Chapter 5

Bayes Estimator and its Admissibility under PPS Sampling
5.1. Introduction

Consider a finite population $U = \{1, \ldots, N\}$ containing $N$ units. Let $y_i$ denote the value of the character under study for the $i$th unit. It is desired to estimate the populations total $Y = \sum_{i \in U} y_i$ on the basis of a sample. When sampling $n$ units with probability proportional to size (pps), with replacement after each draw (ppswr for brevity), the finite population total may be estimated by the estimator $\hat{Y}_{HH} = \sum_{i=1}^{n} y_i / np_i$, where $p_i$ is the probability of selecting the unit occurring at the $i$th draw. Basu (1958) presents for this design an unbiased estimator $\bar{Y}_E = \sum_{i=1}^{u} C_i y_{(i)} / p_{(i)}$ where $u$ is the number of distinct units in the sample, the suffix $(i)$ indicates the $i$th distinct unit in the sample and $C_i$ involves cumbersome computation as $n$ increases, is uniformly more efficient than $\hat{Y}_{HH}$. This estimator is not identical with $\hat{Y}_{HHd} = \sum_{i=1}^{u} y_{(i)} / np_{(i)}$. In fact, $\hat{Y}_{HHd}$ is in general not unbiased. In view of the simplicity of $\bar{Y}_{HHd}$, Subrahmanya (1966) studied the properties of it.

The pps estimator $\hat{Y}_{HH}$ is inadmissible as it depends on multiplicity. Using Rao-Blackwellization, Pathak (1962) gives an improved (rather complicated) estimator which does not admit a simple non-negative variance estimator as does the pps estimator. Moreover the gain in efficiency is considered to be small, unless the sampling fraction is large. Thus, the resulting estimator is less useful in practice than the original pps estimator.

Godambe (1960) shows that the linear estimator $\hat{Y}_L = \sum_{i=1}^{u} \beta_{si} y_{(i)}$ of $Y$ is admissible for $\beta_{si} = 1/p'_{(i)}$, where $p'_{(i)}$ denotes the probability that the $i$th unit is
included in the sample. In this chapter we suggest a generalized difference (GD) estimator based on distinct units of the pps sample. This estimator includes Godambe estimator. In the next section, we obtain Bayes estimator and establish its admissibility using the Bayes risk method given in Lehmann and Casella (1998). For the sake of comparison of the Bayes estimator, in Section 3, we discuss generalize regression (GREG) estimator (see, Särndal et al., 1992) and an optimal (OPT) estimator (see, Montanari, 1998) for distinct units. In Section 4 a small scale Monte Carlo simulation is carried out for the comparison of estimators. Final conclusions are given Section 5.

Denote the random variable \( z \) taking the values

\[
z_i = (y_i - e_i)/p_i \quad \text{with probabilities} \quad p_i, \quad i = 1, \ldots, N,
\]

\[
\hat{e}_{HH} = \frac{1}{n}\sum_{i \in S} e_i / p_i \quad \text{and} \quad e = \sum_{i \in U} e_i.
\]

We suggest the following estimator:

\[
\hat{\gamma}_{GD} = \hat{\gamma}_{HH} + [e - \hat{e}_{HH}]
\]

where \( e = (e_1, \ldots, e_N) \) is a vector of known real numbers.

In particular, if \( e_i = \beta x_i \), for \( i = 1, \ldots, N \), where \( \beta \) is a predetermined constant and \( x_i \) is the known value for unit \( i \) of an auxiliary variable, then \( \hat{\gamma}_{GD} \) takes the form

\[
\hat{\gamma}_{GD} = \hat{\gamma}_{HH} + \beta(X - \bar{X}_{HH})
\]  \hspace{1cm} (5.1.1)

Now onwards, we assume that \( s \) is a fixed effective sample of size \( n \), i.e., \( s \) is a collection of \( n \) distinct units. Let \( S = \{s: \text{fixed effective size of} \ s = n\} \). It has been shown by Basu (1958) that the order statistic \( T = \{y_{(1)}, y_{(2)}, \ldots, y_{(n)}\} \) is
sufficient, where \( y(1), y(2), \ldots, y(n) \) are distinct values of \( y \) variable in sample, arranged in ascending order of their unit-indices. To simplify the notation, we use \( y_1, \ldots, y_n \) in place of \( y(1), y(2), \ldots, y(n) \). For sake of comparison of the suggested estimator we first discuss the following two estimators.

5.2. Random Regression Coefficient

In marketing, industrial organization and transportation economics, hundreds of papers use random coefficient models (see, Train, 2003). In this section, assuming random regression coefficient, we construct a Bayes estimator to estimate the population total with squared error loss function.

5.2.1 The Bayed Estimator

Let \( t(d) \) denote an estimator of \( Y \), where \( t \) depends on \( y \) only through \( d = \{(i, y_i): i \in s\} \). Consider the following normal superpopulation model

\[ \xi: y_1, \ldots, y_N \text{ are independently distributed with joint prior density, given } \theta = (\beta, h), \]

\[ f(y|\theta) \propto \frac{h^{N/2}}{\prod_{i\in U} z_i^{1/2}} \exp \left\{ -h \sum_{i\in U} (y_i - \beta x_i)^2 / 2z_i \right\} \quad (5.2.1) \]

where \( x_i > 0, z_i > 0 \) (\( i = 1, \ldots, N \)) are known and \(-\infty < \beta < \infty \) and \( h > 0 \) are unknown parameters.

Further, assume that the conjugate prior distribution of \( \theta \) is normal-gamma (see, Raiffa and Schlaifer, 1961, pp. 144-145) with the density

\[ g(\theta|\phi) \propto h^{(n_0)/2} \exp\{-hn_0(\beta - \beta_0)^2 / 2\} \cdot h_{m_0}^{1/2} \exp\{-hm_0 a_0 / 2\} \quad (5.2.2) \]
where $\phi = (n_0, \beta_0, m_0, a_0)$ is a vector of known parameters such that $n_0 \geq 0, m_0 \geq 0, a_0 > 0$, and $\delta(n_0) = 0$ if $n_0 = 0, = 1$ if $n_0 > 0$.

The model defined by (5.2.1) and (5.2.2) is designated as the Bayes model. For the Bayes model the Bayes prediction risk of $t$ with respect to squared error loss function is defined as

$$E(t - Y)^2 = E_\theta[E_\xi(t - Y)^2|\theta]$$

Here, $E_\xi(\cdot|\theta)(V_\xi(\cdot|\theta))$ and $E_\theta(\cdot)(V_\theta(\cdot))$ denote, respectively, expectations (variances) over the distributions of $y$ and $\theta$.

Since the posterior distribution of $\theta$ given $d$ is found to be the normal-gamma with the density given in (5.2.2) with $\phi_1 = (n_1, \beta_1, m_1, a_1)$ in place of $\phi$, where

$$n_1 = n_0 + \sum_{i \in s} \frac{x_i^2}{z_i}$$

$$\beta_1 = \left\{ n_0 \beta_0 + (n_1 - n_0) \bar{\beta} \right\}/n_1$$

$$m_1 = m_0 + m_0 + \delta(n_0) - \delta(n_1)$$

$$m_1 a_1 = m_0 a_0 + \sum_{i \in s} \left( y_i - \bar{\beta} x_i \right)^2 / z_i + n_0 (1 - n_0/n_1) (\bar{\beta} - \beta_0)^2$$

and

$$\bar{\beta} = \frac{\sum_{i \in s} x_i y_i / z_i}{\sum_{i \in s} x_i^2 / z_i}$$

Under squared error loss, the Bayes predictor of $Y$ is the posterior expectation of $Y$ given $d$, that is,

$$t_B = E[Y|d] = \sum_{i \in s} y_i + \sum_{j \in s} E_\theta E_\xi(y_i|d, \theta)$$

Since the posterior mean is $\beta_1$ the Bayes predictor is given by
\[ t_B = \sum_{i \in S} y_i + n_0 \beta_0 + \frac{(n_1 - n_0)\bar{\beta}}{n_1} \left( \sum_{j \in S} x_j \right) \] (5.2.3)

The Bayes prediction risk associated with \( t_B \) is given by

\[
E(t_B - Y)^2 = E[V(Y|d)] = E[V_\theta E_\xi(Y|\theta, d) + E_\xi V_\theta(Y|\theta, d)]
\]

\[
= E \left[ V_\theta \left\{ \beta \sum_{j \in S} x_j | d, \phi \right\} + E_\xi \left\{ \frac{1}{h} \sum_{j \in S} z_j | d, \phi \right\} \right]
\]

\[
= \frac{a_1 m_1}{m_1 - 2} \left[ \sum_{j \in S} z_j + \frac{1}{n_1} \left( \sum_{j \in S} x_j \right)^2 \right]
\]

For a diffuse prior distribution, i.e. \( n_0 \to 0 \) and \( m_0 \to 0 \), the limit of the Bayes predictor (5.2.3) is

\[ t_{LB} = \sum_{i \in S} y_i + \bar{\beta} \left( \sum_{j \in S} x_j \right) = \sum_{i \in S} y_i + \frac{\sum_{i \in S} x_i y_i / z_i}{\sum_{i \in S} x_i^2 / z_i} \left( \sum_{j \in S} x_j \right) \] (5.2.4)

### 5.2.2 Admissibility of \( t_{LB} \)

To prove admissibility of \( t_{LB} \) we use the limiting Bayes risk method. A more detailed account of this method is given by Lehmann and Casella (1998, p.325). Without loss of generality assume that \( h \) is known and equals to 1. For Model (2.1) the prediction risk of \( t \) with respect to squared error loss is defined by

\[ R_\xi(\beta, t) = E_\xi (t - Y)^2 \forall \beta \in \mathcal{B} \]

A predictor \( t_0 \in \mathcal{C} \), a class of predictors of \( Y \), is said to be admissible in \( \mathcal{C} \) with respect to \( \xi \) if there exists no predictor \( t_1 \in \mathcal{C} \) such that

\[ R_\xi(\beta, t_1) \leq R_\xi(\beta, t_0) \forall \beta \in \mathcal{B} \]

\[ < R_\xi(\beta, t_0) \text{for at least one } \beta \in \mathcal{B} \]
Theorem 5.1. Suppose that $y_i$, given $\beta$, are independently distributed as $N(\beta x_i, z_i)$, $i = 1, \ldots, N$ and the prior distribution of $\beta$ over $\mathcal{B}$ is $N(\mu, \tau^2)$. Under this model, the Bayes predictor of $Y$ is given by

$$t_B = \sum_{i \in S} y_i + \frac{\sum_{i \in S} x_i y_i / z_i + \mu \sigma^2 / \tau^2}{\sum_{i \in S} x_i^2 / z_i + \sigma^2 / \tau^2} \left( \sum_{j \in S} x_j \right)$$

Moreover, the limit of this Bayes predictor is (5.2.4) and is admissible.

Proof. It is not difficult to show that under the assumed model $t_B$ is the Bayes predictor. Further, to establish the admissibility of $t_{LB}$, suppose $t_{LB}$ is not admissible. Then there exists $t^*$ such that

$$R_\xi(\beta, t^*) \leq R_\xi(\beta, t_{LB}) = \sum_{j \in S} z_j + \frac{\left( \sum_{j \in S} x_j \right)^2}{\sum_{i \in S} x_i^2 / z_i} \quad \forall \beta \in \mathcal{B}$$

and strict inequality for at least one $\beta \in \mathcal{B}$. By continuity of $R_\xi(\beta, t)$, there exist $\varepsilon > 0$ and $\beta' < \beta''$, such that

$$R_\xi(\beta, t^*) \leq R_\xi(\beta, t_{LB}) - \varepsilon \quad \forall \beta' < \beta < \beta''$$

Let $R_\xi^*(\beta, t^*, \tau)$ be the average risk of $t^*$ with respect to the prior distribution $N(0, \tau^2)$ and let $R_\xi(\beta, t_{LB}, \tau)$ be the Bayes prediction risk $t_{LB}$ with $\mu = 0$. It follows that

$$R_\xi(\beta, t_{LB}, \tau) = \sum_{j \in S} z_j + \frac{\left( \sum_{j \in S} x_j \right)^2}{\sum_{i \in S} x_i^2 / z_i + 1 / \tau^2}$$

Hence

$$\frac{R_\xi(\beta, t_{LB}) - R_\xi^*(\beta, t^*, \tau)}{R_\xi(\beta, t_{LB}) - R_\xi(\beta, t_B, \tau)} = \frac{\frac{1}{\sqrt{2\pi\tau}} \int_{-\infty}^{\infty} \left( R_\xi(\beta, t_{LB}) - R_\xi(\beta, t^*) \right) \exp\left(-\beta^2 / 2\tau^2\right) d\beta}{\left( \sum_{j \in S} x_j \right)^2 \left[ \frac{1}{\sum_{i \in S} x_i^2 / z_i} - \frac{1}{\sum_{i \in S} x_i^2 / z_i + 1 / \tau^2} \right]}$$
\[
\geq \frac{1}{\sqrt{2\pi}} \varepsilon \int_{\beta'}^{\beta''} \{\exp(-\beta^2/2\tau^2)\} d\beta
\]
\[
\left(\sum_{j \in S} x_j\right)^2 \left(\sum_{i \in S} x_i^2/z_i\right)^{-1} \left[\tau \sum_{i \in S} x_i^2/z_i + 1/\tau\right]^{-1}
\]

The integrand converges monotonically to 1 as \(\tau \to \infty\) and by the Lebesgue monotone convergence theorem the integrand converges to \(\beta'' - \beta'\), and consequently the ratio converges to \(\infty\). Thus there exists \(\tau < \infty\) such that \(R_\tau(\beta, t_B, \tau) > R_\tau(\beta, t^*, \tau)\), which is a contradiction since \(t_B\) is the Bayes predictor. It follows that \(t_{LB}\) is admissible.

A more general Bayes estimator of \(Y\) can be obtained from \(t_{LB}\) by shifting the origin of \(y\)-values, i.e., substituting \(y_i = y_i - a_i\) for \(i = 1, \ldots, N\) and taking \(z_i = x_i/w_i\) as

\[
t_{LB}^* = \sum_{i \in S} y_i + \sum_{i \in S} a_j + \frac{\sum_{i \in S} w_i (y_i - a_i)}{\sum_{i \in S} w_i x_i} \left(\sum_{j \in S} x_j\right) \quad (5.2.5)
\]

In particular, for the choices \(w_i = (1 - np_i)/np_i\) and \(x_i = np_i\) the above estimator reduces to

\[
t_{LB}^* = \sum_{i \in S} \frac{y_i}{np_i} + \left(\sum_{i \in S} a_i - \sum_{i \in S} \frac{a_i}{np_i}\right)
\]

This is identical to the GD estimator, given in (5.1.1) for \(a_i = \beta x_i\), for \(i = 1, \ldots, N\).

**5.3. Regression coefficient is fixed-unknown**

Two well-known types of regression estimators have appeared in the literature, namely the generalized regression (GREG) estimator (see, Särndal et al., 1992) and the optimal (OPT) estimator (see, Montanari, 1998) for estimating finite population means or totals of survey variables. The asymptotic optimality of the
GREG estimator requires assumed working model to be true and hence the efficiency of it is vulnerable if model is misspecified. On the contrary, in the OPT estimator no superpopulation model is used and its asymptotic optimality is a strictly design based property.

5.3.1 The Optimal Estimator

The generalized difference (GD) estimator (5.1.1) has a design-variance

\[ V_p(\hat{y}_{GD}) = V_p(\hat{y}_{HH}) + \beta^2 V_p(\hat{x}_{HH}) - 2\beta \text{Cov}_p(\hat{y}_{HH}, \hat{x}_{HH}) \]  

(5.3.1)

where

\[ \text{Cov}_p(\hat{y}_{HH}, \hat{x}_{HH}) = N^2 \sum_{i \in U} \sum_{j \in U} \Delta_{ij} y_i x_j \]  

(5.3.2)

with

\[ \Delta_{ij} = (1/p_i - 1)/n \text{ if } i = j \text{ and } = -1/n \text{ if } i \neq j. \]

Here, \( V_p(\hat{y}_{HH}) \) and \( V_p(\hat{x}_{HH}) \) denote variances of the HH estimators of the \( y \) and \( x \) variables, respectively, and these expressions can be deduced from (5.3.2) by obvious modifications.

The variance (5.3.1) is minimized by assuming

\[ \beta = \frac{\text{Cov}_p(\hat{y}_{HH}, \hat{x}_{HH})}{V_p(\hat{x}_{HH})}. \]

The optimum value of \( \beta \) can be estimated in many ways. If we take HH estimators of variances and covariance between unbiased estimators \( \hat{y}_{HH} \) and \( \hat{x}_{HH} \), a design consistent estimator of \( \beta \) is given by

\[ \hat{\beta} = \frac{\text{cov}_p(\hat{y}_{HH}, \hat{x}_{HH})}{V_p(\hat{x}_{HH})} = \frac{\sum_{i \in U} \sum_{j \in U} \tilde{\Delta}_{ij} y_i x_j}{\sum_{i \in U} \sum_{j \in U} \tilde{\Delta}_{ij} x_i x_j} \]

where

\[ \tilde{\Delta}_{ij} = \frac{1}{n^2 p_i^2} \text{ if } i = j \text{ and } = -\frac{1}{n^2 (n - 1)p_i p_j} \text{ if } i \neq j \]
Then replacing $\beta$ by $\bar{\beta}$ in $\bar{Y}_{GD}$, we get

$$\bar{Y}_{OPT} = \bar{Y}_{HH} + \bar{\beta}(X - \bar{X}_{HH})$$

(5.3.3)

The main drawback of this estimator is that it requires estimating sampling variances and covariance. In this case the optimum variance is given by

$$\min V(\bar{Y}_{OPT}) = (1 - \rho^2_{\bar{\beta}}) V(\bar{Y}_{HH})$$

where $\rho_p = \text{cov}_p(\bar{Y}_{HH}, \bar{X}_{HH})/\sqrt{V_p(\bar{Y}_{HH})V_p(\bar{X}_{HH})}$.

5.3.2 The GREG Estimator

One method of using auxiliary information available at estimation stage is through regression estimation. For a single auxiliary variable the GREG estimator it is written as

$$\bar{Y}_{GREG} = \bar{Y}_{HH} + \hat{\beta}(X - \bar{X}_{HH})$$

(5.3.4)

where $\hat{\beta} = (\sum x_i^2/p_i\eta_i)^{-1} (\sum x_i y_i/p_i\eta_i)$ is a consistent regression estimator of $\beta$ for the superpopulation model

$$E_{\xi}(Y_i) = \beta x_i, \quad V_{\xi}(Y_i) = \sigma^2 \eta_i \quad \text{and} \quad C_{\xi}(Y_i, Y_j) = 0 \quad \text{for} \quad i \neq j \in U$$

where $\beta$ and $\sigma^2 > 0$ are the parameters, $\eta_i$ are known constants. Here $E_{\xi}, V_{\xi}$ and $C_{\xi}$ denote expected value, variance and covariance under the model, respectively. See Särndal et al. (1992) for more thorough coverage of the GREG estimator. This estimator is model-assisted estimator.
5.4. Simulation study and comparison of estimators

In this section, the estimators $t_{LB}$, $\hat{y}_{OPT}$ and $\hat{y}_{GREG}$, given in (5.2.4), (5.3.3) and (5.3.4), respectively, were compared empirically on 8 small and 9 large natural populations given in Appendix A, Table A.1.3 and Table A.2.2. For this using the pps sampling scheme, a sample of size n (8 and 20) is drawn from each of the population listed in Table 5.1 and Table 5.2 and the estimators $t_{LB}$, $\hat{y}_{OPT}$ and $\hat{y}_{GREG}$ were computed for each sample. This process was repeated $M = 5000$ times and the performance of the estimators was evaluated in terms of relative percentages bias (RB %) and relative efficiency (RE). Table 1 reports RBs and REs under the pps sampling.

<table>
<thead>
<tr>
<th>Popl.</th>
<th>$\hat{Y}_{HH}$</th>
<th>$\hat{Y}_{OPT}$</th>
<th>$\hat{Y}_{GREG}$</th>
<th>$t_{LB}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.00</td>
<td>1.25</td>
<td>1.17</td>
<td>1.43</td>
</tr>
<tr>
<td></td>
<td>(0.83)*</td>
<td>(0.98)</td>
<td>(1.17)</td>
<td>(-0.80)</td>
</tr>
<tr>
<td>2</td>
<td>1.00</td>
<td>1.29</td>
<td>1.27</td>
<td>1.76</td>
</tr>
<tr>
<td></td>
<td>(2.42)</td>
<td>(1.42)</td>
<td>(1.37)</td>
<td>(-0.23)</td>
</tr>
<tr>
<td>3</td>
<td>1.00</td>
<td>1.80</td>
<td>1.62</td>
<td>4.50</td>
</tr>
<tr>
<td></td>
<td>(0.31)</td>
<td>(0.18)</td>
<td>(0.01)</td>
<td>(1.90)</td>
</tr>
<tr>
<td>4</td>
<td>1.00</td>
<td>2.20</td>
<td>2.21</td>
<td>2.39</td>
</tr>
<tr>
<td></td>
<td>(0.27)</td>
<td>(0.03)</td>
<td>(0.02)</td>
<td>(0.99)</td>
</tr>
<tr>
<td>5</td>
<td>1.00</td>
<td>1.09</td>
<td>1.07</td>
<td>7.63</td>
</tr>
<tr>
<td></td>
<td>(-1.22)</td>
<td>(1.85)</td>
<td>(1.85)</td>
<td>(0.29)</td>
</tr>
<tr>
<td>6</td>
<td>1.00</td>
<td>1.12</td>
<td>1.11</td>
<td>1.43</td>
</tr>
<tr>
<td></td>
<td>(-1.13)</td>
<td>(-0.95)</td>
<td>(-0.91)</td>
<td>(-0.61)</td>
</tr>
<tr>
<td>7</td>
<td>1.00</td>
<td>71.30</td>
<td>64.21</td>
<td>70.46</td>
</tr>
<tr>
<td></td>
<td>(-4.23)</td>
<td>(-0.45)</td>
<td>(-0.55)</td>
<td>(-1.05)</td>
</tr>
</tbody>
</table>
### Table 5.2: RE (%) and RB (%) of estimators (large population)

<table>
<thead>
<tr>
<th>popl.</th>
<th>$\hat{Y}_{HH}$</th>
<th>$\hat{Y}_{OPT}$</th>
<th>$\hat{Y}_{GREG}$</th>
<th>$t^*_{LB}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.00</td>
<td>1.59</td>
<td>1.34</td>
<td>2.14</td>
</tr>
<tr>
<td></td>
<td>(1.77)*</td>
<td>(0.04)</td>
<td>(-0.56)</td>
<td>(-0.18)</td>
</tr>
<tr>
<td>2</td>
<td>1.00</td>
<td>37.41</td>
<td>39.14</td>
<td>66.49</td>
</tr>
<tr>
<td></td>
<td>(6.09)</td>
<td>(0.09)</td>
<td>(-1.54)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.00</td>
<td>78.30</td>
<td>81.76</td>
<td>129.71</td>
</tr>
<tr>
<td></td>
<td>(4.99)</td>
<td>(0.13)</td>
<td>(-0.60)</td>
<td>(-0.90)</td>
</tr>
<tr>
<td>4</td>
<td>1.00</td>
<td>5.00</td>
<td>2.64</td>
<td>4.11</td>
</tr>
<tr>
<td></td>
<td>(5.44)</td>
<td>(-3.24)</td>
<td>(-1.02)</td>
<td>(-2.62)</td>
</tr>
<tr>
<td>5</td>
<td>1.00</td>
<td>1.41</td>
<td>1.01</td>
<td>1.58</td>
</tr>
<tr>
<td></td>
<td>(1.82)</td>
<td>(0.66)</td>
<td>(-0.23)</td>
<td>(-0.20)</td>
</tr>
<tr>
<td>6</td>
<td>1.00</td>
<td>2.77</td>
<td>1.99</td>
<td>3.53</td>
</tr>
<tr>
<td></td>
<td>(3.83)</td>
<td>(1.44)</td>
<td>(1.63)</td>
<td>(-5.53)</td>
</tr>
<tr>
<td>7</td>
<td>1.00</td>
<td>23.85</td>
<td>23.68</td>
<td>24.17</td>
</tr>
<tr>
<td></td>
<td>(-0.14)</td>
<td>0.04</td>
<td>0.04</td>
<td>-0.07</td>
</tr>
<tr>
<td>8</td>
<td>1.00</td>
<td>43.32</td>
<td>42.91</td>
<td>43.39</td>
</tr>
<tr>
<td></td>
<td>(-0.11)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(-0.07)</td>
</tr>
<tr>
<td>9</td>
<td>1.00</td>
<td>1.90</td>
<td>1.90</td>
<td>1.93</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(-0.02)</td>
<td>(-0.02)</td>
<td>(0.26)</td>
</tr>
</tbody>
</table>

* Figures in the parentheses are the RBs of estimators.

**Remark 5.1.** Empirically it is observed that the performance of the estimator $\hat{Y}_{GREG}$ is identical to $\hat{Y}_{OPT}$ under the assumption $V(y_t) \propto x_t$ for all study populations under consideration.

Noteworthy observations from Table 5.1 and 5.2 are:

a) The absolute values of RBs of all estimators are within reasonable range.

b) $\hat{Y}_{OPT}$ Performs well as compared to $\hat{Y}_{GREG}$ except population 3 in table 5.2.

c) Overall the suggested Bayes estimator $t^*_{LB}$ is the best performer.
5.5. Conclusion

In survey sampling auxiliary information about the finite population is often available at the estimation stage. Utilizing this information more efficient estimators may be obtained. There exist several approaches, such as model-based, calibration, Bayesian etc., each of which provides a practical approach to incorporate auxiliary information at the estimation stage.

Hedayat and Sinha (1991, p. 102) contend that it is almost imperative to postulate an explicit relationship between the study variable and the auxiliary variable and it can often be assumed that a linear relationship exists between the two variable. In many survey populations, the relationship between $y_i$ and $x_i$ is often a straight line through the origin (e.g. see, Royall and Cumberland (1981a, b)) with a general variance structure $\nu(y_i) = \sigma_i^2 = \sigma^2 x_i^g$, $1 \leq g \leq 2$, which is usually the case in survey populations (see, e.g. Brewer (1963), Foreman (1991)). We have attempted to incorporate auxiliary information at estimation stage using Bayesian model.

The conclusions emerging from the simulation study can be summarized as follow.

a) Under the assumed Bayesian model, among the three estimators $\hat{Y}_{OPT}$, $\hat{Y}_{GREG}$, and $t^*_L$ with pps sampling, $t^*_L$ is very efficient when the best linear fit goes through the origin and the residuals from it are small. The absolute values of RBs of this estimator are all within reasonable range.
b) $\hat{Y}_{OPT}$ has performed better than $\hat{Y}_{GREG}$. The main drawback of this estimator is that it requires estimating sampling variances and covariance. However, this is not the problem for PPS sampling.