\omega_i)(\overline{Y}_i - \theta)\right| + \left\{\max_{i=1,2,\ldots,\ell}|\hat{\omega}_i-\omega_i|\right\}\ell^{-1}\sum_{i=1}^{\ell}|\overline{Y}_i-\theta|\right\}.$$

By SLLN $\ell^{-1}\sum_{i=1}^{\ell}(1-\omega_i)(\overline{Y}_i-\theta) \xrightarrow{a.s.} 0$ as $\ell \longrightarrow \infty$. Since $E(\ell^{-1}\sum_{i=1}^{\ell}|\overline{Y}_i-\theta|) \leq \delta\tau(1+k\tau)^{1/2} < \infty$, it follows that

$\ell^{-1}\sum_{i=1}^{\ell}|\overline{Y}_i-\theta|$ is finite a.e. Using lemma 1(b) and assuming $\hat{\tau}>0$, $\hat{\theta}-\theta\xrightarrow{a.s.}0$ as $\ell\to\infty$.

3. EMPIRICAL BAYES CONFIDENCE INTERVAL

In section 3 we approximate the HPD interval of $\gamma(\underset{\sim}{Y}_\ell)$ in (1.4) by an EB confidence interval. In section 3.1 we construct the EB confidence interval using a two–stage procedure. In section 3.2 we compare the EB with the HPD interval using asymptotic theory. Approximation, construction and comparisons are made under the marginal distribution of the Y_{ij} in (1.1), (1.2) and (1.3).

3.1 Construction of EB Interval

Suppose ρ is known. Then under the marginal distribution of the Y_{ij}

$$(\gamma(\underset{\sim}{Y}_\ell)-e_B)/\nu_B \sim t_{2\eta} \tag{3.1}$$

where e_B and ν_B are given in (1.4)

At the first stage we assume τ,δ,η are known, and consider the pivotal quantity

$$(\gamma(\underset{\sim}{Y}_\ell)-\hat{e}^*_B)/\nu_{Ba} \tag{3.2}$$

where

$$\hat{e}^*_B=\overline{Y}_\ell-(1-f_\ell)\omega_\ell(\overline{Y}_\ell-\hat{\theta}_*),$$

$$\nu^2_{Ba}=(1-\eta^{-1})\text{var}(\gamma(\underset{\sim}{Y}_\ell)-\hat{e}^*_B)=\nu^2_B+\nu^2_{\hat{\theta}_*},$$

$$\nu^2_{\hat{\theta}_*}=(1-\eta^{-1})\delta(1-f_\ell)^2\omega^2_\ell/\sum_{i=1}^{\ell}n_i\omega_i$$

while $\hat{\theta}_*$ is given by (2.2) and ν^2_B by (1.4). Acting as if the pivotal quantity in (3.2) has a Student t distribution with 2η degrees of freedom, an approximate 100(1–α)% confidence interval for $\gamma(\underset{\sim}{Y}_\ell)$ is

$$\hat{e}^*_B\pm(\nu^2_B+\nu^2_{\hat{\theta}_*})^{1/2}t_{2\eta,\alpha/2} \tag{3.3}$$

where $t_{2\eta,\alpha/2}$ is the 100(1–α/2) percentile point of the Student t distribution.

At the second stage we substitute estimators $\hat{\tau},\hat{\delta},\hat{\eta}$ from (2.3), (2.4), (2.6) into (3.3) to obtain the proposed 100(1–α)% EB confidence interval for $\gamma(\underset{\sim}{Y}_\ell)$

$$\hat{e}_B\pm\hat{\nu}_B t_{2\hat{\eta},\alpha/2}. \tag{3.4}$$

In (3.4),

$$\hat{e}_B=\overline{Y}_\ell-(1-f_\ell)\hat{\omega}_\ell(\overline{Y}_\ell-\hat{\theta})$$

$$\hat{\nu}^2_B=\hat{\nu}^2_B+\hat{\nu}^2_{\hat{\theta}_*}$$

$$\hat{\nu}^2_B=(1-\hat{\eta}^{-1})\hat{\delta}(1-f_\ell)\{f_\ell+(1-f_\ell)(1-\hat{\omega}_\ell)\}/n_\ell$$

and

$$\hat{\nu}^2_{\hat{\theta}_*}=(1-\hat{\eta}^{-1})\hat{\delta}(1-f_\ell)^2\hat{\omega}^2_\ell/\sum_{i=1}^{\ell}n_i\hat{\omega}_i.$$

3.2 Asymptotic Properties of EB Interval

Now we consider how well the EB interval (3.4) approximates the HPD interval (1.4) as $\ell\to\infty$. As a basis of the asymptotics, using the marginal distribution of the Y_{ij} obtained from (1.1), (1.2) and (1.3), we compare the centers, widths, and probability contents of the two intervals. Let W denote the width and P probability content of an interval. Then $W_B=2t_{2\eta,\alpha/2}\,\nu_B$ and $\hat{W}_B=2t_{2\hat{\eta},\alpha/2}\hat{\nu}_B$. Also letting

$$\hat{P}_B=\mathcal{T}\left\{(\hat{e}_B-e_B+t_{2\hat{\eta},\alpha/2}\hat{\nu}_B)\nu_B^{-1}\right\}-\mathcal{T}\left\{(\hat{e}_B-e_B-t_{2\eta,\alpha/2}\hat{\nu}_B)\nu_B^{-1}\right\},$$

the probability content of the EB interval is $P_B=E_{\underset{\sim}{Y}}(\hat{P}_B)$ where expectation is taken over the marginal distribution of $\underset{\sim}{Y}$ obtained from (1.1), (1.2) and (1.3) and $\mathcal{T}(\cdot)$ is the cumulative distribution function of a Student t on 2η degrees of freedom. The centers of the two intervals are e_B and $\hat{e}_B$, which are the Bayes and the empirical Bayes estimators of $\gamma(\underset{\sim}{Y}_\ell)$ respectively.

First, we present Lemma 2.

Lemma 2. Assume $\inf_{i\geq1} n_i=2$ and $\sup_{i\geq1} n_i=k<\infty$. Then as $\ell\to\infty$

(a) $\hat{\nu}^2_{\hat{\theta}_*}\xrightarrow{a.s.}0$ and $E(\hat{\nu}^2_{\hat{\theta}_*})\longrightarrow0$

(b) $\hat{\nu}_B-\nu_B\xrightarrow{a.s.}0$

(c) $E|\hat{\nu}_B-\nu_B|\to0.$

Proof. (a) Using Lemma 1 and the inequality $\hat{\nu}^2_{\hat{\theta}_*}\leq\hat{\delta}(1+k\hat{\tau})/2\ell$, $\hat{\nu}^2_{\hat{\theta}_*}\xrightarrow{a.s.}0$ as $\ell\to\infty$. Now $E(\hat{\nu}^2_{\hat{\theta}_*})\leq(E\hat{\delta}(1+k\hat{\tau}))/2\ell$. Thus by Lemma 1 again there exists $A<\infty$ s.t. $\hat{\delta}(1+k\hat{\tau})<A$ a.e. Thus $E(\hat{\nu}^2_{\hat{\theta}_*})\to0$ as $\ell\to\infty$.

(b) Using the triangular inequality $|\hat{\nu}^2_B-\nu^2_B|\leq\hat{\nu}^2_{\hat{\theta}_*}+|\hat{\nu}^2_B-\nu^2_B|$. Thus by Lemma 2(a) it is only required to show that $\hat{\nu}^2_B-\nu^2_B\xrightarrow{a.s.}0$ as $\ell\to\infty$. Using Lemma 1 and the inequality,

$$|\hat{\nu}^2_B-\nu^2_B|\leq|(1-\hat{\eta}^{-1})\hat{\delta}-(1-\eta^{-1})\delta|+\tfrac{1}{2}\hat{\delta}$$

$\max_{i=1,2,\ldots,\ell} |\hat{\omega}_i - \omega_i|$,

$\hat{\nu}_B^2 - \nu_B^2 \xrightarrow{a.s.} 0$ as $\ell \to \infty$.

(c) It is easy to show that $E|\hat{\nu}_B - \nu_B| \le \sqrt{3}\{E|\hat{\nu}_B^2 - \nu_B^2|\}^{1/2}$. Since $E|\hat{\nu}_B^2 - \nu_B^2| \le \{E(\hat{\nu}_B^2 - \nu_B^2)^2\}^{1/2} + E(\nu^2_{\hat{\theta}_*})$, by Lemma 2(a) it is only required to show that $E(\hat{\nu}_B^2 - \nu_B^2)^2 \to 0$ as $\ell \to \infty$; see Appendix C.

Theorem 1 gives a neat summary of our main results and it establishes that for a large number of small areas the EB interval is expected to be approximately the same as the HPD interval for the finite population mean.

Theorem 1

Assume $\inf_{i \ge 1} n_i = 2$ and $\sup_{i \ge 1} n_i = k < \infty$. Then as $\ell \to \infty$

(a) $E|\hat{e}_B - e_B| \longrightarrow 0$

(b) $E|\hat{W}_B - W_B| \longrightarrow 0$

(c) $E(\hat{P}_B) \longrightarrow 1-\alpha$.

Proof. (a) Since $E|\hat{e}_B - e_B| \le \{E(\hat{e}_B - e_B)^2\}^{1/2}$, we show that $\hat{e}_B - e_B \xrightarrow{a.s.} 0$ as $\ell \to \infty$ and $(\hat{e}_B - e_B)^2$ is uniformly integrable; see Serfling (1980, Section 1.4). Because $(\hat{e}_B - e_B) \le |\hat{\theta} - \theta| + |\bar{Y}_\ell - \theta| \max_{i=1,2,\ldots,\ell} |\hat{\omega}_i - \omega_i|$ and $|\bar{Y}_\ell - \theta|$ is finite a.e., by Lemma 1b,d $\hat{e}_B - e_B \xrightarrow{a.s.} 0$ as $\ell \to \infty$. Appendix B shows that $(\hat{e}_B - e_B)^2$ is uniformly integrable.

(b) It is easy to show

$$E|\hat{W}_B - W_B| \le E\left[|t_{2\hat{\eta},\alpha/2}||\hat{\nu}_B - \nu_B|\right] + \nu_B E|t_{2\hat{\eta},\alpha/2} - t_{2\eta,\alpha/2}|.$$

By using Lemma 1 and the continuity of the inverse cumulative distribution function of the Student t on a degrees of freedom (i.e., $\mathcal{T}_a^{-1}(1-\alpha/2)$ in a, any positive real number)

$t_{2\hat{\eta},\alpha/2} = \mathcal{T}_{2\hat{\eta}}^{-1}(1-\alpha/2) \xrightarrow{a.s.} \mathcal{T}_{2\eta}^{-1}(1-\alpha/2) = t_{2\eta,\alpha/2}$ as $\ell \to \infty$. But since $t_{2\eta,\alpha/2}, t_{2\hat{\eta},\alpha/2} \le t_{4,\alpha/2} = A < \infty$, $t_{2\hat{\eta},\alpha/2} - t_{2\eta,\alpha/2}$ is uniformly bounded and $E(t_{2\hat{\eta},\alpha/2} - t_{2\eta,\alpha/2}) \to 0$. Thus, by Lemma 2(c) $E|\hat{W}_B - W_B| \to 0$ as $\ell \to \infty$.

(c) By lemma 1(c)

$$\hat{P}_B \xrightarrow{a.s.} \mathcal{T}(t_{2\eta,\alpha/2}) - \mathcal{T}(-t_{2\eta,\alpha/2}) = 1-\alpha$$

and since $\hat{P}_B$ is uniformly bounded, $E(\hat{P}_B) \longrightarrow 1-\alpha$.

Finally, we present Corollary 1. The Bayes risk of any estimator, e, of $\gamma(\underset{\sim}{Y}_\ell)$ under squared error loss, r(e), is $r(e) = E_{\underset{\sim}{Y}}\{e - \gamma(\underset{\sim}{Y}_\ell)\}^2$ where expectation is taken over the marginal distribution of $\underset{\sim}{Y}$ obtained from (1.1), (1.2) and (1.3). As in Lemma 3 of Ghosh and Meeden (1986), we have $r(e) - r(e_B) = E(e - e_B)^2$.

Corollary 1

Under the conditions of Theorem 1,

$r(\hat{e}_B) - r(e_B) \longrightarrow 0$ as $\ell \longrightarrow \infty$.

The proof follows immediately from Theorem 1(a).

Corollary 1 shows that $\hat{e}_B$ is asymptotically optimal in the sense of Robbins (1955). This adds credence to the center of the EB interval as an approximation to the center of the HPD interval.

4. CONCLUDING REMARKS

Although our specification in (1.1), (1.2) and (1.3) is a simplification of the structure of a typical finite population, it extends the reslts in Ghosh and Meeden (1986) in three ways. First, here, the sampling variances are assumed to be unequal (with a common inverse gamma distribution). Second, an interval estimator, rather than a point estimator, is obtained. Third, we obtain for our estimators almost sure convergence rather than convergence in probability.

Also one can construct an interval estimator for the $\ell + 1^{st}$ area which has not been sampled. Thus assume observations are obtained from ℓ small areas, all $\ell + 1$ areas follow (1.1), (1.2) and (1.3), and interest is on $\gamma(\underset{\sim}{Y}_{\ell+1}) = \sum_{j=1}^{N_{\ell+1}} Y_{\ell+1,j}/N_{\ell+1}$ where $N_{\ell+1} \ge 1$. Then as an approximation to the 100(1−α)% HPD interval $\theta \pm \{\delta(1-\eta^{-1})\tau\}^{1/2} t_{2\eta,\alpha/2}$ for $\gamma(\underset{\sim}{Y}_{\ell+1})$ we have

$$\hat{\theta} \pm \sqrt{\hat{\delta}}\left\{(1-\hat{\eta}^{-1})\hat{\tau} + 1/\sum_{i=1}^{\ell} n_i\hat{\omega}_i\right\}^{1/2} t_{2\hat{\eta},\alpha/2}$$

and Theorem 1 still holds.

For further research, it is informative to carry out a sampling experiment to assess the coverage property of the EB interval for small to moderate number of areas. One can also assess the parametric point estimators, in particular, $\hat{\eta}$ of η. Comparison of the EB interval with other intervals under (1.1), (1.2) and (1.3) can also be made. One "natural" interval estimator of $\gamma(\underset{\sim}{Y}_\ell)$ is

$$\bar{Y}_\ell \pm \{(1-f_\ell)/n_\ell\}^{1/2} S_\ell t_{n_\ell-1,\alpha/2}$$

which uses data from only the area of interest. A second interval estimator, which "borrows strength," is

$$\bar{\bar{Y}} \pm \sqrt{\hat{\delta}}\,\{f_\ell n_\ell^{-1} + (1-2f_\ell)n_T^{-1} + (1-2n_\ell n_T^{-1} + n_T^{-2}\sum_{i=1}^{\ell} n_i^2)\hat{\tau}\}^{1/2} z_{\alpha/2}$$

where $z_{\alpha/2}$ is the 100(1−α/2) percentile point of the standard normal distribution. Comparisons can also be made with appropriate versions of the intervals given by Hulting and Harville (1991).

Appendix A. Proof of Lemma 1(c)

We show $\hat{\eta}_*^{-1} \xrightarrow{a.s.} (\eta-2)^{-1}$ as $\ell \to \infty$.

First note that

$$\hat{\eta}_*^{-1} = \delta^2\hat{\delta}^{-2}\{H_1(\delta,\hat{\delta}) - H_2(\delta,\hat{\delta}) + (\eta-2)^{-1}\} \quad (A1)$$

where

$$H_1(\delta,\hat{\delta}) = \delta^{-2}\left[\sum_{i=1}^{\ell} h_i(S_i^2 - \hat{\delta})^2 - E\left\{\sum_{i=1}^{\ell} h_i(S_i^2 - \hat{\delta})^2\right\}\right](n_T-\ell)(\ell-1)^{-1}A_\ell$$

$$H_2(\delta,\hat{\delta}) = 2(\hat{\delta}^2\delta^{-2}-1)A_\ell$$

and

$$A_\ell = (\ell-1)/\sum_{i=1}^{\ell}(1-h_i)(n_i+1).$$

Now both A_ℓ and $(n_T-\ell)(\ell-1)^{-1}$ are bounded. It follows by Lemma 1(a) $H_2(\delta,\hat{\delta}) \xrightarrow{a.s.} 0$ and $\delta^2\hat{\delta}^{-2} \xrightarrow{a.s.} 1$ as $\ell \to \infty$. Thus in (A1) we only need to show

$$\sum_{i=1}^{\ell} h_i(S_i^2 - \hat{\delta})^2 - E\left\{\sum_{i=1}^{\ell} h_i(S_i^2 - \hat{\delta})^2\right\} \xrightarrow{a.s.} 0 \text{ as } \ell \to \infty.$$

Now

$$\sum_{i=1}^{\ell} h_i(S_i^2 - \hat{\delta})^2 - E\left\{\sum_{i=1}^{\ell} h_i(S_i^2 - \hat{\delta})^2\right\} = Q_1 - Q_2 \quad (A2)$$

where

$$Q_1 = \sum_{i=1}^{\ell} h_i S_i^4 - E\left[\sum_{i=1}^{\ell} h_i S_i^4\right],$$

and

$$Q_2 = \hat{\delta}^2 - \delta^2 - \mathrm{var}(\hat{\delta}).$$

By SLLN, provided that $\eta > 4$, $Q_1 \xrightarrow{a.s.} 0$ as $\ell \to \infty$.

Also by lemma 1(a) $Q_2 \xrightarrow{a.s.} 0$ as $\ell \to \infty$. Thus $\hat{\eta}_*^{-1} \xrightarrow{a.s.} (\eta-2)^{-1}$ as $\ell \to \infty$.

Appendix B. Proof of the Uniform Integrability of $(\hat{e}_B - e_B)^2$

Since

$$(\hat{e}_B - e_B)^2 < 2\{(\bar{Y}_\ell-\theta)^2 + (\hat{\theta}-\theta)^2\} \quad (B1)$$

we show that $(\bar{Y}_\ell-\theta)^2$ and $(\hat{\theta}-\theta)^2$ are both uniformly integrable (u.i.).

First by (2.1),

$$(\bar{Y}_\ell-\theta)^2 \overset{d}{=} (1-\eta^{-1})\delta\tau(1-\omega_\ell)^{-1}F(1,2\eta) \quad (B2)$$

where $F(1,2\eta)$ has an f distribution. Then by (B2), recalling $\sup_{i\geq 1} n_i = k < \infty$

$$(\bar{Y}_\ell-\theta)^2 \overset{st}{\leq} \delta(k\tau+1)F(1,2\eta)/2$$

and since $\eta > 1$, $(\bar{Y}_\ell-\theta)^2$ is bounded by a random variable with finite expectation. Thus $(\bar{Y}_\ell-\theta)^2$ is u.i., see Serfling (1980, section 1.4). It follows that $\ell^{-1}\sum_{i=1}^{\ell}(\bar{Y}_i-\theta)^2$ is also u.i.

Second

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

$$2\left\{\sum_{i=1}^{\ell}(1-\hat{\omega}_i)(\bar{Y}_i-\theta)/\sum_{i=1}^{\ell}(1-\hat{\omega}_i)\right\}^2. \quad (B3)$$

Using (B3) it is easy to show that

$$(\hat{\theta}-\theta)^2 \leq k^2\ell^{-1}\sum_{i=1}^{\ell}(\bar{Y}_i-\theta)^2. \quad (B4)$$

Then because $\ell^{-1}\sum_{i=1}^{\ell}(\bar{Y}_i-\theta)^2$ is u.i., by (B4) $(\hat{\theta}-\theta)^2$ is u.i..

Appendix C. Completion of proof of Lemma 2(c)

By Minkowski's inequality

$$(E(\hat{\nu}_B^2 - \nu_B^2)^2 \leq \Big[2\{E((1-\hat{\eta}^{-1})\hat{\delta}-(1-\eta^{-1})\delta)^2\}^{1/2}$$

$$+ \{E((1-\hat{\eta}^{-1})\hat{\delta}\hat{\omega}_\ell - (1-\eta^{-1})\delta\omega_i)^2\}^{1/2}\Big]^2 \quad (C1)$$

Thus we show that each term on the right hand side of (C1) $\to 0$ as $\ell \to \infty$.

First, by Minkowski's inequality

$$E[(1-\hat{\eta}^{-1})\hat{\delta} - (1-\eta^{-1})\delta]^2 \leq [\{E(\hat{\delta}-\delta)^2\}^{1/2}$$

$$+ \delta\{E(\hat{\eta}^{-1}-\eta^{-1})^2\}^{1/2}]^2.$$

Thus by Lemma 1(a) we only need to show

$$E(\hat{\eta}^{-1} - \eta^{-1})^2 \to 0 \text{ as } \ell \to \infty. \quad (C2)$$

Since $|\hat{\eta}^{-1} - \eta^{-1}| \leq 1$ by Lemma 1(c)

$E(\hat{\eta}^{-1} - \eta^{-1})^2 \to 0$ as $\ell \to \infty$.

Second, using Minkowski's inequality twice

$$E[(1-\hat{\eta}^{-1})\hat{\delta}\hat{\omega}_\ell - (1-\eta^{-1})\delta\omega_\ell]^2$$

$$\leq \Big[\{E((\hat{\eta}^{-1}-\eta^{-1})^2\hat{\delta}^2)\}^{1/2} + \{E(\hat{\delta}-\delta)^2\}^{1/2}$$

$$+ \delta\{E(\hat{\omega}_\ell-\omega_\ell)^2\}^{1/2}\Big]^2.$$

Thus by Lemma 1(a) and (C2) we must show

$$E(\hat{\omega}_\ell - \omega_\ell)^2 \to 0 \text{ as } \ell \to \infty. \quad (C3)$$

Since $|\hat{\omega}_\ell - \omega_\ell| \leq 1$ and $\max_{i=1,2,\ldots,\ell}|\hat{\omega}_i - \omega_i| \xrightarrow{a.s.} 0$ as

$\ell \to \infty$, (C3) follows.

ACKNOWLEDGEMENT
The author is grateful to Professor Joseph Sedransk for his generous recommendations concerning the presentation.

REFERENCES

Battese, G.E., Harter, R.M. and Fuller, W.A. (1988), "An Error Components Model for Prediction of County Crop Areas Using Survey and Satellite Data," Journal of the American Statistical Association, 83, 28–36.

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

Dempster, A.P., Rubin, D.B. and Tsutakawa, R.K. (1981), "Estimation in Covariance Components Models," Journal of the American Statistical Association, 76, 341–353.

Fay, R.E. and Herriot, R.A. (1979), "Estimates of Income for Small Places: An Application of James–Stein Procedures to Census Data," Journal of the American Statistical Association, 74, 269–277.

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

Hulting, F.L. and Harville, D.A. (1991), "Some Bayesian and Non–Bayesian Procedures for the Analysis of Comparative Experiments and for Small–Area Estimation: Computational Aspects, Frequentists Properties, and Relationships," Journal of the American Statistical Association 86, 557–568.

Laird, N.M. and Louis, T.A. (1987), "Empirical Bayes Confidence Intervals Based on Bootstrap Samples" (with comments), Journal of the American Statistical Association, 82, 739–757.

————— (1989), "Empirical Bayes Confidence Intervals for a Series of Related Experiments," Biometrics, 45, 481–495.

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, eds. G.E.P. Box, T. Leonard, and C.F. Wu, New York: Academic Press, pp. 25–49.

Nandram, B. and Sedransk, J. (1992), "Empirical Bayes Estimation for the Finite Population Mean on the Current Occasion," Journal of the American Statistical Association (accepted), 1–24.

Prasad, N.G.N. and Rao J.N.K. (1990), "The Estimation of the Mean Squared Error of Small Area Estimators," Journal of the American Statistical Association, 85, 163–171.

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, pp. 159–163.

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