Journal of the Korean Data & Information Science Society
2014, **25**(3), 685–696
http://dx.doi.org/10.7465/jkdi.2014.25.3.685
한국데이터정보과학회지

Robust Bayesian inference in finite population sampling with auxiliary information under balanced loss function

Eunyoung Kim[1] · Dal Ho Kim[2]

[12]Department of Statistics, Kyungpook National University
Received 23 April 2014, revised 15 May 2014, accepted 20 May 2014

Abstract

In this paper, we develop Bayesian inference of the finite population mean with the assumption of posterior linearity rather than normality of the superpopulation in the presence of auxiliary information under the balanced loss function. We compare the performance of the optimal Bayes estimator under the balanced loss function with ones of the classical ratio estimator and the usual Bayes estimator in terms of the posterior expected losses, risks and Bayes risks.

Keywords: Auxiliary information, balanced loss function, Bayes risk, finite population mean, posterior expected loss, posterior linearity, risk, superpopulation.

1. Introduction

Consider a finite population $\mathcal{U}$ with units labeled $1, 2, \ldots, N$. Let y_i denote the value of a single characteristic attached to the unit i. The vector $\boldsymbol{y} = (y_1, \cdots, y_N)^T$ is the unknown state of nature, and is assumed to belong to $\Theta = R^N$. A subset s of $\{1, 2, \ldots, N\}$ 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 $\boldsymbol{y} \in \Theta$ and $s = \{i_1, \cdots, i_{n(s)}\}$ with $1 \le i_1 < \cdots < i_{n(s)} \le N$, let $y(s) = \{y_{i_1}, \cdots, y_{i_{n(s)}}\}$. One of the main objectives in sample surveys is to draw inference about $\boldsymbol{y}$ or some function (real or vector valued) $\gamma(\boldsymbol{y})$ of $\boldsymbol{y}$ on the basis of s and $y(s)$. For simplicity, only the case where $p(s) > 0$ if and only if $n(s) = n$ will be considered. This amounts to considering only fixed samples of size n. Here we will be concerned exclusively with $\gamma(\boldsymbol{y}) = N^{-1} \sum_{i=1}^N y_i$.

For most sample surveys, for every unit i in the finite population, information is available for one or more auxiliary characteristics, characteristics other than the one of direct interest. We consider the simplest situation when for every unit i in the population, value of a certain auxiliary characteristic, say x_i (> 0) is known $(i = 1, 2, \ldots, N)$. The classical estimator of $\gamma(\boldsymbol{y})$ in such cases is the ratio estimator $\tilde{\gamma}_R = N^{-1}(\sum_{i\in s} y_i / \sum_{i\in s} x_i) \sum_{i=1}^N x_i$ which seems to incorporate the auxiliary information in a very natural manner.

Bayesian approach for finite population sampling was initiated by Hill (1968) and Ericson (1969). Since then, a huge literature has grown in this area. Ericson (1988), Bolfarine and

[1] Ph.D. candidate, Department of Statistics, Kyungpook National University, Daegu 702-701, Korea.
[2] Corresponding author: Professor, Department of Statistics, Kyungpook National University, Daegu 702-701, Korea. E-mail: dalkim@knu.ac.kr

Zacks (1992), Ghosh and Meeden (1997), and Mukhopadhyay (2000) provide an up-to-date account of Bayesian literature in finite population sampling. Goo and Kim (2012) studied Bayesian inference in finite population sampling under measurement error model. In most of the Bayesian literature in survey sampling, the loss is assumed to be squared error.

However, squared error loss is primarily designed to reflect precision of estimation. As an alternative, we consider in this paper balanced loss functions (BLF) first introduced by Zellner (1988, 1992). This loss is a weighted average of two losses, one the squared distance between the parameters and their estimates, and the other the squared distance between the estimates and the data. The latter reflects the goodness of fit of the estimates. Recently Ghosh *et al.* (2008) considered Bayes estimations under random effects normal ANOVA model setup under BLF.

Ghosh and Meeden (1986) considered empirical Bayes estimation of the finite population mean assuming a normal superpopulation model. The normality assumption in the superpopulation model was relaxed by Ghosh and Lahiri (1987). They assumed that the posterior expectation of a finite population mean is a linear function of the sample observations. Such a property is referred to as posterior linearity. This assumption is met when the superpopulation is normal (see Ericson, 1969). There are other situations, however, when the same assumption holds (see, e.g., Diaconis and Ylvisaker 1979; Goldstein 1975; Hartigan 1969).

In this paper, we consider a finite population sampling in the presence of auxiliary information with the assumption of posterior linearity rather than normality of the superpopulation under the balanced loss function. Our procedures are robust in the sense that the normality assumption in the superpopulation model was relaxed by the posterior linearity. In Section 2, we derive the optimal Bayes estimator of the finite population mean under BLF. Also, we compare the performance of the optimal Bayes estimator with ones of the classical ratio estimator and the usual Bayes estimator under the squared error loss based on the posterior expected losses. Moreover, we seek the dominant conditions for typical estimators by the optimal Bayes estimator with respect to the risk function. In Section 3, we evaluate Bayes risks of estimators analytically. Also, we examine the performance of the proposed estimator with ones of the ratio estimator and the usual Bayes estimator in terms of Bayes risk through the Monte Carlo simulation. In Section 4, we summarize results.

2. Bayes estimation under BLF

2.1. Model and assumptions

We consider the situation that there are two values associated with each unit i in the population. One is a variable of interest y_i that is unknown, the other is an auxiliary variable x_i that is a positive known quantity, to which y_i can be expected to tend to be proportional. Estimator can be improved by accounting for auxiliary information such as age, gender, income and so on.

Let r denote the set of nonsample units. Recall that s denotes a sample. So $\mathcal{U} = s \cup r$. Let $\bar{y}_s = \sum_{i \in s} y_i/n$ and $\bar{y}_r = \sum_{j \in r} y_j/(N-n)$ denote the means in the sample units and nonsample units, respectively. Denote $\bar{x}_s$ and $\bar{x}_r$ similarly. Also denote $\boldsymbol{y}_s = (y_i, i \in s)$ and $\boldsymbol{x}_s = (x_i, i \in s)$.

Our main objective is to estimate of the finite population mean $\gamma(\boldsymbol{y}) = N^{-1} \sum_{i=1}^{N} y_i$. We simply express a quantity of interest $\gamma(\boldsymbol{y})$ as γ. Notice that $\gamma = [n\bar{y}_s + (N-n)\bar{y}_r]/N$. It is a combinations of the known mean for the set s of sample units, say "seen", plus the unknown mean for the set r of nonsample units, say "unseen". Estimation in finite population sampling can be thought of as predicting the unseen from the seen. So, we can define a predictor $\tilde{\gamma}$ for γ as follows:

$$\begin{aligned}\tilde{\gamma} &= \frac{1}{N}\left[\, n\bar{y}_s + (N-n)\tilde{\bar{y}}_r \,\right] \\ &= f\bar{y}_s + (1-f)\tilde{\beta}\bar{x}_r\end{aligned}$$

where $\tilde{\bar{y}}_r$ is a predictor of $\bar{y}_r$ and $f = (N-n)/N$ is the finite population correction factor. Estimating γ is then equivalent to predicting the value $\bar{y}_r$ using $\bar{x}_s, \bar{x}_r$ and $\bar{y}_s$.

We assume the superpopulation model without the normality as follows:

(i) $y_i|\beta \overset{iid}{\sim}$ pdf $f(\cdot|\beta)$ with $E[y_i|\beta] = \beta x_i$ and $V[y_i|\beta] = \mu_2(\beta)x_i, i = 1, \ldots, N$;

(ii) β has prior $\pi(\beta\,)$ with $E(\beta) = \mu$ and $V(\beta) = \tau^2$;

(iii) $0 < \sigma^2 = E[\mu_2(\beta)] < \infty$.

Our basic assumption is the posterior linearity, that is, the posterior expectation of β is a linear function of the sample observations without the normality assumption. So

$$E\big(\beta|\boldsymbol{y}_s\big) = \sum_{i\in s} a_i y_i + b, \tag{2.1}$$

where a_i's and b are constants. From the work of Goldstein(1975), (2.1) leads to

$$E\big(\beta|\boldsymbol{y}_s\big) = a\bar{y}_s + b, \tag{2.2}$$

where a and b are constants. Hereafter we shall use (2.2) rather than (2.1). Then it follows that

$$a = \frac{\tau^2}{\tau^2 + \sigma^2/(n\bar{x}_s)} \cdot \frac{1}{\bar{x}_s} = \frac{n\bar{x}_s}{M + n\bar{x}_s} \cdot \frac{1}{\bar{x}_s} = (1-B)\frac{1}{\bar{x}_s} \quad \text{and} \quad b = B\mu,$$

where $M = \sigma^2/\tau^2$ and $B = M/(M + n\bar{x}_s)$.

2.2. Optimal Bayes estimator under the BLF

Before we find the optimal Bayes estimator of the finite population mean under the BLF, we consider two typical estimators of γ. Notice that the classical ratio estimator of γ is given by $\tilde{\gamma}_R = (1-f)\bar{y}_s + f(\bar{y}_s/\bar{x}_s)\bar{x}_r$, since $\tilde{\beta} = \bar{y}_s/\bar{x}_s$. The usual Bayes estimator of γ under the squared error loss in the assumed model is given by

$$\begin{aligned}\tilde{\gamma}_B &= E[\gamma|\bar{y}_s] \\ &= (1-f)\bar{y}_s + f\Big[(1-B)\frac{\bar{y}_s}{\bar{x}_s} + B\beta\Big]\bar{x}_r,\end{aligned}$$

since $\tilde{\beta}$ is the posterior mean of β in this case.

Now, we consider the balanced loss function in our setup as follows:

$$L_B(\tilde{\gamma}\,, \gamma\,) = w(\boldsymbol{y}_s - \tilde{\beta}\boldsymbol{x}_s)'(\boldsymbol{y}_s - \tilde{\beta}\boldsymbol{x}_s)/n + (1-w)(\tilde{\gamma} - \gamma)^2. \tag{2.3}$$

Since $\tilde{\gamma} = N^{-1}\big[n\bar{y}_s + (N-n)\tilde{\beta}\bar{x}_r\big]$, the balanced loss function given in (2.3) can be simplified by

$$L_B(\tilde{\gamma}\,,\gamma\,) = w\sum_{i\in s}(y_i - \tilde{\beta}x_i)^2/n + \ (1-w)f^2(\bar{y}_r - \tilde{\beta}\bar{x}_r)^2. \tag{2.4}$$

In practice, the choice of w reflects the relative weight which the experimenter wants to assign to goodness of fit and precision of estimation.

To find the optimal Bayes estimator γ under the BLF, we consider the posterior expected loss under $L_B(\tilde{\gamma}\,,\gamma\,)$ as follows:

$$\begin{aligned}
\rho(\tilde{\gamma},\gamma) &= E_{\beta|\boldsymbol{y}_s}\big[L_B(\tilde{\gamma}\,,\gamma\,)\big] \\
&= \frac{w}{n}\Big[\sum_{i\in s} y_i^2 - \frac{\big(\sum_{i\in s} x_i y_i\big)^2}{\sum_{i\in s} x_i^2}\Big] + (1-w)f^2 v \\
&\quad + \frac{w}{n}\Big(\tilde{\beta} - \frac{\sum_{i\in s} x_i y_i}{\sum_{i\in s} x_i^2}\Big)^2 + (1-w)f^2\bar{x}_r^2(\,\tilde{\beta} - \bar{\beta})^2.
\end{aligned} \tag{2.5}$$

where v is the posterior variance of $\bar{y}_r$ and $\bar{\beta}$ is the posterior mean of β. On completing the square on $\tilde{\beta}$ from (2.5), we get

$$\begin{aligned}
\rho(\tilde{\gamma},\gamma) &= \frac{w}{n}\Big[\sum_{i\in s} y_i^2 - \frac{\big(\sum_{i\in s} x_i y_i\big)^2}{\sum_{i\in s} x_i^2}\Big] + (1-w)f^2 v \\
&\quad + w^*(1-w)f^2\bar{x}_r^2\big(\hat{\beta} - \bar{\beta}\big)^2 + \ w_1\big(\tilde{\beta} - \tilde{\beta}^*\big)^2,
\end{aligned} \tag{2.6}$$

where $\hat{\beta} = \sum_{i\in s} x_i y_i / \sum_{i\in s} x_i^2$, $w_1 = w\sum_{i\in s} x_i^2/n + (1-w)f^2\bar{x}_r^2$ and $w^* = (w/w_1)\sum_{i\in s} x_i^2/n$. Here the estimator $\tilde{\beta}^*$ that minimizes the posterior expected value of the loss function is the optimal Bayes estimator of β under the BLF, which is given by

$$\tilde{\beta}^* = w^*\hat{\beta} + (1-w^*)\bar{\beta}.$$

Hence the optimal Bayes estimator of the finite population mean γ under the BLF is given by

$$\tilde{\gamma}_{BLF} = (1-f)\bar{y}_s + f\big[w^*\hat{\beta} + (1-w^*)\bar{\beta}\big]\bar{x}_r.$$

2.3. Comparisons of posterior expected losses relative to the BLF

Now we compare the optimal Bayes estimator $\tilde{\gamma}_{BLF}$ under the BLF with typical estimators, $\tilde{\gamma}_R$ and $\tilde{\gamma}_B$ in terms of the posterior expected loss.

First, the posterior expected loss of each estimator of γ is obtained by substituting each estimator for $\tilde{\beta}$ in (2.6). Hence, differences between posterior expected losses are given by

$$\begin{aligned}
\Delta_\rho(\tilde{\gamma}_R, \tilde{\gamma}_{BLF}) &= \rho(\tilde{\gamma}_R,\gamma) - \rho(\tilde{\gamma}_{BLF},\gamma) \\
&= \omega_1\Big(\frac{\bar{y}_s}{\bar{x}_s} - \tilde{\beta}^*\Big)^2 \\
&= \omega_1\Big[(1-\omega^*)(\bar{\beta} - \hat{\beta}) - \Big(\frac{\bar{y}_s}{\bar{x}_s} - \hat{\beta}\Big)\Big]^2 \geq 0
\end{aligned} \tag{2.7}$$

and

$$\begin{aligned}\Delta_\rho(\tilde{\gamma}_B, \tilde{\gamma}_{BLF}) &= \rho(\tilde{\gamma}_B, \gamma) - \rho(\tilde{\gamma}_{BLF}, \gamma) \\ &= \omega_1 \big(\bar{\beta} - \tilde{\beta}^*\big)^2 \\ &= \omega_1 \big[\omega^*(\bar{\beta} - \hat{\beta})\big]^2 \geq 0. \end{aligned} \tag{2.8}$$

Thus, relative losses of $\tilde{\gamma}_R$ and $\tilde{\gamma}_B$ over $\tilde{\gamma}_{BLF}$, denoted by $RL\big(\tilde{\gamma}_R, \tilde{\gamma}_{BLF}\big)$ and $RL\big(\tilde{\gamma}_B, \tilde{\gamma}_{BLF}\big)$, are given by

$$\begin{aligned} RL\big(\tilde{\gamma}_R, \tilde{\gamma}_{BLF}\big) &= \Delta_\rho(\tilde{\gamma}_R, \tilde{\gamma}_{BLF})/\rho(\tilde{\gamma}_{BLF}, \gamma) \\ &= \frac{\omega_1 \big[(1-\omega^*)(\bar{\beta} - \hat{\beta}) - (\bar{y}_s/\bar{x}_s - \hat{\beta})\big]^2}{\upsilon_a^2 + \omega^*(1-\omega) f^2 \bar{x}_r^2 (\bar{\beta} - \hat{\beta})^2}, \end{aligned} \tag{2.9}$$

and

$$\begin{aligned} RL\big(\tilde{\gamma}_B, \tilde{\gamma}_{BLF}\big) &= \Delta_\rho(\tilde{\gamma}_B, \tilde{\gamma}_{BLF})/\rho(\tilde{\gamma}_{BLF}, \gamma) \\ &= \frac{\omega_1 \big[\omega^*(\bar{\beta} - \hat{\beta})\big]^2}{\upsilon_a^2 + \omega^*(1-\omega) f^2 \bar{x}_r^2 (\bar{\beta} - \hat{\beta})^2}, \end{aligned} \tag{2.10}$$

where

$$\upsilon_a^2 = \frac{\omega}{n}\left[\sum_{i\in s} y_i^2 - \frac{\big(\sum_{i\in s} x_i y_i\big)^2}{\sum_{i\in s} x_i^2}\right] + (1-\omega) f^2 \upsilon.$$

It is easy to show that the differences between posterior expected losses and relative losses are positive except for ω is equal to zero from (2.7)-(2.10). This imply that the optimal Bayes estimator $\tilde{\gamma}_{BLF}$ is superior to typical estimators in terms of the posterior expected loss.

Table 2.1 Two population datasets used in the numerical study

	Symbol	
	Counties 70	Cancer
Description	Counties in NC, SC and GA with fewer than 100,000 households in 1960	Counties in NC, SC, and GA with 1960 white female population<100,000
Source	US Census	x : US Census y : Mason and Mckay (1974)
x	Number of households, 1960	Adult white female population, 1960
y	Population, excluding residents of group quarters, 1970	Breast cancer mortality, 1950-69 (white females)
N	304	301

We also compare the posterior expected loss through the numerical study using real data to show the superiority of $\tilde{\gamma}_{BLF}$. We consider the two datasets among the six real populations that used in Royall and Cumberland (1981) for an empirical study of the ratio estimator and estimates of its variance, described Table 2.1.

In the population named Counties 70, the variable of interest y is the 1970 population, in thousands, in each of $N = 304$ counties in NC, SC and GA, and the auxiliary variable

x is the number of households, in thousands, in each county in 1960. In the other case, the Cancer data consists of the breast cancer mortality in 1950-1969 for white females and the adult white female population in 1960, in thousands, of 301 counties in NC, SC and GA for the auxiliary information.

Table 2.2 Relative losses for various of ω, μ and f in the Counties 70 population

		$f = 0.9$					$f = 0.95$				
	μ	$\omega = .00$	.25	.50	.75	1.00	$\omega = .00$	.25	.50	.75	1.00
	0.5	6.017	0.788	0.458	0.026	0.220	11.873	2.896	0.476	0.002	0.533
	1.0	5.717	0.308	0.082	0.132	0.041	8.870	1.019	0.393	0.119	0.187
RL_R	2.0	2.654	0.368	0.129	0.156	1.133	3.184	1.606	0.225	0.035	0.964
	3.5	0.119	0.021	0.001	0.001	0.053	0.301	0.039	0.072	0.017	0.089
	5.0	0.692	0.172	0.002	0.162	0.150	1.107	0.147	0.009	0.178	0.001
	0.5	0	0.805	1.253	1.869	16.173	0	0.151	1.292	6.400	70.838
	1.0	0	0.949	2.191	3.710	8.611	0	0.605	1.773	4.373	23.779
RL_B	2.0	0	0.530	0.638	1.165	5.971	0	0.174	0.370	3.694	17.883
	3.5	0	0.286	0.291	0.064	0.883	0	0.155	0.183	0.714	0.523
	5.0	0	0.052	0.444	0.662	1.591	0	0.667	0.022	0.573	2.167

Table 2.3 Relative losses for various of ω, μ and f in the Cancer population

		$f = 0.9$					$f = 0.95$				
	μ	$\omega = .00$	.25	.50	.75	1.00	$\omega = .00$	.25	.50	.75	1.00
	0.5	3.919	1.247	0.003	0.144	0.032	6.241	0.957	1.018	0.003	0.099
	1.0	2.239	0.271	0.150	0.022	0.171	3.693	0.435	0.337	0.548	0.108
RL_R	2.0	0.505	0.197	0.003	0.068	0.002	1.058	0.294	0.201	0.087	0.138
	3.5	0.003	0.002	0.030	0.034	0.295	0.042	0.001	0.094	0.012	0.001
	5.0	1.069	0.072	0.073	0.067	0.017	1.662	0.520	0.009	0.067	0.644
	0.5	0	0.127	1.001	1.862	5.404	0	0.146	0.287	2.372	10.599
	1.0	0	0.162	0.411	0.784	3.213	0	0.237	0.356	1.158	4.416
RL_B	2.0	0	0.025	0.922	0.968	0.875	0	0.283	0.143	0.654	0.904
	3.5	0	0.018	0.023	0.039	0.106	0	0.014	0.265	0.266	0.101
	5.0	0	0.146	0.202	0.197	1.199	0	0.108	1.065	0.631	1.368

For each of complete population, we find the population variance $\sigma^2 = (N-1)^{-1}\sum_{i=1}^{N}(y_i - \beta x_i)^2$. To elicit the base prior π for β in the Counties 70, we use the Counties 60 population, which is also one of the six real populations. In this Counties 60, x is the same as one in the Counties 70, but y is the 1960 population of 304 counties in NC, SC and GA. The elicited τ^2 is 0.5995. For the Cancer population, we put τ^2 to be one because of no prior information for the elicitation. We select 10% and 5% simple random samples without replacement from each of the two populations. The sample sizes are $n = 30$ and $n = 15$, respectively. We calculate relative losses of typical estimators, $\tilde{\gamma}_R$ and $\tilde{\gamma}_B$, with respect to the optimal Bayes estimator, $\tilde{\gamma}_{BLF}$, for each of samples, using the given data. Various choices of ω, μ and f are considered.

In the Counties 70 case, Table 2.2 provides relative losses of $\tilde{\gamma}_R$ and $\tilde{\gamma}_B$ over $\tilde{\gamma}_{BLF}$, denoted by RL_R and RL_B respectively. Table 2.3 shows the results for the Cancer population. An inspection of Table 2.2 and Table 2.3 reveal that the optimal Bayes estimator $\tilde{\gamma}_{BLF}$ is superior to the Bayes estimator $\tilde{\gamma}_B$ as well as the classical ratio estimator $\tilde{\gamma}_R$ in terms of the posterior expected loss. For example, in the case of $\omega = 0.5, \mu = 2.0$ and $f = 0.9$, RL_B equals 0.638. That means that posterior expected loss under these conditions is inflated by

63.8% using $\tilde{\gamma}_B$ rather than $\tilde{\gamma}_{BLF}$. Also, if $\omega = 0$, reflecting only the precision of estimation, then $\tilde{\gamma}_{BLF}$ is equivalent to $\tilde{\gamma}_B$. So RL_B is equal to zero in this case.

2.4. Conditions for dominance related to risks

We compute risk functions of $\tilde{\gamma}_R$, $\tilde{\gamma}_B$ and $\tilde{\gamma}_{BLF}$ under the BLF. Then we will find out such conditions that the optimal Bayes estimator $\tilde{\gamma}_{BLF}$ dominates typical estimators, $\tilde{\gamma}_R$ and $\tilde{\gamma}_B$ related to risks. The risk function for any estimator $\tilde{\gamma}$ of γ associated with the BLF is obtained by integrating the BLF given in (2.4) over the samples given the parameter β.

First, the risk function of the classical ratio estimator $\tilde{\gamma}_R$ with respect to the BLF is given by

$$\begin{aligned} R(\tilde{\gamma}_R, \gamma) &= E_{\boldsymbol{y}_s|\beta}\left[L_B(\tilde{\gamma}_R, \gamma)\right] \\ &= \frac{\mu_2(\beta)}{n}\left[wn\bar{x}_s - w\frac{\sum_{i\in s} x_i^2}{n}\frac{1}{\bar{x}_s} + (1-w)f^2\bar{x}_r\frac{n}{N-n} + (1-w)f^2\bar{x}_r^2\frac{1}{\bar{x}_s}\right]. \end{aligned}$$

Also, the risk function of $\tilde{\gamma}_B$ with respect to the BLF is given by

$$\begin{aligned} R(\tilde{\gamma}_B, \gamma) &= E_{\boldsymbol{y}_s|\beta}\left[L_B(\tilde{\gamma}_B, \gamma)\right] \\ &= \frac{\mu_2(\beta)}{n}\left[\omega n\bar{x}_s + \omega_1(1-B)^2\frac{1}{\bar{x}_s} + (1-\omega)f^2\bar{x}_r\frac{n}{N-n} - 2\omega\frac{\sum_{i\in s} x_i^2}{n}(1-B)\frac{1}{\bar{x}_s}\right] \\ &\quad + \omega_1 B^2(\beta-\mu)^2. \end{aligned} \tag{2.11}$$

Next, to obtain the risk function of $\tilde{\gamma}_{BLF}$, the optimal Bayes estimator of β under the BLF, denoted by $\tilde{\beta}^*$ can be written as

$$\tilde{\beta}^* = \omega^*\left(\frac{\sum_{i\in s} x_i y_i}{\sum_{i\in s} x_i^2} - \frac{\bar{y}_s}{\bar{x}_s}\right) + (1-C)\frac{\bar{y}_s}{\bar{x}_s} + C\mu,$$

where $C = (1-\omega^*)B$. Let

$$A_1 = \omega^*\left(\frac{\sum_{i\in s} x_i y_i}{\sum_{i\in s} x_i^2} - \frac{\bar{y}_s}{\bar{x}_s}\right) \text{ and } A_2 = (1-C)\frac{\bar{y}_s}{\bar{x}_s} + C\mu.$$

Then, the risk function of the optimal Bayes estimator $\tilde{\gamma}_{BLF}$ with respect to the BLF is given by

$$\begin{aligned} R(\tilde{\gamma}_{BLF}, \gamma) &= E_{\mathbf{y}_s|\beta}\left[L_B(\tilde{\gamma}_{BLF}, \gamma)\right] \\ &= R_C + \frac{\mu_2(\beta)}{n}\left[\omega\omega^*\left(\frac{\sum_{i\in s} x_i^2}{n\bar{x}_s} - \frac{\sum_{i\in s} x_i^3}{\sum_{i\in s} x_i^2}\right)\right], \end{aligned}$$

where R_C is derived from $R(\tilde{\gamma}_B, \gamma)$ by substituting C for B in (2.11). Hence, it follows that

$$\begin{aligned} R(\tilde{\gamma}_{BLF}, \gamma) = \; & \frac{\mu_2(\beta)}{n}\left[\omega n\bar{x}_s + \omega_1(1-C)^2\frac{1}{\bar{x}_s} + (1-\omega)f^2\bar{x}_r\frac{n}{N-n}\right. \\ & \left. -2\omega\frac{\sum_{i\in s} x_i^2}{n}(1-C)\frac{1}{\bar{x}_s} + \omega\omega^*\left(\frac{\sum_{i\in s} x_i^2}{n\bar{x}_s} - \frac{\sum_{i\in s} x_i^3}{\sum_{i\in s} x_i^2}\right)\right] + \omega_1^2 C^2(\beta-\mu)^2. \end{aligned}$$

Now, we seek dominant conditions for typical estimators by the proposed optimal Bayes estimator in terms of risks. Let $\delta^2 = (\beta-\mu)^2/\big(\mu_2(\beta)/n\big)$. Then the differences between risk functions is obtained by

$$\begin{aligned}\Delta_R\big(\tilde{\gamma}_R, \tilde{\gamma}_{BLF}\big) =&\ R\left(\tilde{\gamma}_R, \gamma\right) - R\left(\tilde{\gamma}_{BLF}, \gamma\right) \\ =&\ \frac{\mu_2(\beta)}{n}\Bigg[2\,(1-\omega)\, f^2\, \bar{x}_r^2\, C\, \frac{1}{\bar{x}_s} - \omega_1\, C^2 \left(\delta^2 + \frac{1}{\bar{x}_s} \right) \\ &- \omega\, \omega^* \,\frac{\sum_{i\in s} x_i^2}{n} \left(\frac{1}{\bar{x}_s} - \frac{n \sum_{i\in s} x_i^3}{\left(\sum_{i\in s} x_i^2\right)^2} \right) \Bigg]\end{aligned}$$

and

$$\begin{aligned}\Delta_R\big(\tilde{\gamma}_B, \tilde{\gamma}_{BLF}\big) =&\ R\left(\tilde{\gamma}_B, \gamma\right) - R\left(\tilde{\gamma}_{BLF}, \gamma\right) \\ =&\ \frac{\mu_2(\beta)}{n}\Bigg[\omega_1\, (\,B^2 - C^2\,) \left(\delta^2 + \frac{1}{\bar{x}_s} \right) - 2\,(1-\omega)\, f^2\, \bar{x}_r^2\, \frac{1}{\bar{x}_s}\,(B-C) \\ &- \omega\, \omega^* \,\frac{\sum_{i\in s} x_i^2}{n} \left(\frac{1}{\bar{x}_s} - \frac{n \sum_{i\in s} x_i^3}{\left(\sum_{i\in s} x_i^2\right)^2} \right) \Bigg].\end{aligned}$$

Thus, we can see that condition for dominance of $\tilde{\gamma}_R$ by $\tilde{\gamma}_{BLF}$ is given by

$$\delta^2 \ <\ \frac{2}{B\,\bar{x}_s} - \left(\frac{w^*}{1-w^*}\right)^2 \frac{1}{B^2}\left(\frac{1}{\bar{x}_s} - \frac{n\sum_{i\in s} x_i^3}{\left(\sum_{i\in s} x_i^2\right)^2}\right) - \frac{1}{\bar{x}_s}.$$

Also, we get the dominant condition for $\tilde{\gamma}_B$ by $\tilde{\gamma}_{BLF}$ as follows:

$$\delta^2 \ >\ \frac{2(1-w^*)}{2-w^*}\,\frac{1}{B\,\bar{x}_s} + \frac{w^*}{2-w^*}\,\frac{1}{B^2}\left(\frac{1}{\bar{x}_s} - \frac{n\sum_{i\in s} x_i^3}{(\sum_{i\in s} x_i^2)^2}\right) - \frac{1}{\bar{x}_s}.$$

As can be seen, the smaller δ^2, the more easily $\tilde{\gamma}_{BLF}$ dominates $\tilde{\gamma}_R$. Whereas, the larger δ^2, the more easily $\tilde{\gamma}_{BLF}$ dominates $\tilde{\gamma}_B$.

3. Monte Carlo simulation for Bayes risks

3.1. Bayes risks

In this subsection, we theoretically evaluate Bayes risks related to BLF of considered estimators to examine the superior performance of the proposed optimal Bayes estimator in terms of Bayes risk.

The Bayes risks of typical estimators, $\tilde{\gamma}_R$ and $\tilde{\gamma}_B$ with respect to the BLF are given by

$$r\left(\tilde{\gamma}_R, \gamma\right) = \frac{\sigma^2}{n}\left[\ \omega\, n\, \bar{x}_s - \omega\, \frac{\sum_{i\in s} x_i^2}{n}\,\frac{1}{\bar{x}_s} + (1-\omega)\, f^2\, \bar{x}_r\, \frac{n}{N-n} + (1-\omega)\ f^2\, \bar{x}_r^2\, \frac{1}{\bar{x}_s}\ \right]$$

and

$$\begin{aligned} r(\tilde{\gamma}_B, \gamma) &= \frac{E[\mu_2(\beta)]}{n}\Big[\omega\, n\, \bar{x}_s + \omega_1 (1-B)^2 \frac{1}{\bar{x}_s} + (1-\omega) f^2 \bar{x}_r \frac{n}{N-n} \\ &\qquad - 2\omega \frac{\sum_{i\in s} x_i^2}{n}(1-B)\frac{1}{\bar{x}_s}\Big] + \omega_1 B^2 E[(\beta-\mu)^2] \\ &= \frac{\sigma^2}{n}\Big[\omega\, n\, \bar{x}_s + (1-\omega) f^2 \bar{x}_r \frac{n}{N-n} - 2\omega \frac{\sum_{i\in s} x_i^2}{n}(1-B)\frac{1}{\bar{x}_s} + \omega_1 (1-B)\frac{1}{\bar{x}_s}\Big]. \end{aligned}$$

And the Bayes risk of $\tilde{\gamma}_{BLF}$ with respect to the BLF also is given by

$$\begin{aligned} r(\tilde{\gamma}_{BLF}, \gamma) = \frac{\sigma^2}{n}\Big[&\omega\, n\, \bar{x}_s + (1-\omega) f^2 \bar{x}_r \frac{n}{N-n} - \omega \frac{\sum_{i\in s} x_i^2}{n}(2-C)\frac{1}{\bar{x}_s} + \omega_1 (1-C)\frac{1}{\bar{x}_s} \\ &+ \omega\,\omega^* \Big(\frac{\sum_{i\in s} x_i^2}{n\,\bar{x}_s} - \frac{\sum_{i\in s} x_i^3}{\sum_{i\in s} x_i^2}\Big)\Big]. \end{aligned}$$

Then, differences between Bayes risks are given by

$$\Delta_r\big(\tilde{\gamma}_R, \tilde{\gamma}_{BLF}\big) = \frac{\sigma^2}{n}\Big[(1-\omega)(1-\omega^*) B f^2 \bar{x}_r^2 \frac{1}{\bar{x}_s} + \omega\omega^* \Big(\frac{\sum_{i\in s} x_i^3}{\sum_{i\in s} x_i^2} - \frac{\sum_{i\in s} x_i^2}{n\,\bar{x}_s}\Big)\Big] \tag{3.1}$$

and

$$\Delta_r\big(\tilde{\gamma}_B, \tilde{\gamma}_{BLF}\big) = \frac{\sigma^2}{n}\Big[\omega\omega^* B \frac{\sum_{i\in s} x_i^2}{n}\frac{1}{\bar{x}_s} + \omega\omega^* \Big(\frac{\sum_{i\in s} x_i^3}{\sum_{i\in s} x_i^2} - \frac{\sum_{i\in s} x_i^2}{n\,\bar{x}_s}\Big)\Big]. \tag{3.2}$$

It can be easily checked that $\Delta_r\big(\tilde{\gamma}_R, \tilde{\gamma}_{BLF}\big)$ and $\Delta_r\big(\tilde{\gamma}_B, \tilde{\gamma}_{BLF}\big)$ are strictly positive except for w equal to zero from (3.1)-(3.2). That means the optimal Bayes estimator, $\tilde{\gamma}_{BLF}$ is superior to typical estimators, $\tilde{\gamma}_R$ and $\tilde{\gamma}_B$ in terms of the Bayes risk. Furthermore, $\Delta_r\big(\tilde{\gamma}_B, \tilde{\gamma}_{BLF}\big)$ is increasing monotonically in ω. Whereas, $\Delta_r\big(\tilde{\gamma}_R, \tilde{\gamma}_{BLF}\big)$ has not monotonicity in ω.

3.2. Monte Carlo simulation

In this subsection, we compare the Bayes risks of $\tilde{\gamma}_{BLF}$ with those of $\tilde{\gamma}_R$ and $\tilde{\gamma}_B$ through the Monte Carlo simulations. The simulated Bayes risks are the average losses given in (2.4) after 10,000 repetitions of an experiment. We considered normal and Poisson cases with the assumption of posterior linearity in the assumed model. The percentage risk improvement of $\tilde{\gamma}_{BLF}$ over $\tilde{\gamma}_R$ and $\tilde{\gamma}_B$, denoted by $PCTIMP_R$ and $PCTIMP_B$ repectively, are provided based on the simulated Bayes risks of each estimator.

We consider the normal case first. The β and y_i's are generated using the RNORM function of the R. In this case, the auxiliary information (x_i) is the number of households in each county in 1960, referred to the previous real data. In the Counties 70 population, we obtained β is equal to 4.14. Therefore, let β be four. The Bayes risks are calculated for various ω, f, τ^2 and σ^2. Table 3.1 and Table 3.2 show the results. As previously mentioned, if $\omega = 0$, then $\tilde{\gamma}_{BLF}$ is equivalent to $\tilde{\gamma}_B$ and $PCTIMP_B$ is equal to zero. From the table, yon can check that $PCTIMP_R$ and $PCTIMP_B$ are positive except for ω equal to zero. This means that the proposed estimator $\tilde{\gamma}_{BLF}$ is better than both $\tilde{\gamma}_C$ and $\tilde{\gamma}_B$ in terms of Bayes risks.

Table 3.1 PCTIMPs of $\tilde{\gamma}_{BLF}$ over $\tilde{\gamma}_C$ and $\tilde{\gamma}_B$ in normal case ($\sigma^2 = 1$)

		$f = 0.8$		$f = 0.9$		$f = 0.95$	
τ^2	ω	$PCTIMP_R$	$PCTIMP_B$	$PCTIMP_R$	$PCTIMP_B$	$PCTIMP_R$	$PCTIMP_B$
	.00	12.712	0.000	26.112	0.000	45.403	0.000
	.25	2.064	2.223	4.156	3.860	9.662	7.112
0.1	.50	2.595	3.098	4.522	5.887	7.456	10.679
	.75	3.030	3.618	5.086	6.970	7.795	13.292
	1.00	3.231	3.886	5.639	7.828	8.674	15.112
	.00	1.525	0.000	3.389	0.000	7.801	0.000
	.25	1.842	1.866	3.063	3.019	4.593	4.080
1.0	.50	2.616	2.685	4.368	4.556	6.178	6.601
	.75	2.911	2.990	5.005	5.254	7.757	8.677
	1.00	3.255	3.347	5.554	5.904	8.715	9.861
	.00	0.749	0.000	1.706	0.000	4.115	0.000
	.25	1.868	1.875	2.960	2.952	4.137	3.747
2.0	.50	2.534	2.571	4.163	4.251	6.121	6.460
	.75	2.976	3.011	5.190	5.340	7.373	7.868
	1.00	3.228	3.265	5.570	5.720	8.467	9.002

Table 3.2 PCTIMPs of $\tilde{\gamma}_{BLF}$ over $\tilde{\gamma}_C$ and $\tilde{\gamma}_B$ in normal case ($f = 0.9$)

		$\sigma^2 = 0.5$		$\sigma^2 = 1.0$		$\sigma^2 = 2.0$	
τ^2	ω	$PCTIMP_R$	$PCTIMP_B$	$PCTIMP_R$	$PCTIMP_B$	$PCTIMP_R$	$PCTIMP_B$
	.00	15.280	0.000	26.112	0.000	40.572	0.000
	.25	3.350	3.161	4.156	3.860	4.809	4.674
0.1	.50	4.456	5.311	4.522	5.887	4.678	6.798
	.75	5.135	6.136	5.086	6.970	5.106	8.256
	1.00	5.644	7.139	5.639	7.828	5.569	9.020
	.00	2.217	0.000	3.389	0.000	7.357	0.000
	.25	3.017	2.982	3.063	3.020	3.234	3.091
1.0	.50	4.213	4.300	4.368	4.556	4.438	4.823
	.75	5.026	5.152	5.005	5.254	5.110	5.644
	1.00	5.767	5.915	5.554	5.904	5.486	6.092
	.00	0.873	0.000	1.706	0.000	3.152	0.000
	.25	2.753	2.706	2.960	2.952	3.224	3.137
2.0	.50	4.112	4.150	4.163	4.251	4.311	4.452
	.75	5.072	5.146	5.190	5.340	4.902	5.140
	1.00	5.604	5.694	5.570	5.720	5.571	5.866

Furthermore, as values of f and σ^2 increase, $PCTIMP_R$ and $PCTIMP_B$ are increasing. Large values of f means small n, and large values of σ^2 means y_i's are very spread out from the mean. The smaller n and the bigger variance of y_i's, the better $\tilde{\gamma}_{BLF}$ than typical estimates. On the contrary, $PCTIMP_C$ and $PCTIMP_B$ decrease with τ^2, because higher values of τ^2 would explain greater uncertainty about the prior. In addition, PCTIMP$_B$ is monotonically increasing in w. However, PCTIMP$_R$ has not monotonicity in w.

Next, the Poisson case is considered. For Monte Carlo simulations, we first generate β from a $gamma(\alpha, p)$, and then we generate y_i's from Poisson($\beta\, x_i$). Here, the auxiliary information (x_i) is the adult white female population in 1960 of 301 counties. Variable type of gamma distributions are considered. The cases $p = 1, p = 0.5$ and $p = 3$ correspond to the exponetial, negatively skewed and positively skewed gammas, respectively.

Table 3.3 provides simulation findings. The results also report that the optimal Bayes estimator $\tilde{\gamma}_{BLF}$ is better than both $\tilde{\gamma}_R$ and $\tilde{\gamma}_B$ in terms of Bayes risks.

Table 3.3 PCTIMPs of $\tilde{\gamma}_{BLF}$ over $\tilde{\gamma}_C$ and $\tilde{\gamma}_B$ in Poisson case

		$f = 0.8$		$f = 0.9$		$f = 0.95$	
$(\alpha,\ p)$	ω	$PCTIMP_R$	$PCTIMP_B$	$PCTIMP_R$	$PCTIMP_B$	$PCTIMP_R$	$PCTIMP_B$
	.00	0.019	0.000	2.169	0.000	3.503	0.000
	.25	1.540	1.501	2.316	2.350	3.510	3.170
(0.5, 0.5)	.50	1.997	1.997	3.295	3.318	5.361	5.469
	.75	2.289	2.302	4.016	4.130	6.511	6.894
	1.00	2.551	2.592	4.734	4.888	7.198	7.580
	.00	1.384	0.000	3.916	0.000	6.741	0.000
	.25	1.376	1.378	2.288	2.243	4.072	3.120
(1.0, 0.5)	.50	1.958	1.974	3.408	3.545	5.154	5.527
	.75	2.305	2.357	3.981	4.172	6.185	6.966
	1.00	2.582	2.652	4.715	5.058	7.564	8.445
	.00	4.573	0.000	11.137	0.000	23.498	0.000
	.25	1.487	1.521	2.592	2.584	5.858	4.094
(4.0, 0.5)	.50	1.957	2.174	3.341	3.864	5.841	7.086
	.75	2.312	2.558	4.122	4.967	6.543	9.006
	1.00	2.409	2.647	4.447	5.258	7.574	10.577
	.00	0.611	0.000	1.639	0.000	4.307	0.000
	.25	1.450	1.419	2.394	2.370	3.583	3.151
(0.5, 1.0)	.50	1.936	1.951	3.376	3.438	5.096	5.188
	.75	2.348	2.395	4.142	4.228	6.403	6.802
	1.00	2.563	2.590	4.573	4.720	7.311	7.812
	.00	1.945	0.000	1.587	0.000	6.075	0.000
	.25	1.414	1.437	2.334	2.256	3.723	3.222
(1.0, 1.0)	.50	1.928	1.979	3.429	3.582	5.017	5.386
	.75	2.258	2.322	4.078	4.326	6.714	7.404
	1.00	2.600	2.673	4.699	4.929	7.561	8.424
	.00	5.181	0.000	9.970	0.000	21.296	0.000
	.25	1.458	1.593	2.836	2.529	6.017	4.193
(4.0, 1.0)	.50	1.933	2.105	3.419	3.849	5.786	7.307
	.75	2.308	2.518	4.145	4.744	6.484	8.784
	1.00	2.465	2.688	4.770	5.660	7.776	10.698
	.00	1.017	0.000	1.025	0.000	3.479	0.000
	.25	1.422	1.437	2.214	2.219	3.569	3.317
(0.5, 3.0)	.50	2.011	2.038	3.201	3.289	5.294	5.475
	.75	2.336	2.362	4.065	4.163	6.354	6.686
	1.00	2.525	2.556	4.548	4.691	7.444	7.903
	.00	1.119	0.000	2.480	0.000	6.592	0.000
	.25	1.339	1.377	2.223	2.228	3.678	3.377
(1.0, 3.0)	.50	2.043	2.088	3.288	3.437	5.316	5.694
	.75	2.316	2.393	4.007	4.193	6.455	7.172
	1.00	2.573	2.622	4.574	4.736	7.292	8.051
	.00	4.940	0.000	10.310	0.000	22.262	0.000
	.25	1.487	1.518	2.781	2.725	5.756	4.611
(4.0, 3.0)	.50	1.998	2.152	3.490	3.907	5.762	6.940
	.75	2.296	2.515	4.073	4.858	6.662	8.908
	1.00	2.530	2.795	4.758	5.639	7.656	10.728

4. Concluding remarks

Throughout this paper, we have considered the balanced loss function in the Bayesian inference for finite population mean. The balanced loss function takes both the goodness of fit and the precision of estimation into account.

We relaxed the normality assumption and assumed the posterior linearity instead. We considered the superpopulation model in the presence of auxiliary information under the assumed model. And, we obtained the optimal Bayes estimator of the finite population mean under the BLF.

We found some conditions for dominance of typical estimators by the optimal Bayes estimator. And It turned out that the optimal Bayes estimator is superior to typical estimators in terms of the posterior expected loss and Bayes risks through the numerical study..

References

Bolfarine, H. and Zacks, S. (1992). *Prediction theory for finite populations*, Springer-Verlag, New York.

Diaconis, P. and Ylvisaker, D. (1979). Conjugate priors for exponential Families. *The Annals of Statistics*, **7**, 269-281.

Ericson, W. A. (1969). Subjective Bayesian models in sampling finite populations (with discussion). *Journal of the Royal Statistical Society B*, **31**, 195-233.

Ericson, W. A. (1988). Bayesian inference in finite populations. In *Handbook of Statistics, Vol. 6 : Sampling*, edited by P.R. Krishnaiah and C.R. Rao, North-Holland, Amsterdam, 213-246.

Ghosh, M., Kim, M. J. and Kim, D. H. (2008). Constrained Bayes and empirical Bayes estimation under random effects normal ANOVA model with balanced loss function. *Journal of Statistical Planning and Inference*, **138**, 2017-2028.

Ghosh, M. and Lahiri, P. (1987). Robust empirical Bayes estimation of means from stratified samples. *Journal of the American Statistical Association*, **82**, 1153-1162.

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 Meeden, G. (1997). *Bayesian methods for finite population sampling*, Chapman and Hall, London.

Goldstein, M. (1975). A note on some Bayesian nonparametric estimates. *The Annals of Statistics*, **3**, 736-740.

Goo, Y. M. and Kim, D. H. (2012). Bayesian inference in finite population sampling under measurement error model. *Journal of the Korean Data & Information Science Society*, **23**, 1241-1247.

Hartigan, J. A. (1969). Linear Bayes methods. *Journal of the Royal Statistical Society B*, **31**, 446-454.

Hill, B. (1968). Posterior distribution of percentiles: Bayes theorem for sampling from a population. *Journal of the American Statistical Association*, **63**, 677-691.

Mukhopadhyay, P. (2000). *Topics in survey sampling*, Springer-Verlag, New York.

Royall, R. M. and Cumberland, W. G. (1981). An empirical study of the ratio estimator and estimators of its variance (with discussion). *Journal of the American Statistical Association*, **76**, 66-88.

Zellner, A. (1988). Bayesian analysis in econometrics. *Journal of Econometrics*, **37**, 27-50.

Zellner, A. (1992). Bayesian and non-Bayesian estimation using balanced loss functions. In *Statistical Decision Theory and Related Topics V*, edited by S. S. Gupta and J. O. Berger, Springer-Verlag, New York, 377-390.