한국통계학회논문집 제2권 2호
1995년 12월 pp. 63-73

Robust Bayes and Empirical Bayes Analysis in Finite Population Sampling

Dal Ho Kim[1]

Abstract

We consider some robust Bayes estimators using ML-II priors as well as certain empirical Bayes estimators in estimating the finite population mean. The proposed estimators are compared with the sample mean and subjective Bayes estimators in terms of "posterior robustness" and "procedure robustness".

1. Introduction

Consider a finite population U with units labeled $1,2,\ldots,N$. Let y_i denote the value of a single characteristic attached to the unit i. The vector $y=(y_1,\ldots,y_N)^T$ is the unknown state of nature, and is assumed to belong to $\Theta=R^N$. A subset s of $\{1,\ldots,N\}$ is called a sample. Let $n(s)$ denote the number of elements belonging to s. The set of all possible samples is denoted by S. A design is a function p on S such that $p(s)\in[0,1]$ for all $s\in S$ and $\sum_{s\in S}p(s)=1$. Given $y\in\Theta$ and $s=\{i_1,\ldots,i_{n(s)}\}$ with $1\le i_1<\ldots<i_{n(s)}\le N$, let $y(s)=\{y_{i_1},\ldots,y_{i_{n(s)}}\}$. One of the main objectives in sample surveys is to draw inference about y or some function (real or vector valued) $\gamma(y)$ of y on the basis of s and $y(s)$.

A unified and elegant formulation of Bayes estimation in finite population sampling was given by Ericson(1969). Since then, there are many papers in the area of Bayes estimation in finite population sampling. However, most of Bayesian literature in survey sampling deals with subjective Bayesian analysis in that the inference procedure is based on a single completely specified prior distribution. Such an 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. We shall see in Section 2 that even in simple examples, failure to specify accurately one or more of parameters of a prior distribution has a serious consequence even from a Bayesian viewpoint, and protection should be taken against such phenomena.

A *robust Bayesian viewpoint* assumes only that subjective information can be quantified

1) Department of Statistics, Kyungpook National University, Taegu, 702-701, KOREA.

only in terms of a class Γ of possible distributions. Inference and decision should be relatively insensitive to deviations as the prior distribition varies over Γ. The robust Bayes idea can be traced back to Good as early as in 1950 (see for example Good(1965)), and has been popularized recently, notably by the stimulating article by Berger(1984). There is a rapidly growing literature in the area of robust Bayesian analysis. Berger(1990) and Wasserman(1992) provide reviews and discussion of the various issues and approaches.

The need for robust Bayesian analysis in survey sampling has been felt by some authors. Godambe and Thompson(1971) adapted a framework whereby the prior information could only be quantified up to a class Γ (C in their notation) of prior distributions. The model assumption there played a very minimal role, the main idea being that model-based inference statements could be replaced, in case of model-failure by design-based inference. In a later study, Godambe(1982) considered the more common phenomenon of specific departures from the assumed model. His contention there was that sampling designs could be a useful companion of model-based inference precedures to generate "near-optimal" and "robust" estimators. However, the basic model assumed in that paper considered $y_1,\dots,y_N$ to be independent, and attention was confined only to design and model unbiased estimators. Royall and Pfefferman(1982) also considered robustness of certain Bayes estimators. However, their main concern is to find out conditions under which the Bayes estimators under an assumed model remain the same under departures from the model.

The purpose of this article is to generate certain robust Bayes and empirical Bayes estimators from a different perspective, and study their performance over a broad class of prior distributions on the parameter space. Specifically, we develop some robust Bayes estimators of the population mean employing ML-II priors (see Berger and Berliner(1986)). Also, certain empirical Bayes estimators are derived. We compare the performance of these estimators with the performance of the subjective Bayes and the classical (sample mean) estimators using the criteria of "posterior robustness" and "procedure robustness". For simplicity of exposition, we have restricted ourselves to the estimation of the population mean, although the methods are equally applicable for estimating other parameters of interest.

For simplicity, in the subsequent sections, only the case when $n(s)\neq n \rightarrow p(s)=0$ is considered. This amounts to considering only fixed samples of size n. Also, throughout the loss is assumed to be squared error.

2. Main Results

We consider the model $y_i=\theta+\varepsilon_i$ $(i=1,\dots,N)$, where θ, $\varepsilon_1,\dots,\varepsilon_N$ are independently distributed with $\theta \sim N(\mu_0,\sigma_0^2)$ and $\varepsilon_1,\dots,\varepsilon_N$ are iid $N(0,\tau^2)$. Write $M_0=\tau^2/\sigma_0^2$,

$B_0=M_0/(M_0+n)$, $\bar{y}(s)=n^{-1}\sum_{i\in s}y_i$, and $y(\bar{s})=\{y_i : i\notin s\}$, the suffixes in $y(\bar{s})$ being arranged in ascending order. From Ericson(1969), it follows that the posterior distribution of $y(\bar{s})$ given s and $y(s)$ is $N((B_0\mu_0+(1-B_0)\bar{y}(s))\,\mathbf{1}_{N-n},\ \tau^2(\,I_{N-n}+(M_0+n)^{-1}J_{N-n}))$, where $\mathbf{1}_u$ denotes a u-component column vector with all elements equal to 1, $J_u=\mathbf{1}_u\ \mathbf{1}_u^T$ and I_u denotes the identity matrix of order u. Then the Bayes estimator of $\gamma(y)=N^{-1}\sum_{i=1}^{N}y_i$ is

$$\delta^0(s,y(s))=E[\gamma(y)\ |s,y(s)]=N^{-1}[n\bar{y}(s)+(N-n)(B_0\mu_0+(1-B_0)\bar{y}(s))]. \tag{2.1}$$

The classical estimator of $\gamma(y)$ is $\delta^C(s,y(s))=\bar{y}(s)$ which is an unbiased estimator of $\gamma(y)$ under any model which assume that the y_i's have a common mean.

To derive robust Bayes estimator of $\gamma(y)$, we first introduce the notion of ε-contaminated priors. Denote by π_0 the $N(\mu_0,\sigma_0^2)$ distribution. The class Γ_* of prior distributions is given by

$$\Gamma_*=\{\pi\ :\ \pi=(1-\varepsilon)\pi_0+\varepsilon q,\ q\in Q\}, \tag{2.2}$$

where $0\le\varepsilon\le1$ is given, and Q is the class of all distribution functions. We denote by $m(y\ |\pi)$ the marginal (predictive) density of y under the prior π. If $\pi\in\Gamma_*$, we can write

$$m(y\ |\pi)=(1-\varepsilon)\ m(y\ |\pi_0)\ +\ \varepsilon\ m(y\ |q).$$

This leads to

$$m(y(s)\ |\pi)=(1-\varepsilon)\ m(y(s)\ |\pi_0)\ +\ \varepsilon\ m(y(s)\ |q).$$

Our objective is to choose the prior π which maximizes $m(y(s)\ |\pi)$ over Γ_*. This amount to maximization of $m(y(s)\ |q)$ over $q\in Q$. Noting that

$$m(y(s)\ |q)=\int_{-\infty}^{\infty}(2\pi)^{-\frac{n}{2}}\exp[-\sum_{i\in s}(y_i-\theta)^2/(2\tau^2)]q(d\theta),$$

it follows that $m(y(s)\ |q)$ is maximized with respect to the prior which is degenerate at $\bar{y}(s)$. We shall denote this prior by $\delta_{\bar{y}(s)}$, δ being the dirac-delta function. The resulting (estimated) prior is now given by

$$\hat{\pi}=(1-\varepsilon)\ \pi_0\ +\ \varepsilon\delta_{\bar{y}(s)}. \tag{2.3}$$

The prior $\hat{\pi}$ is called ML-II prior by Good(1965), and Berger and Berliner(1986). Using the prior $\hat{\pi}$, $y(s)$ is marginally distributed as

$$(1-\varepsilon)N(\mu_0\ \mathbf{1}_n,\ \tau^2\ I_n+\sigma_0^2\ J_n)+\varepsilon F_s$$

where F_s has (improper) pdf

$$f_s(y(s)) = (2\pi\tau^2)^{-\frac{n}{2}} \exp[-\sum_{i\in s}(y_i-\bar{y}(s))^2/(2\tau^2)] .$$

Also, the conditional distribution of $y(\bar{s})$ given $(s, y(s))$ is given by

$$\lambda_{ML}(\bar{y}(s))N((B_0 \mu_0+(1-B_0)\bar{y}(s)) \mathbf{1}_{N-n}, \tau^2(\mathbf{I}_{N-n}+(M_0+n)^{-1} \mathbf{J}_{N-n}))$$
$$+(1-\lambda_{ML}(\bar{y}(s)))N(\bar{y}(s) \mathbf{1}_{N-n}, \tau^2 \mathbf{I}_{N-n}) \quad (2.4)$$

where

$$\lambda_{ML}(\bar{y}(s))=\{1+(\varepsilon/(1-\varepsilon))B_0^{-1/2}\exp(nB_0(\bar{y}(s)-\mu_0)^2/(2\tau^2))\}^{-1}.$$

Hence, under the posterior distribution given in (2.4), the Bayes estimator of $\gamma(y)$ is given by

$$\delta^{RB}(s,y(s))=N^{-1}[n\bar{y}(s)+ (N-n)\{\lambda_{ML}(\bar{y}(s))(B_0\mu_0+(1-B_0)\bar{y}(s))$$
$$+(1-\lambda_{ML}(\bar{y}(s)))\bar{y}(s)\}]. \quad (2.5)$$

Note that for ε close to zero, i.e. when one is very confident about the $N(\mu_0,\sigma_0^2)$ prior for θ, since $\lambda_{ML}(\bar{y}(s))$ is close to 1, it follows that δ^{BR} is very close to δ^0. On the other hand, when ε is close to 1, i.e. there is very little confidence in the $N(\mu_0,\sigma_0^2)$ prior, one gets $\lambda_{ML}(\bar{y}(s))$ very close to the sample mean $\bar{y}(s)$. Thus, the robust Bayes estimator serves as a compromise between the subjective Bayes and the classical estimators.

Next, we compare the performances of δ^0, δ^C and δ^{RB} from the robustness perspective. The main idea is that we want to examine whether these estimators perform satisfactorily over a broad class of priors.

With this end, for a given prior ξ, denote by $\rho(\xi,(s,y(s)),a)$ the posterior risk of an estimator $a(s,y(s))$ of $\gamma(y)$ i.e. $\rho(\xi,(s,y(s),a)=E[\{a(s,y(s))-\gamma(y)\}^2|s,y(s)]$. The following definition is taken from Berger (1984).

Definiton 2.1. An estimator $a_0(s,y(s))$ is ζ-posterior robust with respect to Γ if for the observed $(s,y(s))$,

$$POR_\Gamma(a_0)=\sup_{\xi\in\Gamma} |\rho(\xi,(s,y(s)),a_0)-\inf_{a\in R} \rho(\xi,(s,y(s)),a)| \le \zeta. \quad (2.6)$$

We shall, henceforth, refer to the left hand side of (2.6) as the "posterior robustness index" of the estimator $a_0(s,y(s))$ of $\gamma(y)$ under the class of priors Γ.

Note that $POR_\Gamma(a_0)$ in a sense is the sensitivity index of the estimator a_0 of $\gamma(y)$ as the prior varies over Γ. For any given $\zeta>0$, it is very clear that whether or not posterior robustness exists will often depend on which $(s,y(s))$ is observed. This will be revealed in the examples to follow.

To examine the posterior robustness of δ^0,δ^C and δ^{RB}, we consider the class of $N(\mu,\sigma^2)$ priors, μ (real) and $\sigma^2(>0)$. Write $M=\tau^2/\sigma^2$, $B=M/(M+n)$, where $\tau^2(>0)$ is known. Calculations similar to (2.1) now give the Bayes estimator of $\gamma(y)$ under the $N(\mu,\sigma^2)$ prior (to be denoted by $\xi_{\mu,B}$) as

$$\delta^{\mu,B}(s,y(s))=N^{-1}[n\bar{y}(s)+(N-n)(B\mu+(1-B)\bar{y}(s))]. \tag{2.7}$$

Then the following results hold.

$$\rho(\xi_{\mu,B},(s,y(s)),\delta^{\mu,B})=N^{-2}(N-n)\tau^2(M+N)/(M+n); \tag{2.8}$$

$$\rho(\xi_{\mu,B},(s,y(s)),\delta^0)-\rho(\xi_{\mu,B},(s,y(s)),\delta^{\mu,B})$$
$$=(1-f)^2[B_0(\mu-\mu_0)+(B_0-B)(\bar{y}(s)-\mu)]^2; \tag{2.9}$$

$$\rho(\xi_{\mu,B},(s,y(s)),\delta^C)-\rho(\xi_{\mu,B},(s,y(s)),\delta^{\mu,B})$$
$$=(1-f)^2B^2(\bar{y}(s)-\mu)^2; \tag{2.10}$$

$$\rho(\xi_{\mu,B},(s,y(s)),\delta^{RB})-\rho(\xi_{\mu,B},(s,y(s)),\delta^{\mu,B})$$
$$=(1-f)^2[B_0\lambda_{ML}(\bar{y}(s))(\bar{y}(s)-\mu_0)-B(\bar{y}(s)-\mu)]^2, \tag{2.11}$$

where $f=n/N$ is the sampling fraction.

From (2.9) - (2.11) it is very clear that if we consider the class Γ of all $N(\mu,\sigma^2)$ priors, for each one of the estimators δ^0,δ^C and δ^{RB}, the supremum (over μ (real)) of the left hand side of (2.9) - (2.11) becomes $+\infty$, and all these estimators turn out to be non-robust. One reason why this happens is that the $N(\mu,\sigma^2)$ class of priors for all real μ and $\sigma^2(>0)$ is indeed too big to be practically useful. As a next step, we consider the smaller class $N(\mu_0,\sigma^2)$ of priors, where the mean μ_0 is specified. This is not too unrealistic since, very often from prior experience, one can make a reasonable guess at the center of the distribution.

Note that when $\mu=\mu_0$, denoting $\xi_{\mu_0,B}$ by ξ_B and $\delta^{\mu_0,B}$ by δ^B, (2.9) - (2.11) simplify to

$$\rho(\xi_B,(s,y(s)),\delta^0)-\rho(\xi_B,(s,y(s)),\delta^B)=(1-f)^2(B_0-B)^2(\bar{y}(s)-\mu_0)^2; \tag{2.12}$$

$$\rho(\xi_B,(s,y(s)),\delta^C)-\rho(\xi_B,(s,y(s)),\delta^B)=(1-f)^2B^2(\bar{y}(s)-\mu_0)^2; \tag{2.13}$$

$$\rho(\xi_B,(s,y(s)),\delta^{RB})-\rho(\xi_B,(s,y(s)),\delta^B)=(1-f)^2(B_0\lambda_{ML}(\bar{y}(s))-B)^2(\bar{y}(s)-\mu_0)^2. \tag{2.14}$$

Accordingly, from (2.12) - (2.14),

$$POR(\delta^0) = (1-f)^2 \max(B_0^2, (1-B_0)^2)(\bar{y}(s)-\mu_0)^2. \tag{2.15}$$

$$POR(\delta^C) = (1-f)^2(\bar{y}(s)-\mu_0)^2; \tag{2.16}$$

$$POR(\delta^{RB}) = (1-f)^2 \max(B_0^2\lambda_{ML}^2(\bar{y}(s)), (1-B_0\lambda_{ML}(\bar{y}(s)))^2)(\bar{y}(s)-\mu_0)^2. \tag{2.17}$$

Thus, given any $\zeta > 0$ and $0 < f < 1$, the posterior ζ-robustness of all these procedures depends on the closeness of the sample mean to the prior mean μ_0. Also, it follows from (2.15) - (2.17) that both the subjective and robust Bayes estimators are more posterior robust than the sample mean for the $\{N(\mu_0,\sigma^2),\sigma^2>0\}$ class of priors. A comparison between (2.15) - (2.17) also reveals that the robust Bayes estimator arrived at by employing the ML-II prior enjoys a greater degree of posterior robustness than the subjective Bayes estimator if $B_0\lambda_{ML}(\bar{y}(s)) > 1/2$

We now turn to the empirical Bayes analysis. The empirical Bayes analysis is closely related to the robust Bayes analysis in the sense that in the former analysis, the prior distribution is assumed to belong to some class Γ of distributions. Contrary to the robust Bayes analysis where ε is typically taken to be very small in (2.2), in an empirical Bayes analysis ε is taken as 1. This point is very clearly brought out in Berger and Berliner(1984).

In order to derive the empirical Bayes estimator, we consider the model (2.2) with $\varepsilon=1$, but assume that Q is the class of $\{N(\mu_0,\sigma^2),\sigma^2>0\}$ priors. For a typical member, say $N(\mu_0,\sigma^2)$ in this class, the Bayes estimator of $\gamma(y)$ is given in (2.7). We now estimate B as follows.

Note that marginally $\bar{y}(s)$ is sufficient for σ^2 and $\bar{y}(s) \sim N(\mu_0,\tau^2/nB)$. Hence, $n(\bar{y}(s)-\mu_0)^2 \sim (\tau^2/B)\chi_1^2$, and so $E[n(\bar{y}(s)-\mu_0)^2]=\tau^2/B$. Hence, since τ^2 is known, a sensible estimator of B is $\tau^2/(n(\bar{y}(s)-\mu_0)^2)$. Since $0<B<1$ and this estimator can take values bigger than 1 with positive probability, we estimate B by

$$\hat{B} = \min\{1, \tau^2/(n(\bar{y}(s)-\mu_0)^2)\}. \tag{2.18}$$

The corresponding empirical Bayes estimator of $\gamma(y)$ is

$$\delta^{EB}(s,y(s)) = N^{-1}[n\bar{y}(s)+(N-n)(\hat{B}\mu_0+(1-\hat{B})\bar{y}(s))]. \tag{2.19}$$

Also,

$$\rho(\xi_B,(s,y(s)),\delta^{EB}) - \rho(\xi_B,(s,y(s)),\delta^B)$$
$$= (1-f)^2(\hat{B}-B)^2(\bar{y}(s)-\mu_0)^2, \tag{2.20}$$

so that

$$POR(\delta^{EB}) = (1-f)^2 \max(\widehat{B}^2, (1-\widehat{B})^2)(\overline{y}(s)-\mu_0)^2. \tag{2.21}$$

Thus, the empirical Bayes estimator also has greater posterior robustness index than the sample mean. The performance of δ^{EB} in comparison with δ^0 or δ^{RB} depends on the actual $(s,y(s))$ observed.

Although the Bayesian thinks conditionally on $(s,y(s))$, and accordingly in terms of posterior robustness, it is certainly possible to use the overall Bayes risk in defining a suitable robustness criterion. This may not be totally unappealing even to a Bayesian at the preexperimental stage when he perceives that y will be occurring according to a certain marginal distribution. The overall performance of an estimator $a(s,y(s))$ of $\gamma(y)$ when the prior is ξ is evaluated by $r(\xi,a) = E[\rho(\xi,(s,y(s)),a)]$, the expectation being taken with respect to the resulting marginal distribution of $y(s)$. The following method of measuring the robustness of a procedure is given in Berger(1984).

Definition 2.2. An estimator $a_0(s,y(s))$ of $\gamma(y)$ is said to be ζ-procedure robust with respect to Γ if

$$PR(a_0) = \sup_{\xi\in\Gamma} |r(\xi,a_0) - \inf_{a\in R} r(\xi,a)| < \zeta. \tag{2.22}$$

We shall henceforth refer to $PR(a_0)$ as the "procedure robustness index" of a_0.

Next we examine how the four estimators $\delta^0, \delta^C, \delta^{RB}$ and δ^{EB} compare according to the PR criterion as given (2.22) when we consider the $\{N(\mu_0,\sigma^2), \sigma^2>0\}$ class of priors. Using the same notation ξ_B as before for the $N(\mu_0,\sigma^2)$ prior, from (2.9) - (2.11), it follows that

$$r(\xi_B,\delta^0) - r(\xi_B,\delta^B) = (1-f)^2 (B_0-B)^2\tau^2/(nB); \tag{2.23}$$

$$r(\xi_B,\delta^C) - r(\xi_B,\delta^B) = (1-f)^2 B\tau^2/n; \tag{2.24}$$

$$r(\xi_B,\delta^{RB}) - r(\xi_B,\delta^B) = (1-f)^2 E[(B_0\lambda_{ML}(\overline{y}(s))-B)^2(\overline{y}(s)-\mu_0)^2]; \tag{2.25}$$

$$r(\xi_B,\delta^{EB}) - r(\xi_B,\delta^B) = (1-f)^2 E[(\widehat{B}-B)^2(\overline{y}(s)-\mu_0)^2]. \tag{2.26}$$

It is clear from (2.23) - (2.26) that

$$PR(\delta^{EB}) = (1-f)^2 \sup_{0<B<1} E[(\widehat{B}-B)^2(\overline{y}(s)-\mu_0)^2]. \tag{2.30}$$

From (2.27) and (2.28) it is clear that the subjective Bayes estimator lacks completely in procedure robustness, while the sample mean is quite procedure robust. To examine the procedure robustness of δ^{RB}, we proceed as follows.

Theorem 2.1. For every $\varepsilon>0$, $E[(B_0\lambda_{ML}(\bar{y}(s))-B)^2(\bar{y}(s)-\mu_0)^2]=O_e(B^{1/2})$, where O_e denote the exact order.

Proof. Noting that $n(\bar{y}(s)-\mu_0)^2 \sim (\tau^2/B)\chi_1^2$, it follows from (2.29) that

$$E\,[\{B_0\lambda_{ML}(\bar{y}(s))-B\}^2(\bar{y}(s)-\mu_0)^2]$$

$$=\int_0^\infty \left(\frac{B_0}{1+g\exp(B_0u/2B)}-B\right)^2\frac{\tau^2}{nB}u\exp\left(-\frac{u}{2}\right)\frac{u^{1/2-1}}{2^{1/2}\Gamma(1/2)}du, \qquad (2.31)$$

where $g=(\varepsilon/(1-\varepsilon))B_0^{-1/2}$. Next observe that

$$rhs\ of\ (2.31) \le \int_0^\infty \left[\frac{B_0^2}{\{1+g\exp(B_0u/2B)\}^2}+B^2\right]\frac{\tau^2}{nB}u\exp\left(-\frac{u}{2}\right)\frac{u^{1/2-1}}{2^{3/2}\Gamma(3/2)}du$$

$$\le \frac{\tau^2}{n}\left[\frac{B_0^2}{2g}B^{-1}\int_0^\infty \exp\left(-\frac{u}{2}(1+B_0B^{-1})\right)\frac{u^{3/2-1}}{2^{3/2}\Gamma(3/2)}du + B\right]$$

$$= \frac{\tau^2}{n}[B^{-1}(B_0^2/2g)(1+B_0B^{-1})^{-3/2}+B]$$

$$= O(B^{1/2}). \qquad (2.32)$$

Again, writing $g' = \max(g,1)$,

$$rhs\ of\ (2.31) \ge \frac{\tau^2}{n}\int_0^\infty \left[\frac{B_0^2B^{-1}}{\{2g'\ \exp(B_0u/2B)\}^2}-\frac{2B_0}{g\exp(-uB_0/2B)}+B\right]$$

$$\exp\left(-\frac{u}{2}\right)\frac{u^{3/2-1}}{2^{3/2}\Gamma(3/2)}du$$

$$= \frac{\tau^2}{n}[(B_0/2g')^2B^{-1}(1+2B_0B^{-1})^{-3/2}-2B_0g^{-1}(1+B_0B^{-1})^{-3/2}+B]$$

$$= O(B^{1/2}). \qquad (2.33)$$

Combining (2.32) and (2.33) the result follows.

Thus, unlike the subjective Bayes estimator, the robust Bayes estimator using the ML-II prior dose not suffer from lack of procedure robustness. For small B, the classical estimator δ^C has a certain edge over δ^{RB} from the PR index. This is, however, natural to expect since small B signifies small variance ratio τ^2/σ^2 which amounts to instability in the assessment of the prior distribution of θ. It is not surprising that in such a situation, it is safer to use $\bar{y}(s)$ in estimating $\gamma(y)$ if one is seriously concerned about long-run performance of an estimator.

Next we study $PR(\delta^{EB})$. First write $MSB=n(\bar{y}(s)-\mu_0)^2$, with $MSB \sim (\tau^2/B)\chi_1^2$. We prove the theorem.

Theorem 2.2. $E\,[(\hat{B}-B)^2(\bar{y}(s)-\mu)^2]=O(B^{1/2-\eta})$ where η $(0<\eta<1/2)$ can be made arbitrarily small.

Proof. First using (2.18) one gets,

$$E\,[(\hat{B}-B)^2(\bar{y}(s)-\mu)^2]=n^{-1}E\,[(\hat{B}-B)^2 MSB]$$

$$= n^{-1}[(1-B)^2\tau^2 P(\chi_1^2<B) + B^2(\tau^2/B)E\,[(1/\chi_1^2-1)^2\chi_1^2\, I_{[\chi_1^2>B]}]. \quad (2.34)$$

Now,

$$P(\chi_1^2<B)=\int_0^B \exp(-u/2)\frac{u^{-1/2}}{2^{1/2}\Gamma(1/2)}du$$

$$\le \int_0^B \frac{u^{-1/2}}{2^{1/2}\Gamma(1/2)}du = (2/\pi)^{1/2}B^{1/2}. \quad (2.35)$$

Moreover,

$$E\,[(1/\chi_1^2-1)^2\chi_1^2\, I_{[\chi_1^2>B]}]$$

$$=\int_B^\infty \frac{1}{u}e^{-\frac{u}{2}}\frac{u^{1/2-1}}{2^{1/2}\Gamma(1/2)}du-2P(\chi_1^2>B)+\int_B^\infty ue^{-\frac{u}{2}}\frac{u^{-1/2}}{2^{1/2}\Gamma(1/2)}du$$

$$\le B^{-1/2-\eta}\int_B^\infty e^{-u/2}\frac{u^{\eta-1}}{2^{1/2}\Gamma(1/2)}du+\int_B^\infty \frac{e^{-u/2}u^{1/2}}{2^{3/2}\Gamma(3/2)}du$$

$$\le B^{-1/2-\eta}\frac{\Gamma(\eta)}{\sqrt{2\pi}}+1. \quad (2.36)$$

Combining (2.34) - (2.36) the result follow.

Theorem 2.1 and 2.2 clearly demonstrate the procedure robustness of δ^{RB} and δ^{EB} as $B \rightarrow 0$. The superior performance of the *POR* index of these estimators relative to the *POR* index of δ^C has already been established. On the other hand δ^0 which performs better than δ^C on its *POR* index, fails miserably in its procedure robustness. This also shows that the average long-term performance of a procedure can sometimes be highly misleading when used as a yardstick for measuring its performance in a given situation.

In practice, however, τ^2 is not usually known. In such a situation, one can conceive of an inverse gamma prior for τ^2 independent of the prior for θ to derive a Bayes estimator of $\gamma(y)$ (see Ericson(1969)). In a robust Bayes approach, if one assumes a mixture of a normal-gamma prior and a ML-II prior for (θ,τ^2), then the ML-II prior puts its entire mass on the point $(\overline{y}(s),\ \sum_{i\in s}(y_i-\overline{y}(s))^2/n)$. In an empirical Bayes approach, one estimates τ^2/B once again by $MSB=n(\overline{y}(s)-\mu_0)^2$, and τ^2 by $MSW=\sum_{i\in s}(y_i-\overline{y}(s))^2/(n-1)$. The latter can be justified on the ground that $\sum_{i\in s}(y_i-\overline{y}(s))^2 \sim \tau^2\chi^2_{n-1}$ so that $E[MSW]=\tau^2$. Hence, an estimator of B is given in this case by $\widehat{B_*}=\min(1,\ MSW/MSB)$, and the corresponding empirical Bayes estimator is obtained by substituting $\widehat{B_*}$ for $\widehat{B}$ in (2.18). We have not studied the robustness of these estimators.

3. An Example

This section concerns the analysis of real data set from *Fortune* magazine, May 1975 and May 1976 to illustrate the methods suggested in Section 2. The data set consists of 331 corporations in US with 1975 gross sales, in billions, between one-half billion and ten billion dollars. For the complete finite population, we find the population mean to be 1.7283 and the population variance 2.2788. The population variance is assumed to be known for us. We select 10% simple random sample without replacement from this finite population. So the sample size is n=33. We use gross sales in previous year as prior information to elicit the base prior π_0. The elicited prior π_0 is the $N(1.6614,\ 6.4351\times10^{-3})$ distribution. Under this elicited prior π_0, we obtain the subjective Bayes estimate. But we have some uncertainty in π_0 and the prior information, so we choose $\varepsilon=.1$ in Γ_* and we get the robust Bayes estimate. Also we can obtain easily the empirical Bayes estimate. A number of samples were tried and we have reported our analysis for one sample. Table 3.1 provides the classical estimate δ^C, the subjective estimate δ^0, the robust Bayes estimate δ^{RB} and the empirical Bayes estimate δ^{EB}. Table 3.1 also provides the posterior robustness index for each estimate which is in a sense the sensitivity index of the estimate as the prior varies over Γ.

From Table 3.1, we may find that the subjective Bayes estimate δ^0 is closest to $\gamma(y)$, but not much good in the posterior robustness. The robust Bayes estimate δ^{RB} and the empirical

Bayes estimate δ^{EB} are well behaved in the sense that they are closer to $\gamma(y)$ than at least the classical estimate δ^{C} and good in the posterior robustness index. Finally note that using gross sales in previous year as auxiliary information the ratio estimate was 1.5040 and $|\gamma(y)-\delta| = 0.2243$.

Table 3.1. Estimates and Posterior Robustness Index

	Estimate	$\|\gamma(y) - \delta\|$	POR
δ^{C}	2.1266	0.3983	0.1754
δ^{0}	1.7435	0.0152	0.1467
δ^{RB}	1.8689	0.1407	0.0664
δ^{EB}	1.9929	0.2646	0.0813

References

[1] Berger, J. (1984). The robust Bayesian viewpoint, *Robustness in Bayesian Statistics* (J. Kadane, ed.), North Holland, Amsterdam.

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

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

[4] Ericson, W. A. (1969). Subjective Bayesian models in sampling finite populations (with discussion), *The Journal of the Royal Statistical Society*, Ser. B, 31, 195-23] Godambe, V. P. (1982). Estimation in survey sampling: Robustness and optimality, *Journal of the American Statistical Association*, Vol. 77, 393-406.

[6] Godambe, V. P. and Thompson, M. E. (1971). The specification of prior knowledge by distributions in survey sampling estimation, *Foundations of Statistical Inference* (V. P. Godambe and D. A. Sprott eds.), Holt, Rinehart and Winston, Toronto, 243-258.

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

[8] Royall, R. M. and Pfeffermann, D. (1982). Balanced samples and robust Bayesian inference in finite population sampling, *Biometrika*, Vol. 69, 401-409.

[9] Wasserman, L. (1992). Recent methodological advances in robust Bayesian inference, *Bayesian Statistics* 4 (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith eds.), Oxford Science Publications, Oxford, 483-502.