Directed by Dr. Sat Narain Gupta and Dr. Javid Shabbir. 117 pp.

The first order approximation of the theoretical mean square error and assumption of bivariate normality are very often used for the ratio type estimators for the population mean and variance. We have examined the adequacy of the first order approximation and the robustness of various ratio type estimators. We observed that the first order approximation for ratio type mean estimators and ratio type variance estimators works well if the sampling fraction is small and that departure from the assumption of bivariate normality is not a problem for large samples. We have also proposed some generalized mixture estimators which are combinations of the commonly used estimators. We have also extended the proposed generalized mixture estimators to the case when the study variable is sensitive and a non sensitive auxiliary variable is available. We have shown that the proposed generalized mixture estimators are more efficient than other commonly used estimators. An extensive simulation study and numerical examples are also presented.
GENERALIZED MIXTURE ESTIMATORS FOR THE FINITE POPULATION MEAN

by

Tanja Zatezalo

A Dissertation Submitted to the Faculty of The Graduate School at The University of North Carolina at Greensboro in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

Greensboro 2016

Approved by

Committee Co-Chair

Committee Co-Chair
To my husband, Milan

and our beautiful son, Stefan
This dissertation written by Tanja Zatezalo has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Co-Chair
Sat Narain Gupta

Committee Co-Chair
Javid Shabbir

Committee Members
Scott James Richter
Haimeng Zhang
Shanmugathasan Suthaharan

Date of Acceptance by Committee

Date of Final Oral Examination
ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisors Prof. Sat Narain Gupta and Prof. Javid Shabbir for their guidance, support and encouragement during this dissertation work. I would like especially to thank Dr. Sat Gupta for mentoring me through of all my PhD study. His continuous support, guidance, and dedication have helped me to grow as a researcher and finish this long journey. I would like to thank Prof. Javid Shabbir for collaboration in all of my research projects in this dissertation, his dedication and believing in me as a researcher.

Also, I would like to thank the rest of my thesis committee: Prof. Scott James Richter, Prof. Shanmugathasan Suthaharan, and Prof. Haimeng Zhang, for their insightful comments, encouragement, and input during this dissertation work. Thanks goes to our Director of Graduate Studies Prof. Greg Bell for his selfless guidance and all of his help.

I would like to thank my co-author Prof. Subhash Kumar Yadav from Department of Mathematics and Statistics (A Centre of Excellence) Dr. RML. Avadh University, Faizabad, India whose collaboration improved this work.

Last but not the least, I would like to thank my family for supporting and be motivation through all my hard work and long studies.
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>vii</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1. General Discussion on Generalized Mixture Estimators</td>
<td>1</td>
</tr>
<tr>
<td>1.2. Basic Ratio, Product, Regression and Exponential Estimators</td>
<td>3</td>
</tr>
<tr>
<td>1.3. Randomized Response Methodology</td>
<td>8</td>
</tr>
<tr>
<td>1.4. Motivation for this Work and Outline of the Dissertation</td>
<td>17</td>
</tr>
<tr>
<td>II. ADEQUACY OF THE FIRST ORDER APPROXIMATION FOR RATIO ESTIMATORS OF THE MEAN AND VARIANCE</td>
<td>19</td>
</tr>
<tr>
<td>2.1. Introduction</td>
<td>19</td>
</tr>
<tr>
<td>2.2. Some Ratio Estimators of the Mean</td>
<td>19</td>
</tr>
<tr>
<td>2.3. Some Ratio Estimators of the Variance</td>
<td>24</td>
</tr>
<tr>
<td>2.4. Comparisons of the Theoretical and Empirical Mean Square Errors</td>
<td>29</td>
</tr>
<tr>
<td>2.5. Robustness Against Departure From the Bivariate Normal Assumption</td>
<td>46</td>
</tr>
<tr>
<td>III. THE NEW GENERALIZED MIXTURE ESTIMATORS OF THE MEAN</td>
<td>47</td>
</tr>
<tr>
<td>3.1. Introduction</td>
<td>47</td>
</tr>
<tr>
<td>3.2. Some Existing Generalized Estimators</td>
<td>49</td>
</tr>
<tr>
<td>3.3. The Proposed Generalized Mixture Estimator I</td>
<td>52</td>
</tr>
<tr>
<td>3.4. The Proposed Generalized Mixture Estimator II</td>
<td>57</td>
</tr>
<tr>
<td>3.5. Efficiency Comparisons</td>
<td>64</td>
</tr>
<tr>
<td>3.6. Numerical Examples</td>
<td>65</td>
</tr>
<tr>
<td>3.7. Conclusion</td>
<td>68</td>
</tr>
<tr>
<td>IV. IMPROVED GENERALIZED MIXTURE ESTIMATORS WITH RRT MODELS</td>
<td>69</td>
</tr>
<tr>
<td>4.1. Introduction</td>
<td>69</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Modified Ratio and Product Estimators of Population Mean With the Characterising Constant, Bias and Mean Square Errors</td>
<td>21</td>
</tr>
<tr>
<td>2</td>
<td>The Modified Ratio Estimators of Population Variance With the Corresponding Characterising Constant, Bias and the Mean Square Error</td>
<td>27</td>
</tr>
<tr>
<td>3</td>
<td>Population Statistics for Various Bivariate Distributions with Positive Coefficient of Correlations</td>
<td>30</td>
</tr>
<tr>
<td>4</td>
<td>Population Statistics for Various Bivariate Distributions with Negative Coefficients of Correlation</td>
<td>31</td>
</tr>
<tr>
<td>5</td>
<td>The Ratio of the Empirical Mean Square Errors and the Theoretical Mean Square Errors for Some Ratio Estimators of Population Mean</td>
<td>31</td>
</tr>
<tr>
<td>6</td>
<td>The Ratio of the Empirical Mean Square Errors and the Theoretical Mean Square Errors for Some Product Estimators of Population Mean</td>
<td>38</td>
</tr>
<tr>
<td>7</td>
<td>The Ratio of the Empirical Mean Square Errors and the Theoretical Mean Square Errors for Some Ratio Estimators of Population Variance</td>
<td>42</td>
</tr>
<tr>
<td>8</td>
<td>Summary Statistics for the Real Populations Used in Comparing $\hat{\mu}<em>{GM}$ and $\hat{\mu}</em>{GMR}$ with other Mean Estimators</td>
<td>66</td>
</tr>
<tr>
<td>9</td>
<td>The Theoretical Percent Relative Efficiency for Various Mean Estimators</td>
<td>67</td>
</tr>
<tr>
<td>10</td>
<td>Summary Statistics for the Real Populations Used in Comparing $\hat{\mu}_{GRR}$ with other RRT Mean Estimators</td>
<td>80</td>
</tr>
<tr>
<td>11</td>
<td>The Theoretical Percent Relative Efficiency for RRT Mean Estimators</td>
<td>81</td>
</tr>
</tbody>
</table>
Table 12. Summary Statistics Used in Comparing $\hat{\mu}_{GM}$, $\hat{\mu}_{GMR}$ and $\hat{\mu}_{GRR}$ with other Mean Estimators ........................................ 85

Table 13. The Theoretical Percent Relative Efficiency (PRET) and the Empirical Percent Relative Efficiency (PREE) for the non-RRT Mean Estimators ........................................ 86

Table 14. The Theoretical Percent Relative Efficiency (PRET) and the Empirical Percent Relative Efficiency (PREE) for the RRT Mean Estimators ........................................ 88
CHAPTER I
INTRODUCTION

1.1 General Discussion on Generalized Mixture Estimators

The purpose of a sample survey is to obtain information about the population based on a random sample. By “population” we mean a group of units defined according to the objective of the survey. Thus the population may consist of all the fields under a specified crop or all the agricultural holdings larger than a specified size, as in an agricultural survey; or all the households having four or more children, as in a socio-economic survey. Of course, the population may also refer to the human population of a country. The information we want to get maybe, for example, the total number of units, such as the number of farms that grow corn; or aggregate values of various characteristics, such as the total area under corn. We may also look for the mean of various characteristics per unit, such as the mean household size; or the proportion of units which have certain characteristics, or the proportion of households having income over a given level or having five or more children.

In survey research, there are situations in which the information is available for every unit in the population. If a variable’s value is known for every unit of the population, then it is not a variable of direct interest. Instead it maybe employed to improve the sampling plan or to improve the estimation of another variable of interest. Such a variable is called an auxiliary variable. Ratio, product, and regression type estimators rely on the use of an auxiliary variable to estimate parameters of the study
variable. Auxiliary variables have been used by various authors in various estimation situations.

Cochran (1940)[6] introduced the use of an auxiliary variable at the estimation stage and proposed the ratio estimator for the population mean. It is well known that the ratio type mean estimator ensures better efficiency than the sample mean estimator if the study variable and an auxiliary variable have strong positive correlations. For situations when they are negatively correlated, the product estimator was introduced by Robson (1957)[47]. The product estimator is also more efficient than the sample mean estimator.

The regression estimator is used when the regression line between the study variable and the auxiliary variable does not pass through the origin. It is a well known that the regression estimator is more efficient than the ratio estimator and the sample mean estimators for $\rho_{yx} \neq 0$. Modified ratio, product, and regression type estimators have been introduced by different authors.

Development continued in the form of exponential estimators for different situations such as the work by Singh and Vishwakarma (2007) [67] in double sampling and Sanaullah et al. (2014) [50] in stratified two-phase random sampling. Singh et al. (2008) [64] proposed a ratio-product type exponential estimator which is more efficient than ordinary exponential ratio and product type estimators of Bahl & Tuteja (1991).

Grover & Kaur (2011) [12] introduced a regression-exponential type estimator of the mean. Subramani (2013) [77] proposed a generalized modified ratio estimator for estimating the population mean using the known population parameters of an auxiliary variable such as coefficient of variation, coefficient of kurtosis, coefficient of skewness, the coefficient of correlation, and various quartiles.

Asghar’s et al. (2014) [1] proposed the generalized exponential type estimator for the population variance. Following them, Shabbir and Gupta (2015) [53] proposed a new generalized exponential type estimator for the population variance which performs better than Asghar (2014) et al. estimator.

In this dissertation, some new generalized mixture estimators of the population mean of the study variable by combining the ratio, product, exponential, and regression estimators will be proposed. The main aim is to gain efficiency in comparisons to the existing generalized mixture estimators, and also use these estimators to estimate the population mean of the sensitive study variable when a non sensitive auxiliary variable is used.

1.2 Basic Ratio, Product, Regression and Exponential Estimators

Let $U = \{U_1, \ldots, U_N\}$ be a finite population of size $N$ and let $(y_i, x_i)$ be the values of the study variable $Y$ and an auxiliary variable $X$ on the $i$th unit $U_i$, $i = 1, \ldots, N$. 

Let a sample of size $n$ be drawn from this population, using simple random sampling without replacement. The goal is to estimate the population mean $\bar{Y} = \frac{1}{N} \sum_{i=1}^{N} y_i$. Let $S^2_y = \frac{1}{N-1} \sum_{i=1}^{N} (y_i - \bar{Y})^2$ be the population variance of the study variable $Y$. Let $\bar{X} = \frac{1}{N} \sum_{i=1}^{N} x_i$ be the population mean and $S^2_x = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{X})^2$ be the population variances of the auxiliary variable $X$. Let $\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$ and $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ be the sample means of the study variable and an auxiliary variable respectively. Let $S_{xy} = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{Y})(x_i - \bar{X})$ be the population covariance between the study variable and the auxiliary variable. We assume that the population mean $\bar{X}$ and the population variance $S^2_x$ of an auxiliary variable are known. Let $\rho_{yx}$ be the correlation coefficient between the study variable and an auxiliary. Also, assume $C_x = \frac{S_x}{\bar{X}}$ and $C_y = \frac{S_y}{\bar{Y}}$ are the coefficients of variation of the study variable $Y$ and an auxiliary variable $X$, and $C_{xy} = \frac{S_{xy}}{\bar{Y}\bar{X}}$ is the coefficient of covariance between $Y$ and $X$.

It is well known that the variance of the sample mean, the unbiased estimator, is given by $Var(\bar{y}) = \lambda \bar{Y}^2 C^2_y$, where $\lambda = \frac{1-f}{n}$ and $f = \frac{n}{N}$ is the sampling fraction. We give below some other commonly known men estimators.

### 1.2.1 The Ratio Estimator

The ordinary ratio estimator for the population mean $\bar{Y}$ of the study variable is given by Cochran (1940)[6] as:

$$t_R = \frac{\bar{y} \bar{X}}{\bar{x}}$$

(1.1)
The bias and the mean square error respectively of this estimator, up to the first order approximation, are given by:

$$\text{Bias}(t_R) = E(t_R - \bar{Y}) \approx \lambda \bar{Y} \left( C_x^2 - \rho_{yx} C_x C_y \right), \quad (1.2)$$

$$MSE(t_R) = E(t_R - \bar{Y})^2 \approx \lambda \bar{Y}^2 \left( C_x^2 - 2 \rho_{yx} C_x C_y + C_y^2 \right). \quad (1.3)$$

If the sample size $n$ is sufficiently large, then up to the first order of approximation, the ratio estimator will be more efficient than the ordinary sample mean estimator if

$$\rho_{yx} > \frac{C_x}{2C_y}. \quad (1.4)$$

For situations where $C_x \approx C_y$, condition (1.4) becomes $\rho_{yx} > \frac{1}{2}$.

### 1.2.2 The Product Estimator

The product estimator is used when the study variable $Y$ and the auxiliary variable $X$ are negatively correlated. The estimator introduced by Robson (1957)[47], and revised by Murthy (1964) [34] is given by:

$$t_P = \frac{\bar{y} \bar{x}}{\bar{X}}. \quad (1.5)$$

The exact bias of the product estimator is given by:

$$\text{Bias}(t_P) = E(t_p - \bar{Y}) = \lambda \frac{S_{yx}}{\bar{X}}. \quad (1.6)$$
The mean square error, up to the first order of approximation, is given by:

\[ MSE(t_P) = E(t_P - \bar{Y})^2 \approx \lambda\bar{Y}^2 \left( C_x^2 + 2\rho_{yx} C_x C_y + C_y^2 \right). \tag{1.7} \]

Up to the first order of approximation, the product estimator is more efficient than the ordinary sample mean if

\[ MSE(t_P) < Var(\bar{y}), \]

or if \( \rho_{yx} < -\frac{C_x}{2C_y}, \)

or if \( \rho_{yx} < -\frac{1}{2} \) when \( C_x \approx C_y. \tag{1.8} \)

\subsection*{1.2.3 The Regression Estimator}

The ratio type estimators often result in increased precision if the line of best fit of \( Y \) on \( X \) is linear and passes through the origin. If the line does not pass through the origin, it is better to use the regression estimator given by:

\[ t_{\text{Reg}} = \bar{y} + \hat{\beta}_{yx} (X - \bar{x}), \tag{1.9} \]

where \( \hat{\beta}_{yx} = \frac{s_{xy}}{s_x^2} \) is the sample regression coefficient between \( Y \) and \( X \). The bias of the regression estimator, up to the first order of approximation, is given by:

\[ \text{Bias}(t_{\text{Reg}}) = E(t_{\text{Reg}} - \bar{Y}) \approx -\lambda\beta_{yx} \left\{ \frac{\mu_{12}}{\mu_{11}} - \frac{\mu_{03}}{\mu_{02}} \right\}, \tag{1.10} \]

where \( \beta_{yx} = \frac{s_{yx}}{s_x^2} \) is the population regression coefficient between the study variable \( Y \) and the auxiliary variable \( X \), and \( \mu_{rs} = \frac{1}{N-1} \sum_{i=1}^{N} (y_i - \bar{Y})^r (x_i - \bar{X})^s \). Also \( s_x^2 = \)
\[ \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \] is the sample variance of \( X \) and \( s_{xy} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x}) (y_i - \bar{y}) \) is the sample covariance between \( X \) and \( Y \).

The mean square error, up to the first order of approximation, is given by:

\[ MSE(t_{Reg}) = E (t_{Reg} - \bar{Y})^2 \approx \lambda \bar{Y}^2 C_y^2 (1 - \rho_{yx}^2), \quad (1.11) \]

The conditions under which the regression estimator is more efficient than the ordinary sample mean and ratio estimator are given below:

1. the regression estimator \( t_{Reg} \) is more efficient than the ordinary sample mean \( \bar{y} \) if

\[ MSE(t_{Reg}) < Var(\bar{y}), \quad \text{or if} \quad C_y^2 - C_y (1 - \rho_{yx}^2) > 0, \quad \text{or if} \quad \rho_{xy} \neq 0, \quad \text{and} \]

2. the regression estimator \( t_{Reg} \) is more efficient than the ratio estimator \( t_R \) if

\[ MSE(t_{Reg}) < MSE(t_R), \quad \text{if} \quad C_x^2 - 2\rho_{yx} C_x C_y + \rho_{yx}^2 C_y^2 > 0, \quad \text{or if} \quad (C_x - \rho_{yx} C_y)^2 > 0. \]

If the relationship between \( Y \) and \( X \) is linear, and passes through the origin, then the two estimators are equally efficient.

1.2.4 Bahl & Tuteja Exponential Estimators

The exponential type estimators are often used to improve efficiencies of the ratio and product type estimators and were introduced by Bahl and Tuteja (1991)\[2\] as:

\[ t_{ER} = \bar{y} \exp \left( \frac{\bar{X} - \bar{x}}{X + \bar{x}} \right), \quad \text{and} \] (1.12)

\[ t_{EP} = \bar{y} \exp \left( \frac{\bar{x} - \bar{X}}{X + \bar{x}} \right). \] (1.13)
The exponential part helps, since it captures the auxiliary variable effect for a longer duration. The bias of the exponential estimators, up to the first order of approximation, are given by:

\[
Bias(t_{ER}) = (t_{ER} - \bar{Y}) \approx \lambda \bar{Y} \left( \frac{3}{8} C_x^2 - \frac{1}{2} \rho_y x y C_y C_x \right), \quad \text{and} \quad (1.14)
\]

\[
Bias(t_{EP}) = (t_{EP} - \bar{Y}) \approx \lambda \bar{Y} \left( \frac{1}{2} \rho_y x y C_y C_x - \frac{1}{8} C_x^2 \right). \quad (1.15)
\]

The mean square error of the exponential ratio and product type estimators, up to the first order of approximation, are given by:

\[
MSE(t_{ER}) = (t_{ER} - \bar{Y})^2 \approx \frac{1}{4} \lambda \bar{Y}^2 \left( 4C_y^2 + C_x^2 - 4\rho_y x y C_y C_x \right), \quad \text{and} \quad (1.16)
\]

\[
MSE(t_{EP}) = (t_{EP} - \bar{Y})^2 \approx \frac{1}{4} \lambda \bar{Y}^2 \left( 4C_y^2 + 4\rho_y x y C_y C_x + C_x^2 \right). \quad (1.17)
\]

### 1.3 Randomized Response Methodology

In this dissertation, we also want to discuss mean estimators in situations where the study variable is sensitive and cannot be observed directly. This is one of the most important issues in behavioral and social sciences. The respondents are sometimes asked a sensitive question such as their personal income, experiencing feelings of low self-worth and powerlessness, sexual orientation, number of sexual partners in last two years, number of miscarriages or abortions etc. To circumvent the social desirability bias (the tendency in people to present themselves in a socially acceptable light when
they are confronted with a sensitive question) several methods have been developed. The randomized response technique (RRT) is one such method. In this dissertation our focus is on situations when the study sensitive variable $Y$ can not be observed directly, but a highly correlated non-sensitive auxiliary variable $X$ is observed directly. For example, the study variable may be the number of miscarriages or abortions, and the non-sensitive auxiliary question may be the number of children for a woman. The Optional RRT models are models in which a respondent who considers a question sensitive provides a scrambled answer, and the rest provide a true answer to the sensitive question. This was discussed in Gupta et al. (2002,2006,2010)[14,15,18] and Kalucha et al.(2015)[27]. Our focus here is on non-optional RRT models and we will introduce a few of these models proposed by different authors.

1.3.1 Warner’s (1971) model

Warner’s (1971) [84] model is the quantitative additive version of Warner (1965)[83] Binary Randomized Response Technique and works as follows:

For a simple random sample with replacement, let $Y$ the true response, and $S$ be a scrambling variable with known mean $E(S) = \mu_s$ and known variance $\sigma_s^2$. The population mean $\mu_Y$ and the population variance $\sigma_Y^2$ of the study variable are unknown. Also, assume that the true response $Y$ and scrambling variable $S$ are independent. The reported response $Z$ is the sum of the true response and the scrambling variable, and is given by:

$$Z = Y + S.$$ (1.18)
Since

\[ E(Z) = \mu_Y + \mu_S, \]  

(1.19)

it follows that an unbiased estimator of the sensitive variable mean is given by:

\[ \hat{\mu}_Y = Z - \mu_S. \]  

(1.20)

The variance of this estimator is given by:

\[ Var(\hat{\mu}_Y) = Var(\bar{Z}) = \frac{\sigma_Y^2}{n} + \frac{\sigma_S^2}{n}. \]  

(1.21)

The second term in (1.21) is the “penalty” for randomizing. Also, note that an unbiased estimator for the variance is given by:

\[ \hat{Var}(\hat{\mu}_Y) = \frac{s_z^2}{n}, \]  

where \( s_z^2 \) is the sample variance of the reported responses.

### 1.3.2 Sousa et al. (2010) Ratio Estimator

Many authors have estimated the mean of a sensitive variable when the primary variable is sensitive and there is no auxiliary variable available. Sousa et al. (2010)[69] proposed the ratio estimator of the mean for the sensitive variable \( Y \) which has a strong positive correlation with a non-sensitive variable \( X \). The model works as follows:

Let \( Y \) be the sensitive study variable, and \( X \) be the non-sensitive variable which is strongly (positively) correlated with \( Y \). Let \( S \) be a scrambling variable independent of \( Y \) and \( X \). Assume that the population mean \( \bar{X} \) and the population variance \( S_x^2 \)
of the auxiliary variable are known. Also, assume that the population mean and the
population variance of the scrambling variable are known, and given as $\mu_S = 0$ and
$\sigma^2_s$. The population mean $Y$ and the population variance $S^2_y$ of the sensitive study
variable are unknown. The respondent is asked to report a scrambled response for $Y$,
but is asked to provide a true response for $X$. The reported response is the sum of
the sensitive variable and the scrambling variable, and is given by:

$$Z = Y + S.$$  \hspace{1cm} (1.22)

Note that $E(Z) = E(Y)$ since $\mu_s = 0$.

An estimator of the mean of the sensitive variable ($Y$) when the information on
($X$) is ignored, is the ordinary sample mean given by:

$$\hat{\mu}_Y = \bar{z}.$$  \hspace{1cm} (1.23)

The mean square error of this unbiased estimator, when sampling is without replace-
ment, is given by:

$$MSE(\hat{\mu}_Y) = \lambda S^2_z = \lambda \left( S^2_y + \sigma^2_s \right),$$  \hspace{1cm} (1.24)

where $S^2_y = \frac{1}{N-1} \sum_{i=1}^{N} (y_i - \bar{Y})^2$, $\sigma^2_s = \frac{1}{N-1} \sum_{i=1}^{N} (s_i - \mu_s)^2$ and $S^2_z = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{X})^2$ are the population variances of the study variable, the scrambling variable and
the auxiliary variable, respectively and $\lambda = \frac{1-f}{n}$ where $f = \frac{n}{N}$.

Sousa et al.(2010)[69] proposed the ratio estimator given by:

$$\hat{\mu}_R = \bar{z} \left( \frac{\bar{X}}{\bar{x}} \right).$$  \hspace{1cm} (1.25)
The bias of this estimator, up to the second order of approximation is given as:

\[
Bias^{(2)}(\hat{\mu}_R) = E(\hat{\mu}_R - \bar{Y}) \approx Bias^{(1)}(\hat{\mu}_R) + 3\lambda^3\bar{Y} \left(C_x^4 - \rho_{zx}C_xC_z^3\right), \tag{1.26}
\]

where the bias up to the first order of approximation, is given by:

\[
Bias^{(1)}(\hat{\mu}_R) \approx \lambda\bar{Y} \left(C_x^2 - \rho_{zx}C_xC_z\right). \tag{1.27}
\]

The mean squared error, correct up to the second order of approximation, is given by:

\[
MSE^{(2)}(\hat{\mu}_R) \approx MSE^{(1)}(\hat{\mu}_R) + 3\lambda^2\bar{Y}^2C_x^2 \left[1 + 2\rho_{zx}^2\right]C_x^2 + 3C_x^2 - 6\rho_{zx}C_xC_z \tag{1.28}
\]

where

\[
MSE^{(1)}(\hat{\mu}_R) \approx \lambda^2\bar{Y}^2 \left(C_x^2 + C_x^2 - 2\rho_{zx}^2C_x^3\right) \tag{1.29}
\]

is the corresponding mean square error up to the first of order of approximation. The difference between the two approximations for the mean square errors is given by:

\[
3\lambda^2\bar{Y}^2C_x^2 \left[(1 + 2\rho_{zx}^2)C_x^2 + 3C_x^2 - 6\rho_{zx}C_xC_z\right] \tag{1.30}
\]

The difference (1.30) converges to zero as \(n \to N\).

### 1.3.3 Gupta et al. (2012) Ordinary Regression Estimator Using RRT

Gupta et al.(2012)[19] proposed an ordinary regression estimator where the RRT estimator of the population mean \(\bar{Y}\) of the sensitive study variable is improved by using a non-sensitive auxiliary variable \(X\). The RRT regression estimator is given by:

\[
\hat{\mu}_{Reg} = \bar{z} + \hat{\beta}_{zx}(\bar{X} - \bar{x}) \tag{1.31}
\]
where $\hat{\beta}_{zx} = \frac{s_{zx}}{s_x^2}$ is the sample regression coefficient between $Z$ and $X$, and $Z = Y + S$ is the scrambled response, where $Y$ is the true response and $S$ is a scrambling variable.

The bias, up to the first order of approximation, is given by:

$$
Bias(\hat{\mu}_{Reg}) \approx -\beta_{zx} \lambda \left\{ \frac{\mu_{12}}{\mu_{11}} - \frac{\mu_{03}}{\mu_{02}} \right\} 
$$

(1.32)

where $\beta_{zx} = \frac{s_{zx}}{s_x^2}$ is the population regression coefficient and $\mu_{rs} = \frac{1}{N-1} \sum_{i=1}^{N} (z_i - \bar{Z})(x_i - \bar{X})$. Also note that the following holds:

$$
\beta_{zx} = \frac{S_{zx}}{S_x^2} = \frac{S_{yx}}{S_x^2} = \rho_{yx} \frac{S_y}{S_x} \quad \text{and} \quad \rho_{zx} = \frac{\rho_{yx}}{\sqrt{1 + \frac{s_y^2}{s_x^2}}}, 
$$

(1.33)

where $\rho_{yx}$ and $\rho_{zx}$ are the coefficients of correlation between $y$ and $x$, and $z$ and $x$, respectively.

The mean square error, up to the first order of approximation, is given by:

$$
MSE(\hat{\mu}_{Reg}) \approx \lambda Y^2 C_z^2 (1 - \rho_{zx}^2) = \lambda S_y^2 \left[ \left(1 + \frac{s_x^2}{s_y^2}\right) - \rho_{yx}^2 \right].
$$

(1.34)

The conditions under which the RRT regression estimator is more efficient than the RRT ratio estimator and the RRT sample mean are given by:

(1) The RRT regression estimator $\hat{\mu}_{Reg}$ is more efficient than the RRT sample mean estimator if $\rho_{yx}^2 > 0$, and

(2) The RRT regression estimator $\hat{\mu}_{Reg}$ is more efficient than the RRT ratio estimator $\hat{\mu}_R$ if $(C_x - C_z \rho_{zx})^2 > 0$. 

13
These conditions will always hold, indicating that up to the first order of approximation, the regression estimator performs better than the ordinary RRT sample mean and the RRT ratio estimator.

1.3.4 Gupta et al. (2012) Generalized Regression-Cum-Ratio Estimator

Many authors have used regression-cum-ratio estimators that combine the regression estimator and the ratio estimator. These include Ray and Singh (1981)[44], Perri (2004)[39], and Kadilar and Cingi (2004)[23]. Gupta et al. (2012)[19] proposed a similar hybrid estimator, as a generalized regression-cum-ratio estimator. The main idea was to see if further gains can be achieved by using a generalized regression-cum-ratio estimator, as compared to the Gupta et al (2012)[19] RRT regression estimator. The model works under the same conditions as the RRT regression estimator, and is given by:

\[
\hat{\mu}_{GRR} = \left[ k_1 \bar{Z} + k_2 (\bar{X} - \bar{x}) \right] \left( \frac{\bar{X}}{\bar{x}} \right),
\]  

where \( k_1 \) and \( k_2 \) are suitably chosen parameters. The bias, up to the first order of approximation, is given by:

\[
\text{Bias}(\hat{\mu}_{GRR}) \approx (k_1 - 1)\bar{Z} + \lambda k_1 \bar{Z} \left( C_x^2 - \rho_{zx} C_z C_x \right) + \lambda k_2 \bar{X} C_x^2. \]

The minimum mean square error of the generalized regression-cum-ratio estimator, at the optimum values of \( k_1 \) and \( k_2 \) i.e.,

\[
k_{1(\text{opt})} = \frac{1 - \lambda C_x^2}{1 - \lambda [C_x^2 - C_z^2 (1 - \rho_{zx}^2)]}, \quad \text{and} \quad k_{2(\text{opt})} = \frac{\bar{Y}}{\bar{X}} \left[ 1 + k_{1(\text{opt})} \left( \frac{\rho_{zx} C_z}{C_x} - 2 \right) \right],
\]
is given by:

\[
MSE(\hat{\mu}_{GRR})_{min} \approx \bar{Y}^2 \frac{\lambda C_z^2 [1 - \rho_{zx}^2] [1 - \lambda C_x^2]}{\lambda C_z^2 [1 - \rho_{zx}^2] + [1 - \lambda C_x^2]}.
\] (1.37)

The conditions under which the generalized regression-cum-ratio estimator is more efficient than the ordinary RRT sample mean, RRT ratio estimator, and RRT regression estimator are given below:

(1) 

\[
MSE(\hat{\mu}_{GRR})_{min} < MSE(\hat{\mu}_Y) \quad \text{if} \quad \lambda (S_y^2 + S_s^2) > 0 \quad (1.38)
\]

(2) 

\[
MSE(\hat{\mu}_{GRR})_{min} < MSE(\hat{\mu}_R) \quad \text{if} \quad \left( \frac{C_x}{C_z} - \rho_{zx} \right)^2 + \frac{\lambda C_z^2 (1 - \rho_{zx}^2)}{\lambda C_z^2 (1 - \rho_{zx}^2) + (1 - \lambda C_x^2)} > 0, \quad \text{and} \quad (1.39)
\]

(3) 

\[
MSE(\hat{\mu}_{GRR})_{min} < MSE(\hat{\mu}_{Reg}) \quad \text{if} \quad \lambda C_z^2 (1 - \rho_{zx}^2) > 0. \quad (1.40)
\]

From these conditions, which always hold true, we can conclude that the generalized regression-cum-ratio estimator with optimal coefficients is always better than the ordinary RRT sample mean, RRT regression and RRT ratio estimators.

1.3.5 Koyuncu et al. (2014) Generalized Exponential Estimator

Many authors have studied exponential type estimators when the study variable \(Y\) is non sensitive. These include Bahl & Tuteja (1991)[2], Shabbir and Gupta
(2007)[51], Grover and Kaur (2011)[12] and Koyuncu (2012)[31]. Following Gupta et al. (2012)[19] and Bahl & Tuteja (1991)[2], Koyuncu et al. (2014)[32] proposed a generalized exponential type estimator of the mean $\bar{Y}$ of the sensitive study variable utilizing a non-sensitive auxiliary variable $X$. The model works under the same assumptions as the ordinary regression estimator and the generalized regression-cum-ratio estimator, and is given by:

$$\hat{\mu}_{GE} = \left[ w_1 \bar{z} + w_2 (\bar{X} - \bar{x}) \right] \exp \left( \frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right),$$  \hspace{1cm} (1.41)$$

where $w_1$ and $w_2$ are the model parameters. The bias of this generalized exponential estimator, up to the first order of approximation, is given as:

$$\text{Bias}(\hat{\mu}_{GE}) \approx (w_1 - 1) \bar{Y} + \lambda w_1 \bar{Y} \left( \frac{3}{8} C_x^2 - \frac{1}{2} \rho_{zx} C_x C_{zx} \right) + \frac{1}{2} w_2 \lambda \bar{X} C_x^2,$$  \hspace{1cm} (1.42)$$

The minimum mean square error of generalized exponential estimator, up to the first order of approximation, at the optimum values of $w_1$ and $w_2$ i.e.,

$$w_{1(\text{opt})} = \frac{1 - \frac{1}{8} \lambda C_x^2}{1 + \lambda C_x^2 (1 - \rho_x^2)}, \hspace{1cm} \text{and}$$

$$w_{2(\text{opt})} = \frac{\bar{Y}}{\bar{X}} \left[ \frac{1}{2} - w_{1(\text{opt})} \left( 1 - \rho_{zx} \frac{C_z}{C_x} \right) \right],$$
is given by:

\[
MSE_{\text{min}}(\hat{\mu}_{GE}) \approx \bar{Y}^2 \left[ \left(1 - \frac{1}{4} \lambda C_x^2\right) - \frac{(1 - \frac{1}{8} \lambda C_x^2)^2}{1 + \lambda C_x^2 (1 - \rho_{zx}^2)} \right], \quad \text{or} \quad (1.43)
\]

\[
MSE_{\text{min}}(\hat{\mu}_{GE}) \approx \left\{ \frac{MSE(\hat{\mu}_{\text{Reg}})}{1 + \frac{MSE(\hat{\mu}_{\text{Reg}})}{\bar{Y}^2}} - \frac{\lambda C_x^2 \left[MSE(\hat{\mu}_{\text{Reg}}) + \lambda \frac{1}{16} C_x^2 \bar{Y}^2 \right]}{4 \left[1 + \frac{MSE(\hat{\mu}_{\text{Reg}})}{\bar{Y}^2} \right]} \right\}. \quad (1.44)
\]

1.4 Motivation for this Work and Outline of the Dissertation

Many generalized mean estimators have been discussed in the earlier sections. The main motivation of this dissertation is to improve efficiency of some existing estimators by introducing some new generalized mixture estimators. The second motivation is to use the proposed generalized mixture estimators in the situations when the sensitive study variable cannot be observed directly and a non-sensitive auxiliary variable is available.

An outline of the work discussed in various chapters is given below.

Chapter 1 provides an introduction to the ordinary ratio, product, regression and exponential type estimators. Also some generalized estimators are discussed. We also provide an introduction to RRT models for the quantitative response. Modification of these RRT estimators using non-sensitive auxiliary variable is also introduced.

Chapter 2 focuses on the comparisons of the empirical mean square errors and the corresponding theoretical mean square errors for various ratio and product type mean estimators, as well as for some ratio type variance estimators. The purpose is to examine the adequacy of the first order approximation which is generally used in the calculation of mean square errors for rato estimators. Also we examine the robustness
of these estimators since the main assumption for such situations is that $Y$ and $X$ have a bivariate normal distribution. A simulation study shows that the mean square error, up to the first order of approximation, generally works well whenever the sampling fraction ($\frac{n}{N}$) is small. Also, we have observed the departure from the assumption of bivariate normality is not a problem when the sample size is large.

Chapter 3 introduces the generalized mixture estimators proposed by Zatezalo et al. (2016) [87], with mathematical derivations for the bias and mean square errors, up to the first order of approximation. The optimum values of the parameters involved, and the optimum mean square error, up to the first order of approximation, are derived. Corresponding theoretical and empirical comparisons with some commonly used generalized estimators are also presented.

Chapter 4 discusses the ordinary ratio, regression and some generalized mixture estimators when the study variable is sensitive in nature and a non-sensitive auxiliary variable is available. The mathematical derivations for the bias and the mean square error, up to the first order of approximation, are presented. Also, the minimum mean square errors for two special cases are derived. The efficiency comparisons with some existing RRT estimators of the sensitive variable in the presence of an auxiliary variable are also presented. Results of a numerical study are given at the end of this chapter.

Chapter 5 presents simulation results where we compare the estimators proposed in Chapter III and Chapter IV with different existing estimators.

Chapter 6 gives some concluding remarks and future research directions.
CHAPTER II

ADEQUACY OF THE FIRST ORDER APPROXIMATION FOR RATIO ESTIMATORS OF THE MEAN AND VARIANCE

2.1 Introduction

In many studies the first order approximation for the theoretical mean square error has been used for ratio type estimators for the population mean and variance. Bivariate normality is another commonly used assumption. The main focus of this chapter is on examining the adequacy of the first order approximation and also on examining the robustness of ratio estimators against departure from bivariate normality. We have calculated the theoretical mean square errors for many ratio type estimators, based on first order approximation, and the corresponding empirical mean square errors. We observed that the first order approximation for the ratio type mean and variance estimators generally works well as long as the sampling fraction is small. We also observed that departure from the assumption of bivariate normality is not a serious handicap for large samples. We will use the terminology introduced in Chapter I.

2.2 Some Ratio Estimators of the Mean

Often the characteristic \( Y \) under study is closely related to an auxiliary variable \( X \), and summary data on \( X \), such as the population mean \( \bar{X} \) and the population variance \( S^2_x \), are readily available. In such a situation it is convenient to consider estimators of the population mean \( \bar{Y} \) and population variance \( S^2_y \) that use information about \( X \). Those estimators are generally more efficient than those based on a sample of \( Y \).
alone if the correlation between $X$ and $Y$ is strong. Many modifications of the ratio and product estimators have been done to improve their efficiency by using a number of known parameters of the auxiliary variable such as the coefficient of variation $C_x$, the coefficient of kurtosis $\beta_2(x)$, standard deviation $\sigma_x$, the coefficient of skewness $\beta_1(x)$, the correlation coefficient between the study variable and an auxiliary variable $\rho_{yx}$ and the quartiles $Q_i$’s. Sisodia and Dwivedi (1981)[68] have suggested a modified ratio estimator using the coefficient of variation $C_x$ of an auxiliary variable $X$ for estimating the population mean $\bar{Y}$. Upadhyaya and Singh (1999)[80] suggested another modified ratio estimator using a linear combination of the coefficient of variation $C_x$ and coefficient of the kurtosis $\beta_2(x)$. Singh and Tailor (2003) [62] proposed another estimator using the correlation coefficient $\rho_{yx}$ between $X$ and $Y$. By using the population variance $S_x^2$ of an auxiliary variable $X$, Singh (2003)[56] proposed another modified ratio estimator. Also, Singh used a linear combination of the coefficient of kurtosis $\beta_2(x)$ and standard deviation $\sigma_x$, and the coefficient of skewness $\beta_1(x)$ and standard deviation $\sigma_x$ for estimating the population mean of the study variable $\bar{Y}$. Motivated by Singh (2003)[56], Yan and Tian (2010)[89] used a linear combination of the coefficient of kurtosis $\beta_2(x)$ and the coefficient of skewness $\beta_1(x)$, and the coefficient of variation $C_x$ and the coefficient of skewness $\beta_1(x)$ of the auxiliary variable $X$. More recently, Subramani and Kumarapandiyan (2013)[76] suggested a new modified ratio estimator using known population median $M_d$ of an auxiliary variable. Subramani and Kumarapandiyan (2012, 2013)[71–73] have also suggested modified ratio estimators using the known median and the coefficient of kurtosis, median and coefficient of skewness, median and the coefficient of variation, and median and the coefficient of correlation.

We now introduce some of the ratio type estimators and product type estimators with corresponding mean square errors for the purpose of examining the adequacy of the first order of approximation and robustness. These include the classical ratio estimator $t_R$, the product estimator $t_P$ and the exponential ratio and product type estimator $t_{ER}$, all introduced in Chapter I. We also include many modified ratio and product estimators with corresponding characterising constants, the bias and the mean square error as given in Table 1. This was also discussed by Zatezalo et al. (2016)[86].

Table 1. Modified Ratio and Product Estimators of Population Mean With the Characterising Constant, Bias and Mean Square Errors

<table>
<thead>
<tr>
<th>Estimator</th>
<th>Constant $\theta_i$</th>
<th>Bias</th>
<th>Mean squared error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_1 = \bar{y} \left( \frac{\bar{x} + C_x}{\bar{x}} \right)$</td>
<td>$\theta_1 = \frac{\bar{x}}{\bar{x} + C_x}$</td>
<td>$\lambda \bar{Y} \left[ \theta_1 C_x^2 (\theta_1 - C) \right]$</td>
<td>$\lambda \bar{Y}^2 \left[ C_y^2 + \theta_1 C_x^2 (\theta_1 - 2C) \right]$</td>
</tr>
<tr>
<td>$t_2 = \bar{y} \left( \frac{\bar{x} + C_x}{\bar{x}} \right)$</td>
<td>$\theta_2 = \frac{\bar{x}}{\bar{x} + C_x}$</td>
<td>$\lambda \bar{Y} \left[ \theta_2 C_x^2 (\theta_2 + C) \right]$</td>
<td>$\lambda \bar{Y}^2 \left[ C_y^2 + \theta_2 C_x^2 (\theta_2 + 2C) \right]$</td>
</tr>
<tr>
<td>$t_3 = \bar{y} \left( \frac{\beta_3(x) \bar{x} + C_x}{\beta_2(x) \bar{x} + C_x} \right)$</td>
<td>$\theta_3 = \frac{\beta_3(x) \bar{x}}{\beta_2(x) \bar{x} + C_x}$</td>
<td>$\lambda \bar{Y} \left[ \theta_3 C_x^2 (\theta_3 + C) \right]$</td>
<td>$\lambda \bar{Y}^2 \left[ C_y^2 + \theta_3 C_x^2 (\theta_3 + 2C) \right]$</td>
</tr>
</tbody>
</table>

Continued on next page
<table>
<thead>
<tr>
<th>Estimator</th>
<th>Constant $\theta_i$</th>
<th>Bias</th>
<th>Mean square error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_4 = \bar{y} \left( \frac{C_x \bar{x} + \beta_2(x)}{C_x \bar{x} + \beta_2(x)} \right)$</td>
<td>$\theta_4 = \frac{C_x \bar{x}}{C_x \bar{x} + \beta_2(x)}$</td>
<td>$\lambda \bar{Y} \theta_4 C_x^2 (\theta_4 + C)$</td>
<td>$\lambda \bar{Y}^2 \left[ C_y^2 + \theta_4 C_x^2 (\theta_4 + 2C) \right]$</td>
</tr>
<tr>
<td>$t_5 = \bar{y} \left( \frac{\bar{x} + \sigma_x}{\bar{x} + \sigma_x} \right)$</td>
<td>$\theta_5 = \frac{\bar{x}}{\bar{x} + \sigma_x}$</td>
<td>$\lambda \bar{Y} \theta_5 C_x^2 (\theta_5 + C)$</td>
<td>$\lambda \bar{Y}^2 \left[ C_y^2 + \theta_5 C_x^2 (\theta_5 + 2C) \right]$</td>
</tr>
<tr>
<td>$t_6 = \bar{y} \left( \frac{\beta_1(x) \bar{x} + \sigma_x}{\beta_1(x) \bar{x} + \sigma_x} \right)$</td>
<td>$\theta_6 = \frac{\beta_1(x) \bar{x}}{\beta_1(x) \bar{x} + \sigma_x}$</td>
<td>$\lambda \bar{Y} \theta_6 C_x^2 (\theta_6 + C)$</td>
<td>$\lambda \bar{Y}^2 \left[ C_y^2 + \theta_6 C_x^2 (\theta_6 + 2C) \right]$</td>
</tr>
<tr>
<td>$t_7 = \bar{y} \left( \frac{\beta_2(x) \bar{x} + \sigma_x}{\beta_2(x) \bar{x} + \sigma_x} \right)$</td>
<td>$\theta_7 = \frac{\beta_2(x) \bar{x}}{\beta_2(x) \bar{x} + \sigma_x}$</td>
<td>$\lambda \bar{Y} \theta_7 C_x^2 (\theta_7 + C)$</td>
<td>$\lambda \bar{Y}^2 \left[ C_y^2 + \theta_7 C_x^2 (\theta_7 + 2C) \right]$</td>
</tr>
<tr>
<td>$t_8 = \bar{y} \left( \frac{\bar{x} + \rho x}{\bar{x} + \rho x} \right)$</td>
<td>$\theta_8 = \frac{\bar{x}}{\bar{x} + \rho x}$</td>
<td>$\lambda \bar{Y} \theta_8 C_x^2 (\theta_8 - C)$</td>
<td>$\lambda \bar{Y}^2 \left[ C_y^2 + \theta_8 C_x^2 (\theta_8 - 2C) \right]$</td>
</tr>
<tr>
<td>$t_9 = \bar{y} \left( \frac{\bar{x} + \rho x}{\bar{x} + \rho x} \right)$</td>
<td>$\theta_9 = \frac{\bar{x}}{\bar{x} + \rho x}$</td>
<td>$\lambda \bar{Y} \theta_9 C_x^2 (\theta_9 + C)$</td>
<td>$\lambda \bar{Y}^2 \left[ C_y^2 + \theta_9 C_x^2 (\theta_9 + 2C) \right]$</td>
</tr>
<tr>
<td>$t_{10} = \bar{y} \left( \frac{\bar{x} + \beta_2(x)}{\bar{x} + \beta_2(x)} \right)$</td>
<td>$\theta_{10} = \frac{\bar{x}}{\bar{x} + \beta_2(x)}$</td>
<td>$\lambda \bar{Y} \theta_{10} C_x^2 (\theta_{10} - C)$</td>
<td>$\lambda \bar{Y}^2 \left[ C_y^2 + \theta_{10} C_x^2 (\theta_{10} - 2C) \right]$</td>
</tr>
<tr>
<td>$t_{11} = \bar{y} \left( \frac{\bar{x} + \beta_2(x)}{\bar{x} + \beta_2(x)} \right)$</td>
<td>$\theta_{11} = \frac{\bar{x}}{\bar{x} + \beta_2(x)}$</td>
<td>$\lambda \bar{Y} \theta_{11} C_x^2 (\theta_{11} + C)$</td>
<td>$\lambda \bar{Y}^2 \left[ C_y^2 + \theta_{11} C_x^2 (\theta_{11} + 2C) \right]$</td>
</tr>
<tr>
<td>$t_{12} = \bar{y} \left( \frac{\beta_2 \bar{x} + \beta_1(x)}{\beta_2 \bar{x} + \beta_1(x)} \right)$</td>
<td>$\theta_{12} = \frac{\beta_2 \bar{x}}{\beta_2 \bar{x} + \beta_1(x)}$</td>
<td>$\lambda \bar{Y} \theta_{12} C_x^2 (\theta_{12} - C)$</td>
<td>$\lambda \bar{Y}^2 \left[ C_y^2 + \theta_{12} C_x^2 (\theta_{12} - 2C) \right]$</td>
</tr>
<tr>
<td>$t_{13} = \bar{y} \left( \frac{\bar{x} + \sigma_x}{\bar{x} + \sigma_x} \right)$</td>
<td>$\theta_{13} = \frac{\bar{x}}{\bar{x} + \sigma_x}$</td>
<td>$\lambda \bar{Y} \theta_{13} C_x^2 (\theta_{13} - C)$</td>
<td>$\lambda \bar{Y}^2 \left[ C_y^2 + \theta_{13} C_x^2 (\theta_{13} - 2C) \right]$</td>
</tr>
<tr>
<td>$t_{14} = \bar{y} \left( \frac{\bar{x} + \rho M}{\bar{x} + \rho M} \right)$</td>
<td>$\theta_{14} = \frac{\bar{x}}{\bar{x} + \rho M}$</td>
<td>$\lambda \bar{Y} \theta_{14} C_x^2 (\theta_{14} - C)$</td>
<td>$\lambda \bar{Y}^2 \left[ C_y^2 + \theta_{14} C_x^2 (\theta_{14} - 2C) \right]$</td>
</tr>
<tr>
<td>$t_{15} = \bar{y} \left( \frac{\beta_2 \bar{x} + C_x}{\beta_2 \bar{x} + C_x} \right)$</td>
<td>$\theta_{15} = \frac{\beta_2 \bar{x}}{\beta_2 \bar{x} + C_x}$</td>
<td>$\lambda \bar{Y} \theta_{15} C_x^2 (\theta_{15} - C)$</td>
<td>$\lambda \bar{Y}^2 \left[ C_y^2 + \theta_{15} C_x^2 (\theta_{15} - 2C) \right]$</td>
</tr>
<tr>
<td>$t_{16} = \bar{y} \left( \frac{C_x \bar{x} + \rho x}{C_x \bar{x} + \rho x} \right)$</td>
<td>$\theta_{16} = \frac{C_x \bar{x}}{C_x \bar{x} + \rho x}$</td>
<td>$\lambda \bar{Y} \theta_{16} C_x^2 (\theta_{16} - C)$</td>
<td>$\lambda \bar{Y}^2 \left[ C_y^2 + \theta_{16} C_x^2 (\theta_{16} - 2C) \right]$</td>
</tr>
<tr>
<td>$t_{17} = \bar{y} \left( \frac{C_x \bar{x} + M_x}{C_x \bar{x} + M_x} \right)$</td>
<td>$\theta_{17} = \frac{C_x \bar{x}}{C_x \bar{x} + M_x}$</td>
<td>$\lambda \bar{Y} \theta_{17} C_x^2 (\theta_{17} - C)$</td>
<td>$\lambda \bar{Y}^2 \left[ C_y^2 + \theta_{17} C_x^2 (\theta_{17} - 2C) \right]$</td>
</tr>
</tbody>
</table>

Continued on next page
The bias and theoretical mean square errors for the estimators $t_1$ to $t_{21}$, up to first order of approximation, can be represented in a single expressions as:

$$Bias(t_i) = \lambda \bar{Y} \left[ \theta_i C_x^2 (\theta_i \pm C) \right], \quad i = 1 \ldots 21$$  \hspace{1cm} (2.1)$$

and

$$MSE(t_i) = \lambda \bar{Y}^2 \left[ C_y^2 + \theta_i C_x^2 (\theta_i \pm 2C) \right], \quad i = 1 \ldots 21,$$  \hspace{1cm} (2.2)$$
where (+) sign is used for the product estimators and (−) sign is used for the ratio estimators. Also,

\[ C_y = \frac{S_y}{\bar{Y}}, \quad C_x = \frac{S_x}{\bar{X}}, \quad S_y^2 = \frac{1}{N-1} \sum_{i=1}^{N} (Y_i - \bar{Y})^2, \quad S_x^2 = \frac{1}{N-1} \sum_{i=1}^{N} (X_i - \bar{X})^2, \quad (2.3) \]

\[ \lambda = \frac{1-f}{n}, \quad f = \frac{n}{N} \rho_{xy} = \frac{S_{yx}}{S_y S_x}, \quad S_{xy} = \frac{1}{N-1} \sum_{i=1}^{N} (Y_i - \bar{Y})(X_i - \bar{X}), \quad C = \rho_{xy} \frac{C_y}{C_x}. \]

### 2.3 Some Ratio Estimators of the Variance

The ratio type variance estimators are used to improve the precision of the sample variance estimator when the study variable \( Y \) is positively correlated with an auxiliary variable \( X \). Isaki (1983)[20] proposed a ratio type variance estimator of the population variance \( S_y^2 \) when the population variance \( S_x^2 \) of an auxiliary variable \( X \) is known. Further improvements over the classical ratio estimator are also achieved by introducing a number of modified ratio estimators with the use of known parameters such as the coefficient of variation \( C_x \) and coefficient of kurtosis \( \beta_2(x) \). The problem of constructing efficient estimators for the population variance has been widely discussed by various authors such as Das and Tripathi (1978)[8], Wolter (1985)[85], Prasad and Singh (1990)[41], Garcia and Cebrain (1997)[10], Upadhyaya and Singh (2006)[81], Gupta and Shabbir (2008)[17], Bhushan (2012)[4], Subramani and Kumarapandiyam (2012b, 2012c)[74, 75], and Singh et al.(1988,2003)[57, 58]. Motivated by Sisoda and Dwivedi (1981)[68], Uphadhyaya and Singh (1999)[80] and Singh et al.(2004)[63], Kadilar and Cingi (2006)[24,25] suggested four types of variance estimators using known values of the coefficient of variation \( C_x \) and the coefficient of kurtosis \( \beta_2(x) \) of an auxiliary variable \( X \). Singh et al. (2011)[59] proposed the exponential ra-
ratio type estimator for the population variance with the aim to improve the efficiency of the existing ratio estimators. Also, Subramani and Kumarpandiyan (2012c)[75] suggested the modified ratio type estimators using the quartiles of the auxiliary variable. The modified ratio type estimators are biased, but have smaller mean squared errors compared to the traditional ratio type variance estimator. Following Kadilar and Cingi (2006)[25], Subramani and Kumarapandiyan (2013)[76] proposed the ratio type estimators of the population variance \( S^2_y \) using known values of the coefficient of variation \( C_x \) and the population median \( Q_2 \) of an auxiliary variable \( X \). Recently Khan and Shabbir (2013)[28] proposed another ratio type estimator of the population variance using known values of the coefficient of correlation \( \rho_{yx} \) and the population upper quartile \( Q_3 \) of an auxiliary variable. Following Singh et al. (2011)[59] and motivated by Upadhyaya et al. (2011)[82], Yadav and Kadilar (2013)[88] proposed an improved generalized ratio exponential type estimator of the population variance.

We discuss some of these estimators below.

The sample variance estimator of the population variance is defined as:

\[
\hat{S}^2_y = s^2_y, \tag{2.4}
\]

which is an unbiased estimator. Its variance is given by:

\[
V(\hat{S}^2_y) = \gamma S^4_y \left( \lambda_4 - 1 \right), \tag{2.5}
\]

where

\[
\lambda_{rs} = \frac{\mu_{rs}}{\mu^{20}_{20}}, \quad \mu_{rs} = \frac{1}{N-1} \sum_{i=1}^{N} (Y_i - \bar{Y})^r (X - \bar{X})^s, \quad \text{and} \quad \gamma = \frac{1}{n}. \tag{2.6}
\]
The classical ratio type estimator for the population variance $S_y^2$, when the population variance $S_x^2$ of an auxiliary variable $X$ is known is proposed by Isaki (1983) [20], and is given by:

$$\hat{S}_R^2 = s_y^2 \frac{S_x^2}{s_x^2},$$

(2.7)

where $s_y^2$ is the sample variance of the study variable and $s_x^2$ is the sample mean of the auxiliary variable. The bias and mean square error of this estimator, up to the first order of approximation, are given by:

$$Bias(\hat{S}_R^2) \approx \frac{1}{n} S_y^2 [ (\beta_2(x) - 1) - (\lambda_{22} - 1) ], \quad \text{and} \quad (2.8)$$

$$MSE(\hat{S}_R^2) \approx \frac{1}{n} S_y^4 \left[ (\beta_2(y) - 1) + (\beta_2(x) - 1) - 2 (\lambda_{22} - 1) \right], \quad (2.9)$$

respectively, where

$$\beta_2(y) = \frac{\mu_{04}}{\mu_{02}}, \quad \beta_2(x) = \frac{\mu_{40}}{\mu_{20}}, \quad \text{and} \quad \lambda_{22} = \frac{\mu_{22}}{\mu_{02} \mu_{20}}. \quad (2.10)$$

Singh et al. (2011) [59] proposed an exponential ratio type estimator for the population variance which is given by:

$$\hat{S}_{EXP}^2 = s_y^2 \exp \left( \frac{S_x^2 - s_x^2}{S_x^2 + s_x^2} \right).$$

(2.11)
Its bias and mean square error respectively, up to first order of approximation, are given by:

\[
\text{Bias}(\hat{S}_{\text{EXP}}^2) \approx \frac{1}{n} S_y^2 \left[ \frac{3}{8} (\beta_2(x) - 1) - \frac{1}{2} (\lambda_{22} - 1) \right], \quad \text{and} \quad (2.12)
\]

\[
MSE(\hat{S}_{\text{EXP}}^2) \approx \frac{1}{n} S_y^4 \left[ (\beta_2(y) - 1) + \left( \frac{\beta_2(x) - 1}{4} \right) - (\lambda_{22} - 1) \right]. \quad (2.13)
\]

The following table gives various modified ratio estimators of the population variance using known population parameters of an auxiliary variable. For the ease of presentation, the following notations are used:

\[
\beta = \beta_2(x) - 1 \quad \text{and} \quad \lambda = \lambda_{22} - 1. \quad (2.14)
\]

<table>
<thead>
<tr>
<th>Estimator</th>
<th>Constant $R_i$</th>
<th>Bias</th>
<th>Mean square error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{S}_1^2 = s_y^2 \left[ \frac{s_x^2 + C_x}{s_x^2 + \beta_2(x)} \right]$</td>
<td>$R_1 = \frac{s_x^2}{s_x^2 + C_x}$</td>
<td>$\frac{1}{n} S_y^2 R_1 [R_1 \beta - \lambda]$</td>
<td>$\frac{1}{n} S_y^4 [\beta + R_1^2 \beta - 2R_1 \lambda]$</td>
</tr>
<tr>
<td>$\hat{S}_2^2 = s_y^2 \left[ \frac{s_x^2 + \beta_2(x)}{s_x^2 + \beta_2(x)} \right]$</td>
<td>$R_2 = \frac{s_x^2}{s_x^2 + \beta_2(x)}$</td>
<td>$\frac{1}{n} S_y^2 R_2 [R_2 \beta - \lambda]$</td>
<td>$\frac{1}{n} S_y^4 [\beta + R_2^2 \beta - 2R_2 \lambda]$</td>
</tr>
</tbody>
</table>

Continued on next page
For convenience, the biases and mean square errors, up to first order of approximation, of the modified ratio type variance estimators $\hat{S}_i^2$ shown in Table 2 are represented in a single expressions as:

\[
\text{Bias}(S_i^2) \approx \frac{1}{n} S_y^2 R_i \left[ R_i \left( \beta_{2(x)} - 1 \right) - \left( \lambda_{22} - 1 \right) \right], \quad i = 1, \ldots, 11 \quad (2.15)
\]

\[
\text{MSE}(S_i^2) \approx \frac{1}{n} S_y^4 \left[ (\beta_{2(y)} - 1) + R_i^2 \left( \beta_{2(x)} - 1 \right) - 2 R_i (\lambda_{22} - 1) \right], \quad i = 1, \ldots, 11 \quad (2.16)
\]
For both the mean and variance estimators, we will now check how the empirical mean square errors and the approximated theoretical mean square errors compare.

2.4 Comparisons of the Theoretical and Empirical Mean Square Errors

In this section we compare the empirical and approximate theoretical mean square errors of various ratio type mean and variance estimators by carrying out a simulation study. We calculate the ratios of two mean square errors using the expressions:

\[ R(t_i) = 100 \times \frac{MSEE(t_i)}{MSET(t_i)} \quad \text{and} \quad R(\hat{S}_i^2) = 100 \times \frac{MSEE(\hat{S}_i^2)}{MSET(\hat{S}_i^2)} \] (2.17)

where \( MSEE \) is the empirical mean square error and \( MSET \) is the corresponding theoretical mean square error, correct to first order of approximation. In order to study the effect of departure from bivariate normal assumption on these comparisons, we consider three distributions - bivariate normal, bivariate Poisson and bivariate gamma for \((X, Y)\) with parameters as given in Tables 3 and 4. We generated bivariate normal distributions with mean as \( \mu = [4 \quad 6] \) and the standard deviation as \( \sigma = [2 \quad 3] \). We also used three correlation levels between \( X \) and \( Y \) as \( \rho_{yx} = 0.8; 0.2; 0.5 \). For the purpose of simulation, we generated 10,000 values from each distribution and used that as our finite population. In doing so, our means, standard deviations and the coefficient of correlation shift a little bit from the original distribution values. The same approach was used for generating bivariate Poisson and gamma distributions. Tables 5 and 6 give the ratios between the empirical and approximated theoretical mean square errors for the ratio and product type mean estimators, respectively. Table 7 does the same for the variance estimators. The population size used is \( N = 5000 \) with sample size \( n = 100, 200 \) and 500. The results are averaged over 10,000 trials. These distributions were generated using the software package R and the code
is given in Appendix A. Note that the ratios for all estimators are greater than one hundred, indicating that the first order approximations underestimate the true mean square errors. Also, the first order approximation works better when the sampling fraction is smaller. For the variance estimators, almost all ratios are close to 100, indicating that the first order approximations for the theoretical mean square errors for variance estimators are generally good.

Table 3. Population Statistics for Various Bivariate Distributions with Positive Coefficient of Correlations

<table>
<thead>
<tr>
<th></th>
<th>Normal distribution</th>
<th>Poisson distribution</th>
<th>Gamma distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pop.1</td>
<td>Pop.2</td>
<td>Pop.3</td>
</tr>
<tr>
<td>$\rho_{yx}$</td>
<td>0.7929 0.1742 0.5077</td>
<td>0.7779 0.2027 0.4931</td>
<td>0.7931 0.1965 0.4905</td>
</tr>
<tr>
<td>$\bar{Y}$</td>
<td>5.9468 6.0264 5.9642</td>
<td>5.9960 5.9538 6.0286</td>
<td>5.9882 5.9554 5.9699</td>
</tr>
<tr>
<td>$\bar{X}$</td>
<td>3.9783 4.0038 4.0078</td>
<td>3.9906 4.0156 4.0072</td>
<td>4.0268 4.0403 4.0350</td>
</tr>
<tr>
<td>$C_y$</td>
<td>0.4921 0.4921 0.5002</td>
<td>0.4083 0.4138 0.4132</td>
<td>0.4959 0.4971 0.4968</td>
</tr>
<tr>
<td>$C_x$</td>
<td>0.4991 0.4929 0.4982</td>
<td>0.5048 0.5034 0.5042</td>
<td>0.4926 0.4907 0.4915</td>
</tr>
<tr>
<td>$\beta_1(x)$</td>
<td>0.0417 0.0437 0.0514</td>
<td>0.5110 0.5046 0.5508</td>
<td>0.9461 0.9445 0.4915</td>
</tr>
<tr>
<td>$\beta_1(y)$</td>
<td>-0.0028 0.0153 -0.0141</td>
<td>0.4334 0.4007 0.4231</td>
<td>0.9998 1.0532 1.0337</td>
</tr>
<tr>
<td>$\beta_2(x)$</td>
<td>2.9266 2.9253 3.0240</td>
<td>3.2403 3.2322 3.3187</td>
<td>4.1891 4.1458 4.1813</td>
</tr>
<tr>
<td>$\beta_2(y)$</td>
<td>3.1279 3.0059 2.9742</td>
<td>3.2559 3.1523 3.0777</td>
<td>4.6017 4.9625 4.8376</td>
</tr>
</tbody>
</table>
### Table 4. Population Statistics for Various Bivariate Distributions with Negative Coefficients of Correlation

<table>
<thead>
<tr>
<th></th>
<th>Normal distribution</th>
<th>Poisson distribution</th>
<th>Gamma distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pop.1</td>
<td>Pop.2</td>
<td>Pop.3</td>
</tr>
<tr>
<td>$\rho_{yx}$</td>
<td>-0.7960</td>
<td>-0.2168</td>
<td>-0.4854</td>
</tr>
<tr>
<td>$\bar{Y}$</td>
<td>6.0463</td>
<td>5.9800</td>
<td>5.9969</td>
</tr>
<tr>
<td>$\bar{X}$</td>
<td>3.9785</td>
<td>3.9914</td>
<td>3.9938</td>
</tr>
<tr>
<td>$C_y$</td>
<td>0.4932</td>
<td>0.5011</td>
<td>0.4962</td>
</tr>
<tr>
<td>$C_x$</td>
<td>0.5053</td>
<td>0.5103</td>
<td>0.5022</td>
</tr>
<tr>
<td>$\beta_1(x)$</td>
<td>-0.0150</td>
<td>-0.0221</td>
<td>0.0155</td>
</tr>
<tr>
<td>$\beta_1(y)$</td>
<td>0.0299</td>
<td>0.0032</td>
<td>-0.0028</td>
</tr>
<tr>
<td>$\beta_2(x)$</td>
<td>2.9358</td>
<td>3.0516</td>
<td>3.0137</td>
</tr>
<tr>
<td>$\beta_2(y)$</td>
<td>2.9369</td>
<td>3.0078</td>
<td>2.9492</td>
</tr>
</tbody>
</table>

### Table 5. The Ratio of the Empirical Mean Square Errors and the Theoretical Mean Square Errors for Some Ratio Estimators of Population Mean

<table>
<thead>
<tr>
<th>Estimator</th>
<th>Distribution</th>
<th>$n$</th>
<th>$\rho_{yx} = 0.8$</th>
<th>$\rho_{yx} = 0.2$</th>
<th>$\rho_{yx} = 0.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_R$</td>
<td>normal</td>
<td>100</td>
<td>100.2321</td>
<td>103.2854</td>
<td>103.8200</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>105.0483</td>
<td>104.3566</td>
<td>105.8758</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>112.0853</td>
<td>113.84</td>
<td>110.0194</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>102.4274</td>
<td>103.6298</td>
<td>101.1414</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.5659</td>
<td>105.8989</td>
<td>106.7388</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.9086</td>
<td>111.1239</td>
<td>111.2031</td>
</tr>
</tbody>
</table>

*Continued on next page*
<table>
<thead>
<tr>
<th>Estimator</th>
<th>Distribution</th>
<th>$n$</th>
<th>$\rho_{yx} = 0.8$</th>
<th>$\rho_{yx} = 0.2$</th>
<th>$\rho_{yx} = 0.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>gamma</td>
<td>normal</td>
<td>100</td>
<td>103.4866</td>
<td>103.5309</td>
<td>103.8604</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>105.7128</td>
<td>106.5458</td>
<td>105.6162</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>111.2953</td>
<td>109.2713</td>
<td>108.9397</td>
</tr>
<tr>
<td>$t_{ER}$</td>
<td>normal</td>
<td>100</td>
<td>99.38368</td>
<td>102.4767</td>
<td>102.8346</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.3322</td>
<td>103.0252</td>
<td>106.6609</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>109.5924</td>
<td>114.084</td>
<td>112.0102</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>103.9735</td>
<td>102.7874</td>
<td>100.6577</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.9174</td>
<td>104.7845</td>
<td>106.7614</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.6656</td>
<td>112.2963</td>
<td>109.7985</td>
</tr>
<tr>
<td></td>
<td>gamma</td>
<td>100</td>
<td>102.4864</td>
<td>102.9836</td>
<td>103.0244</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>105.5767</td>
<td>105.387</td>
<td>105.7716</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.4721</td>
<td>109.686</td>
<td>109.0082</td>
</tr>
<tr>
<td>$t_1$</td>
<td>normal</td>
<td>100</td>
<td>99.91239</td>
<td>102.9766</td>
<td>103.5206</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.9526</td>
<td>103.9713</td>
<td>106.004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>111.6351</td>
<td>113.9252</td>
<td>112.0102</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>102.2853</td>
<td>103.2257</td>
<td>100.7495</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.3092</td>
<td>105.2909</td>
<td>106.9095</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.547</td>
<td>111.3049</td>
<td>110.8211</td>
</tr>
<tr>
<td></td>
<td>gamma</td>
<td>100</td>
<td>103.3101</td>
<td>103.3526</td>
<td>103.6305</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>105.6329</td>
<td>106.3251</td>
<td>105.6657</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>111.9976</td>
<td>109.2644</td>
<td>108.7708</td>
</tr>
<tr>
<td>$t_8$</td>
<td>normal</td>
<td>100</td>
<td>99.79767</td>
<td>103.1607</td>
<td>103.516</td>
</tr>
</tbody>
</table>

Continued on next page
<table>
<thead>
<tr>
<th>Estimator</th>
<th>Distribution</th>
<th>$n$</th>
<th>$\rho_{yx} = 0.8$</th>
<th>$\rho_{yx} = 0.2$</th>
<th>$\rho_{yx} = 0.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>poisson</td>
<td></td>
<td>200</td>
<td>104.9001</td>
<td>104.2067</td>
<td>106.0066</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>111.3602</td>
<td>113.8724</td>
<td>110.4547</td>
</tr>
<tr>
<td>gamma</td>
<td></td>
<td>100</td>
<td>102.3332</td>
<td>103.4441</td>
<td>100.7559</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.8723</td>
<td>105.7712</td>
<td>106.9063</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.4137</td>
<td>111.2337</td>
<td>110.8286</td>
</tr>
<tr>
<td>$t_{10}$</td>
<td>normal</td>
<td>100</td>
<td>99.49451</td>
<td>102.4014</td>
<td>102.8629</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.5105</td>
<td>103.0878</td>
<td>106.524</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>109.936</td>
<td>114.07</td>
<td>111.7853</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>103.6949</td>
<td>102.5935</td>
<td>100.4773</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.3297</td>
<td>104.7673</td>
<td>106.8626</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.4786</td>
<td>112.0259</td>
<td>109.8373</td>
</tr>
<tr>
<td></td>
<td>gamma</td>
<td>100</td>
<td>102.4247</td>
<td>102.8487</td>
<td>102.9197</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>105.6019</td>
<td>105.3316</td>
<td>105.7317</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.4758</td>
<td>109.6692</td>
<td>109.0421</td>
</tr>
<tr>
<td>$t_{12}$</td>
<td>normal</td>
<td>100</td>
<td>100.2444</td>
<td>103.2739</td>
<td>103.8076</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>105.0515</td>
<td>104.3431</td>
<td>105.8796</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>112.0971</td>
<td>113.8429</td>
<td>110.0345</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>102.3442</td>
<td>103.4837</td>
<td>100.9874</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>105.2978</td>
<td>105.2216</td>
<td>106.7997</td>
</tr>
</tbody>
</table>

*Continued on next page*
<table>
<thead>
<tr>
<th>Estimator</th>
<th>Distribution</th>
<th>$n$</th>
<th>$\rho_{yx} = 0.8$</th>
<th>$\rho_{yx} = 0.2$</th>
<th>$\rho_{yx} = 0.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.7798</td>
<td>111.2238</td>
<td>111.0676</td>
</tr>
<tr>
<td>gamma</td>
<td></td>
<td>100</td>
<td>103.3994</td>
<td>103.4378</td>
<td>103.7428</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>105.6642</td>
<td>106.4382</td>
<td>105.6381</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>111.2953</td>
<td>109.2624</td>
<td>108.8474</td>
</tr>
<tr>
<td>$t_{13}$</td>
<td>normal</td>
<td>100</td>
<td>99.57341</td>
<td>102.5288</td>
<td>103.0421</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.6762</td>
<td>103.3053</td>
<td>106.3713</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.4195</td>
<td>114.0579</td>
<td>111.4054</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>103.0156</td>
<td>102.7045</td>
<td>100.4017</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.6184</td>
<td>104.0080</td>
<td>105.0635</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.2619</td>
<td>111.7367</td>
<td>110.1086</td>
</tr>
<tr>
<td></td>
<td>gamma</td>
<td>100</td>
<td>102.8818</td>
<td>103.0759</td>
<td>103.2414</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>105.6227</td>
<td>105.8209</td>
<td>105.7647</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.7136</td>
<td>109.4011</td>
<td>108.7263</td>
</tr>
<tr>
<td>$t_{14}$</td>
<td>normal</td>
<td>100</td>
<td>99.43659</td>
<td>102.3166</td>
<td>102.7426</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.3603</td>
<td>102.9423</td>
<td>106.6093</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>109.5952</td>
<td>114.0473</td>
<td>112.0008</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>103.9902</td>
<td>102.5738</td>
<td>100.5521</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>105.2389</td>
<td>104.5556</td>
<td>106.7036</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.6473</td>
<td>112.2362</td>
<td>109.765</td>
</tr>
<tr>
<td></td>
<td>gamma</td>
<td>100</td>
<td>102.5017</td>
<td>102.8907</td>
<td>102.9737</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>105.6104</td>
<td>105.418</td>
<td>105.7475</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.5067</td>
<td>109.6157</td>
<td>108.9717</td>
</tr>
</tbody>
</table>

*Continued on next page*
<table>
<thead>
<tr>
<th>Estimator</th>
<th>Distribution</th>
<th>$n$</th>
<th>$\rho_{yx} = 0.8$</th>
<th>$\rho_{yx} = 0.2$</th>
<th>$\rho_{yx} = 0.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_{15}$</td>
<td>normal</td>
<td>100</td>
<td>100.1001</td>
<td>103.1645</td>
<td>103.7071</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>105.0131</td>
<td>104.2113</td>
<td>105.9148</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>111.9379</td>
<td>113.8713</td>
<td>110.1646</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>102.345</td>
<td>103.4841</td>
<td>100.9994</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.8734</td>
<td>105.0988</td>
<td>106.7947</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.7813</td>
<td>111.2237</td>
<td>111.0787</td>
</tr>
<tr>
<td></td>
<td>gamma</td>
<td>100</td>
<td>103.4393</td>
<td>103.4799</td>
<td>103.7964</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>105.6844</td>
<td>106.4887</td>
<td>105.6272</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>111.2492</td>
<td>109.2652</td>
<td>108.8884</td>
</tr>
<tr>
<td>$t_{16}$</td>
<td>normal</td>
<td>100</td>
<td>99.62308</td>
<td>103.0511</td>
<td>103.3062</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.7514</td>
<td>104.0687</td>
<td>106.1474</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.6907</td>
<td>113.9032</td>
<td>110.8416</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>102.7141</td>
<td>103.2923</td>
<td>100.5444</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.5655</td>
<td>105.2231</td>
<td>107.0185</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.2492</td>
<td>111.2798</td>
<td>110.5362</td>
</tr>
<tr>
<td></td>
<td>gamma</td>
<td>100</td>
<td>102.9826</td>
<td>103.3795</td>
<td>103.4657</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>105.6195</td>
<td>106.3627</td>
<td>105.7135</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.7885</td>
<td>109.2624</td>
<td>108.6981</td>
</tr>
<tr>
<td>$t_{17}$</td>
<td>normal</td>
<td>100</td>
<td>99.31717</td>
<td>102.184</td>
<td>102.4745</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.0298</td>
<td>102.763</td>
<td>106.6999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>109.0659</td>
<td>113.8599</td>
<td>112.3249</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>104.6171</td>
<td>102.6304</td>
<td>100.9547</td>
</tr>
</tbody>
</table>

Continued on next page
<table>
<thead>
<tr>
<th>Estimator</th>
<th>Distribution</th>
<th>$n$</th>
<th>$\rho_{yx} = 0.8$</th>
<th>$\rho_{yx} = 0.2$</th>
<th>$\rho_{yx} = 0.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>105.6679</td>
<td>104.7766</td>
<td>105.7763</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>111.2693</td>
<td>112.8358</td>
<td>109.7016</td>
</tr>
<tr>
<td>gamma</td>
<td></td>
<td>100</td>
<td>102.0796</td>
<td>102.5983</td>
<td>102.6325</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>105.5362</td>
<td>104.8809</td>
<td>105.5878</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.372</td>
<td>109.9758</td>
<td>109.4835</td>
</tr>
<tr>
<td>$t_{18}$</td>
<td>normal</td>
<td>100</td>
<td>99.36481</td>
<td>102.3168</td>
<td>102.4113</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>103.4991</td>
<td>102.9427</td>
<td>106.6942</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>108.7809</td>
<td>114.0474</td>
<td>112.3512</td>
</tr>
<tr>
<td>poisson</td>
<td></td>
<td>100</td>
<td>102.7024</td>
<td>103.2929</td>
<td>100.5716</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.3997</td>
<td>105.8712</td>
<td>105.0036</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.2501</td>
<td>111.2796</td>
<td>110.5812</td>
</tr>
<tr>
<td>gamma</td>
<td></td>
<td>100</td>
<td>103.2043</td>
<td>103.445</td>
<td>103.6204</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>105.6174</td>
<td>106.4471</td>
<td>105.6685</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.9938</td>
<td>109.2627</td>
<td>108.7648</td>
</tr>
<tr>
<td>$t_{19}$</td>
<td>normal</td>
<td>100</td>
<td>99.52536</td>
<td>102.1462</td>
<td>102.749</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.5819</td>
<td>102.7529</td>
<td>106.6053</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.1281</td>
<td>113.7218</td>
<td>111.997</td>
</tr>
<tr>
<td>poisson</td>
<td></td>
<td>100</td>
<td>103.3636</td>
<td>102.6733</td>
<td>100.5634</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.2235</td>
<td>104.9812</td>
<td>106.6796</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.3442</td>
<td>113.0128</td>
<td>109.7573</td>
</tr>
<tr>
<td>gamma</td>
<td></td>
<td>100</td>
<td>102.7525</td>
<td>102.4615</td>
<td>102.9356</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>105.6237</td>
<td>104.6741</td>
<td>105.7368</td>
</tr>
</tbody>
</table>

*Continued on next page*
<table>
<thead>
<tr>
<th>Estimator</th>
<th>Distribution</th>
<th>$n$</th>
<th>$\rho_{yx} = 0.8$</th>
<th>$\rho_{yx} = 0.2$</th>
<th>$\rho_{yx} = 0.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_{20}$</td>
<td>normal</td>
<td>500</td>
<td>110.6315</td>
<td>110.1277</td>
<td>109.0208</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>99.39377</td>
<td>102.1017</td>
<td>102.4807</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.2428</td>
<td>102.817</td>
<td>106.6998</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>109.3767</td>
<td>113.4286</td>
<td>112.3211</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>100.287</td>
<td>102.8040</td>
<td>100.9687</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>105.3788</td>
<td>104.8766</td>
<td>105.7411</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.8782</td>
<td>113.4029</td>
<td>109.7043</td>
</tr>
<tr>
<td></td>
<td>gamma</td>
<td>100</td>
<td>102.3522</td>
<td>102.1922</td>
<td>102.6315</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>105.5919</td>
<td>104.3254</td>
<td>105.5872</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.4494</td>
<td>110.396</td>
<td>109.4851</td>
</tr>
<tr>
<td>$t_{21}$</td>
<td>normal</td>
<td>500</td>
<td>109.0208</td>
<td>108.0308</td>
<td>107.0416</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>99.57239</td>
<td>102.5207</td>
<td>103.0416</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.6744</td>
<td>103.2917</td>
<td>106.3718</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.4135</td>
<td>110.0596</td>
<td>111.4067</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>102.9972</td>
<td>102.7115</td>
<td>100.4024</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.8978</td>
<td>104.9876</td>
<td>107.0661</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.2593</td>
<td>111.7243</td>
<td>110.1212</td>
</tr>
<tr>
<td></td>
<td>gamma</td>
<td>100</td>
<td>102.9103</td>
<td>103.092</td>
<td>103.2612</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>105.6219</td>
<td>105.8558</td>
<td>105.762</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.7338</td>
<td>109.386</td>
<td>108.7168</td>
</tr>
</tbody>
</table>
The Table 6 is given below:

Table 6. The Ratio of the Empirical Mean Square Errors and the Theoretical Mean Square Errors for Some Product Estimators of Population Mean

<table>
<thead>
<tr>
<th>Estimator</th>
<th>Distribution</th>
<th>( n )</th>
<th>( \rho_{yx} = -0.8 )</th>
<th>( \rho_{yx} = -0.2 )</th>
<th>( \rho_{yx} = -0.5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_P )</td>
<td>normal</td>
<td>100</td>
<td>100.7915</td>
<td>102.4633</td>
<td>102.924</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.3677</td>
<td>103.1064</td>
<td>105.2511</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>109.6164</td>
<td>110.7636</td>
<td>111.7535</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>500</td>
<td>102.5464</td>
<td>100.7025</td>
<td>100.4814</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.6183</td>
<td>105.0594</td>
<td>103.6489</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>111.0469</td>
<td>108.6186</td>
<td>112.588</td>
</tr>
<tr>
<td></td>
<td>gamma</td>
<td>100</td>
<td>100.4532</td>
<td>100.4111</td>
<td>102.8959</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.186</td>
<td>103.6635</td>
<td>104.9024</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.5243</td>
<td>109.9268</td>
<td>108.8254</td>
</tr>
<tr>
<td>( t_{EP} )</td>
<td>normal</td>
<td>100</td>
<td>102.6857</td>
<td>102.6083</td>
<td>102.1444</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>102.8142</td>
<td>102.7849</td>
<td>103.6092</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>109.9355</td>
<td>111.3013</td>
<td>112.0234</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>102.239</td>
<td>99.46373</td>
<td>101.3195</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>103.5144</td>
<td>104.2884</td>
<td>102.1749</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>108.0874</td>
<td>109.676</td>
<td>110.333</td>
</tr>
<tr>
<td></td>
<td>gamma</td>
<td>100</td>
<td>99.84798</td>
<td>99.98673</td>
<td>102.9191</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>103.1241</td>
<td>104.4126</td>
<td>104.2602</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>112.3393</td>
<td>110.814</td>
<td>109.1679</td>
</tr>
<tr>
<td>( t_2 )</td>
<td>normal</td>
<td>100</td>
<td>102.2507</td>
<td>102.4889</td>
<td>102.7886</td>
</tr>
</tbody>
</table>

Continued on next page
<table>
<thead>
<tr>
<th>Estimator</th>
<th>Distribution</th>
<th>$n$</th>
<th>$\rho_{yx} = -0.8$</th>
<th>$\rho_{yx} = -0.2$</th>
<th>$\rho_{yx} = -0.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>103.6927</td>
<td>102.9077</td>
<td>104.9994</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>109.5666</td>
<td>110.8968</td>
<td>111.824</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>102.6845</td>
<td>100.4653</td>
<td>100.5029</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.3465</td>
<td>104.9586</td>
<td>103.3465</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.3771</td>
<td>108.7617</td>
<td>112.1137</td>
</tr>
<tr>
<td></td>
<td>gamma</td>
<td>100</td>
<td>100.1934</td>
<td>100.2999</td>
<td>102.8969</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>103.6414</td>
<td>103.7711</td>
<td>104.7753</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.9058</td>
<td>110.0245</td>
<td>108.6992</td>
</tr>
<tr>
<td>$t_3$</td>
<td>normal</td>
<td>100</td>
<td>102.6067</td>
<td>102.4167</td>
<td>102.8792</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.1152</td>
<td>103.0319</td>
<td>105.1696</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>109.5909</td>
<td>110.8102</td>
<td>111.7785</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>102.5997</td>
<td>100.6269</td>
<td>100.4818</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.5343</td>
<td>105.0295</td>
<td>103.5587</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.8402</td>
<td>108.6615</td>
<td>112.4456</td>
</tr>
<tr>
<td></td>
<td>gamma</td>
<td>100</td>
<td>100.3809</td>
<td>100.3809</td>
<td>102.8958</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.0563</td>
<td>103.6894</td>
<td>104.8699</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.6187</td>
<td>109.9488</td>
<td>108.7841</td>
</tr>
<tr>
<td>$t_4$</td>
<td>normal</td>
<td>100</td>
<td>102.4197</td>
<td>102.6283</td>
<td>101.959</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>103.0729</td>
<td>102.9747</td>
<td>104.2773</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.0457</td>
<td>111.3666</td>
<td>112.0318</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>101.7583</td>
<td>99.24205</td>
<td>101.8122</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>103.4367</td>
<td>103.9872</td>
<td>104.89</td>
</tr>
</tbody>
</table>

*Continued on next page*
<table>
<thead>
<tr>
<th>Estimator</th>
<th>Distribution</th>
<th>$n$</th>
<th>$\rho_{yx} = -0.8$</th>
<th>$\rho_{yx} = -0.2$</th>
<th>$\rho_{yx} = -0.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>108.0828</td>
<td>110.0665</td>
<td>110.0189</td>
</tr>
<tr>
<td>gamma</td>
<td></td>
<td>100</td>
<td>100.0999</td>
<td>100.0955</td>
<td>102.9719</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.4244</td>
<td>104.8327</td>
<td>104.0842</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>112.6744</td>
<td>111.3918</td>
<td>109.9204</td>
</tr>
<tr>
<td>$t_5$</td>
<td>normal</td>
<td>100</td>
<td>102.9923</td>
<td>102.6381</td>
<td>102.4300</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>105.8345</td>
<td>104.6867</td>
<td>104.2223</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>109.6884</td>
<td>111.159</td>
<td>111.956</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>102.606</td>
<td>99.9174</td>
<td>100.7893</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>103.7097</td>
<td>104.6424</td>
<td>102.634</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>108.7927</td>
<td>109.1969</td>
<td>111.0331</td>
</tr>
<tr>
<td></td>
<td>gamma</td>
<td>100</td>
<td>99.86107</td>
<td>100.1111</td>
<td>102.9124</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>102.179</td>
<td>104.0898</td>
<td>104.4759</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>111.8073</td>
<td>110.3774</td>
<td>108.7577</td>
</tr>
<tr>
<td>$t_6$</td>
<td>normal</td>
<td>100</td>
<td>102.5952</td>
<td>102.1691</td>
<td>101.6073</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.3076</td>
<td>105.0189</td>
<td>103.2927</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.4568</td>
<td>111.1449</td>
<td>111.9344</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>102.1137</td>
<td>99.54014</td>
<td>101.2400</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>103.434</td>
<td>104.3345</td>
<td>102.1657</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>108.0895</td>
<td>109.6098</td>
<td>110.3643</td>
</tr>
<tr>
<td></td>
<td>gamma</td>
<td>100</td>
<td>99.85309</td>
<td>100.1022</td>
<td>102.9142</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>102.0757</td>
<td>104.1151</td>
<td>104.4547</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>111.8658</td>
<td>110.4081</td>
<td>108.7816</td>
</tr>
</tbody>
</table>

Continued on next page
<table>
<thead>
<tr>
<th>Estimator</th>
<th>Distribution</th>
<th>$n$</th>
<th>$\rho_{yx} = -0.8$</th>
<th>$\rho_{yx} = -0.2$</th>
<th>$\rho_{yx} = -0.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_7$</td>
<td>normal</td>
<td>100</td>
<td>100.0679</td>
<td>102.5509</td>
<td>102.745</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>103.5048</td>
<td>102.8624</td>
<td>104.9155</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>109.5654</td>
<td>110.9326</td>
<td>111.8439</td>
</tr>
<tr>
<td>poisson</td>
<td></td>
<td>100</td>
<td>102.7039</td>
<td>100.4166</td>
<td>100.511</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.2795</td>
<td>104.9354</td>
<td>103.3014</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.2112</td>
<td>108.7943</td>
<td>112.0436</td>
</tr>
<tr>
<td>gamma</td>
<td></td>
<td>100</td>
<td>100.194</td>
<td>100.3005</td>
<td>102.8969</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>103.6431</td>
<td>103.7703</td>
<td>104.7749</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.9047</td>
<td>110.0238</td>
<td>108.699</td>
</tr>
<tr>
<td>$t_9$</td>
<td>normal</td>
<td>100</td>
<td>100.479</td>
<td>100.4261</td>
<td>103.0476</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>105.6751</td>
<td>103.2183</td>
<td>105.4638</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>100.2139</td>
<td>110.6987</td>
<td>111.6723</td>
</tr>
<tr>
<td>poisson</td>
<td></td>
<td>100</td>
<td>102.1747</td>
<td>100.8053</td>
<td>100.5172</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.978</td>
<td>105.0968</td>
<td>103.9425</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>111.9261</td>
<td>108.5645</td>
<td>113.0575</td>
</tr>
<tr>
<td>gamma</td>
<td></td>
<td>100</td>
<td>100.9835</td>
<td>100.4608</td>
<td>102.8984</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.7845</td>
<td>103.625</td>
<td>105.0288</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>110.0227</td>
<td>109.8964</td>
<td>109.0511</td>
</tr>
<tr>
<td>$t_{11}$</td>
<td>normal</td>
<td>100</td>
<td>102.7542</td>
<td>102.6559</td>
<td>102.2548</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>104.7896</td>
<td>104.7075</td>
<td>103.9014</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>109.8039</td>
<td>111.2546</td>
<td>111.9976</td>
</tr>
<tr>
<td>poisson</td>
<td></td>
<td>100</td>
<td>102.3278</td>
<td>99.61978</td>
<td>101.1517</td>
</tr>
</tbody>
</table>

*Continued on next page*
Table 7. The Ratio of the Empirical Mean Square Errors and the Theoretical Mean Square Errors for Some Ratio Estimators of Population Variance

<table>
<thead>
<tr>
<th>Estimator</th>
<th>Distribution</th>
<th>n</th>
<th>$\rho_{yx} = -0.8$</th>
<th>$\rho_{yx} = -0.2$</th>
<th>$\rho_{yx} = -0.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>103.4949</td>
<td>104.4081</td>
<td>102.2342</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>108.2487</td>
<td>109.5119</td>
<td>110.4586</td>
</tr>
<tr>
<td>gamma</td>
<td></td>
<td>100</td>
<td>99.88689</td>
<td>100.0447</td>
<td>102.9379</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>101.1153</td>
<td>104.4351</td>
<td>104.251</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>112.373</td>
<td>110.8222</td>
<td>109.186</td>
</tr>
<tr>
<td>$\hat{S}_R^2$</td>
<td>normal</td>
<td>100</td>
<td>102.6034</td>
<td>100.1753</td>
<td>100.2699</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>100.0565</td>
<td>101.8364</td>
<td>103.8158</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>101.3344</td>
<td>102.3967</td>
<td>101.6947</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>100.8989</td>
<td>100.32592</td>
<td>100.3245</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>102.8938</td>
<td>100.6621</td>
<td>99.9786</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>101.7800</td>
<td>101.2609</td>
<td>100.3021</td>
</tr>
<tr>
<td></td>
<td>gamma</td>
<td>100</td>
<td>100.6676</td>
<td>101.1580</td>
<td>100.5213</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>100.7265</td>
<td>102.1337</td>
<td>102.2242</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>101.9045</td>
<td>100.9870</td>
<td>100.6419</td>
</tr>
<tr>
<td>$\hat{S}_{EXP}^2$</td>
<td>normal</td>
<td>100</td>
<td>99.14575</td>
<td>99.1708</td>
<td>101.5117</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>99.86147</td>
<td>99.66083</td>
<td>99.85453</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>102.5893</td>
<td>102.1652</td>
<td>100.7433</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>101.7435</td>
<td>99.7546</td>
<td>99.4442</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>100.6136</td>
<td>100.7263</td>
<td>99.9876</td>
</tr>
</tbody>
</table>

Continued on next page
<table>
<thead>
<tr>
<th>Estimator</th>
<th>Distribution</th>
<th>n</th>
<th>$\rho_{yx} = 0.8$</th>
<th>$\rho_{yx} = 0.2$</th>
<th>$\rho_{yx} = 0.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{S}_1^2$</td>
<td>gamma</td>
<td>500</td>
<td>100.2186</td>
<td>100.3465</td>
<td>100.1730</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>101.8146</td>
<td>100.5106</td>
<td>99.8945</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>102.2507</td>
<td>100.9564</td>
<td>101.0135</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>100.5558</td>
<td>101.9936</td>
<td>100.8295</td>
</tr>
<tr>
<td></td>
<td>normal</td>
<td>100</td>
<td>102.2496</td>
<td>100.9403</td>
<td>103.8919</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>99.45776</td>
<td>101.5443</td>
<td>102.4047</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>102.3915</td>
<td>102.1696</td>
<td>101.2589</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>100.93423</td>
<td>101.3237</td>
<td>99.89328</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>102.0616</td>
<td>100.1365</td>
<td>99.3841</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>101.3153</td>
<td>99.8032</td>
<td>100.0044</td>
</tr>
<tr>
<td></td>
<td>gamma</td>
<td>100</td>
<td>100.6576</td>
<td>100.6597</td>
<td>102.1390</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>100.9911</td>
<td>101.7138</td>
<td>101.8544</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>102.4915</td>
<td>101.9453</td>
<td>102.3454</td>
</tr>
<tr>
<td>$\hat{S}_2^2$</td>
<td>normal</td>
<td>100</td>
<td>98.75806</td>
<td>102.4983</td>
<td>100.7022</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>99.46771</td>
<td>102.3254</td>
<td>99.73651</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>102.5981</td>
<td>101.9765</td>
<td>100.5811</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>100.5542</td>
<td>99.6985</td>
<td>99.2566</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>100.5009</td>
<td>99.5748</td>
<td>101.5866</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>100.2759</td>
<td>100.0909</td>
<td>100.3321</td>
</tr>
<tr>
<td></td>
<td>gamma</td>
<td>100</td>
<td>101.6496</td>
<td>100.2611</td>
<td>99.7620</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>101.1869</td>
<td>100.7350</td>
<td>100.8760</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>100.3941</td>
<td>101.9003</td>
<td>100.1419</td>
</tr>
</tbody>
</table>

*Continued on next page*
<table>
<thead>
<tr>
<th>Estimator</th>
<th>Distribution</th>
<th>$n$</th>
<th>$\rho_{yx} = 0.8$</th>
<th>$\rho_{yx} = 0.2$</th>
<th>$\rho_{yx} = 0.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{S}_3^2$</td>
<td>normal</td>
<td>100</td>
<td>103.6515</td>
<td>101.2764</td>
<td>105.2261</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>99.7874</td>
<td>102.3254</td>
<td>103.2793</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>102.3452</td>
<td>102.3012</td>
<td>101.5259</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>100.9856</td>
<td>100.5016</td>
<td>100.1651</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>102.5866</td>
<td>102.0570</td>
<td>99.3830</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>101.6146</td>
<td>99.8975</td>
<td>100.1987</td>
</tr>
<tr>
<td></td>
<td>gamma</td>
<td>100</td>
<td>100.6440</td>
<td>101.0134</td>
<td>100.4092</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>100.7819</td>
<td>102.0199</td>
<td>102.1244</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>102.8929</td>
<td>100.9120</td>
<td>101.4923</td>
</tr>
<tr>
<td>$\hat{S}_4^2$</td>
<td>normal</td>
<td>100</td>
<td>98.15762</td>
<td>101.8818</td>
<td>100.2661</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>99.92439</td>
<td>98.8897</td>
<td>98.9731</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>102.4372</td>
<td>102.0219</td>
<td>100.4837</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>100.5581</td>
<td>100.8977</td>
<td>99.8810</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>99.88426</td>
<td>99.9987</td>
<td>99.6844</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>99.86211</td>
<td>100.2740</td>
<td>100.3467</td>
</tr>
<tr>
<td></td>
<td>gamma</td>
<td>100</td>
<td>100.02922</td>
<td>100.5105</td>
<td>99.8400</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>102.3324</td>
<td>100.5051</td>
<td>100.5786</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>100.1122</td>
<td>102.4766</td>
<td>100.4888</td>
</tr>
<tr>
<td>$\hat{S}_5^2$</td>
<td>normal</td>
<td>100</td>
<td>98.326</td>
<td>101.9862</td>
<td>100.5991</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>99.2388</td>
<td>98.93484</td>
<td>99.526</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>102.4762</td>
<td>101.8357</td>
<td>100.6721</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>100.9719</td>
<td>99.6923</td>
<td>99.3741</td>
</tr>
</tbody>
</table>

*Continued on next page*
<table>
<thead>
<tr>
<th>Estimator</th>
<th>Distribution</th>
<th>$n$</th>
<th>$\rho_{yx} = 0.8$</th>
<th>$\rho_{yx} = 0.2$</th>
<th>$\rho_{yx} = 0.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>100.3009</td>
<td>100.8454</td>
<td>100.5001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>99.86105</td>
<td>101.9002</td>
<td>99.5100</td>
</tr>
<tr>
<td></td>
<td>gamma</td>
<td>100</td>
<td>101.3022</td>
<td>100.1843</td>
<td>99.8754</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>101.9050</td>
<td>100.9605</td>
<td>101.8705</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>100.9568</td>
<td>101.5943</td>
<td>100.9341</td>
</tr>
<tr>
<td>$\hat{S}_6^2$</td>
<td>normal</td>
<td>100</td>
<td>109.2415</td>
<td>91.42963</td>
<td>93.9322</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>111.478</td>
<td>88.76051</td>
<td>92.56915</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>113.919</td>
<td>91.64913</td>
<td>94.16197</td>
</tr>
<tr>
<td></td>
<td>poisson</td>
<td>100</td>
<td>111.0124</td>
<td>89.1428</td>
<td>90.0183</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>110.2576</td>
<td>88.9592</td>
<td>93.4996</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>112.6517</td>
<td>89.8431</td>
<td>92.9386</td>
</tr>
<tr>
<td></td>
<td>gamma</td>
<td>100</td>
<td>112.0999</td>
<td>90.1924</td>
<td>93.8665</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>111.4981</td>
<td>90.2828</td>
<td>94.5559</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>100.8711</td>
<td>91.9146</td>
<td>92.5532</td>
</tr>
</tbody>
</table>
2.5 Robustness Against Departure From the Bivariate Normal Assumption

We have considered bivariate normal distributions with the three positive and three negative coefficients of correlation $\rho_{yx}$ whose statistics are given in Tables 3 and 4. Also, we have considered the bivariate Poisson and bivariate Gamma distributions for similar coefficients of correlation $\rho_{yx}$. It was shown that the departure from the bivariate normal assumption does not produce any serious issue if the sample size is large. Tables 5 and 6 are for the ratio and product type mean estimators respectively, and Table 7 is for the ratio type variance estimators. We can see that corresponding ratios for the six bivariate distributions are very similar. Thus we can say that these estimators are robust with respect to the assumption of bivariate normality.
CHAPTER III
THE NEW GENERALIZED MIXTURE ESTIMATORS OF THE MEAN

3.1 Introduction

In this chapter we propose new generalized mixture estimators of the population mean of the study variable by utilizing an auxiliary variable. The aim is to get a more efficient estimator than the existing mixture estimators. As was mentioned before, many modification have been done on the ratio, product, and regression estimators of the population mean of the study variable $Y$ using an auxiliary variable $X$ to improve efficiency of these estimators. Bahl & Tuteja [2] introduced the exponential ratio and product type estimators which we will also use in the proposed estimators. Combining modified ratio type estimators and the exponential ratio type estimator, Singh et al. (2009)[65] suggested a generalized ratio type estimator. The special cases of this estimator are exponential ratio type, exponential product type and also Bedi (1996)[3] transformed estimators. Also, many authors have suggested several transformed ratio-type estimators for estimating the finite population mean by utilizing auxiliary information. Khoshnevisan et al. (2007)[29] proposed a general class of estimators that includes several modified ratio type estimators. Shabbir and Gupta (2010)[52] proposed a regression ratio type exponential estimator by combining Rao’s (1991)[42] and Bedi’s (1996)[3] estimators. Following these works, Grover & Kaur (2011)[12] introduced a regression exponential type estimator. Subramani (2013)[77] proposed a generalized modified ratio estimator for estimation of finite population mean. The ordinary ratio estimator, the linear regression estimator and the existing
modified ratio estimators are special cases of that estimator. Also, more recently Grover & Kaur (2014)[13] proposed a generalized class of ratio type exponential estimators by combining Rao’s (1991)[42] and Singh’s et al. (2009)[65] generalized ratio type exponential estimator. Our approach is as follows.

Let \( U = \{U_1, \ldots, U_N\} \) be a finite population of size \( N \) and let \((y_i, x_i)\) be the value of the study variable \( Y \) and the auxiliary variable \( X \) on \( i \)th unit \( U_i, i = 1, \ldots N \). Let \( \bar{Y} \) and \( \bar{X} \) be population means of the study variable \( Y \) and the auxiliary variable \( X \) respectively. We assume that the population mean \( \bar{X} \) and the population variance \( S_x^2 \) of the auxiliary variable are known. Let \( S_y^2 \) be the population variance of the study variable \( Y \). Let the correlation coefficient between the study variable and the auxiliary variable be \( \rho_{yx} \). Also, let \( C_y = \frac{S_y}{Y} \) and \( C_x = \frac{S_x}{X} \) be the coefficients of variation of the study variable \( Y \) and the auxiliary variable \( X \), and \( C_{yx} = \frac{S_{yx}}{YX} \) be the coefficient of covariance between \( Y \) and \( X \) with \( S_{yx} = \frac{1}{N-1} \sum_{i=1}^{N} (y_i - \bar{Y})(x_i - \bar{X}) \).

To obtain the asymptotic properties of the estimators, we define the following error terms, as in Sukathme and Sukathme (1970)[78]:

\[
e_y = \frac{\bar{y} - \bar{Y}}{Y} \quad \text{and} \quad e_x = \frac{\bar{x} - \bar{X}}{X}, \tag{3.1}
\]

for which the following holds true:

\[
E(e_y) = E(e_x) = 0, \quad E(e_y^2) = \lambda C_y^2, \quad E(e_x^2) = \lambda C_x^2, \quad E(e_y e_x) = \lambda C_{yx} = \lambda \rho_{yx} C_y C_x, \quad \text{where} \quad \lambda = \frac{1-f}{n} \quad \text{and} \quad f = \frac{n}{N}.
\tag{3.2}
\]
We will use these expressions to derive the bias and the mean square error of the proposed estimators, up to the first order of approximation.

3.2 Some Existing Generalized Estimators

There are many generalized estimators of the population mean of the study variable utilizing an auxiliary variable. We discuss a few of these generalized estimators first.

3.2.1 Rao (1991) Regression Estimator

Rao (1991)[42] introduced the generalized regression type estimator to improve efficiency of the ordinary regression estimator. The estimator is given by:

\[ \hat{\mu}_{R,Reg} = k_1 \bar{y} + k_2 \left( \bar{X} - \bar{x} \right), \]  

(3.3)

where \( k_1 \) and \( k_2 \) are suitably chosen constants. The minimum mean square error of this estimator, up to the first order of approximation, with optimum values of \( k_1 \) and \( k_2 \) i.e.,

\[ k_{1(\text{opt})} = \frac{1}{1 + \lambda \left( 1 - \rho_{yx}^2 \right) C^2_y}, \]  

(3.4)

\[ k_{2(\text{opt})} = k_{1(\text{opt})} \frac{\bar{Y}}{\bar{X}} \frac{\rho_{yx} C_y}{C_x}, \]  

(3.5)

is given by:

\[ MSE_{\text{min}}(\hat{\mu}_{R,Reg}) \approx \bar{Y}^2 \left[ 1 - \frac{1}{1 + \lambda \left( 1 - \rho_{yx}^2 \right) C^2_y} \right]. \]  

(3.6)
3.2.2 Singh et al. (2008) Estimator

Following Bahl & Tuteja (1991)[2], Singh et al. (2008)[64] proposed a ratio product type exponential estimator given by:

\[
\hat{\mu}_S = \bar{y} \left[ \alpha \exp \left( \frac{X - \bar{x}}{X + \bar{x}} \right) + (1 - \alpha) \exp \left( \frac{\bar{x} - X}{X + \bar{x}} \right) \right],
\]

(3.7)

where \( \alpha \) is suitably chosen constant. The minimum mean square error, up to the first order of approximation, at optimum value of \( \alpha \), i.e.,

\[
\alpha_{(\text{opt})} = \frac{1}{2} + \frac{\rho_{yx} C_y}{C_x},
\]

(3.8)

is given by:

\[
MSE_{\text{min}}(\hat{\mu}_S) \approx \lambda \bar{Y}^2 \left( 1 - \rho_{yx}^2 \right) C_y^2 = MSE(\hat{\mu}_{\text{Reg}}).
\]

(3.9)

3.2.3 Grover & Kaur (2011) Estimator

Following Rao (1991) [42] and Bahl & Tuteja [2], Grover & Kaur (2011) [12] suggested a regression exponential type estimator given by:

\[
\hat{\mu}_{\text{GK}} = \left[ l_1 \bar{y} + l_2 (\bar{X} - \bar{x}) \right] \exp \left( \frac{\bar{X} - \bar{x}}{X + \bar{x}} \right),
\]

(3.10)

where \( l_1 \) and \( l_2 \) are suitably chosen constants.
The minimum mean square error of this estimator, at the optimum values of $l_1$ and $l_2$, i.e.,

$$l_{1(\text{opt})} = \frac{-1 + \frac{1}{8} \lambda C_x^2}{-1 + \lambda (1 - \rho_{yx}^2) C_y^2},$$  \hspace{1cm} (3.11)

$$l_{2(\text{opt})} = \frac{\bar{Y}}{X} \left[ \frac{1}{2} - l_{1(\text{opt})} \left( 1 - \rho_{yx} \frac{C_y}{C_x} \right) \right],$$  \hspace{1cm} (3.12)

is given by:

$$MSE_{\text{min}}(\hat{\mu}_{GK}) \approx \frac{\lambda \bar{Y}^2 \left[ \lambda C_x^4 - 16 \left( \rho_{yx}^2 - 1 \right) (-4 + \lambda C_x^2) C_y^2 \right]}{64 \left[-1 + \lambda \left( \rho_{yx}^2 - 1 \right) C_y^2 \right]}.$$  \hspace{1cm} (3.13)

It turns out that

$$MSE_{\text{min}}(\hat{\mu}_{GK}) \approx MSE(\hat{\mu}_{\text{Reg}}) - \frac{\lambda^2 \bar{Y}^2 \left[ C_x^2 + 8 \left( 1 - \rho_{yx}^2 \right) C_y^2 \right]^2}{64 \left[ 1 + \lambda \left( \rho_{yx}^2 - 1 \right) C_y^2 \right]}.$$  \hspace{1cm} (3.14)

We can see that Grover & Kaur (2011)[12] estimator is more efficient than the linear regression estimator $\hat{\mu}_{\text{Reg}}$ if

$$1 + \lambda \left( \rho_{yx}^2 - 1 \right) C_y^2 > 0, \text{ or } \rho_{yx}^2 > 1 - \frac{1}{\lambda C_y^2}.$$  \hspace{1cm} (3.15)

This condition is very likely to hold true since $\left( 1 - \frac{1}{\lambda C_y^2} \right)$ is typically small. For example, if $N = 5000, n = 200$ and $C_y = 1.5$, this expression equals -150. Hence (3.15) will hold true for all correlation values. Since the linear regression estimator is always better than the sample mean, ratio, product, and exponential estimators,
we can say that Grover & Kaur (2011) estimator is also always better than these estimators.

3.3 The Proposed Generalized Mixture Estimator I

In this section we propose a new generalized mixture estimator by combining the ratio, product, regression, and exponential ratio type estimators. The estimator is given as:

\[
\hat{\mu}_{GM} = \left\{ d_1 \bar{y} \left[ \frac{1}{2} \left( \frac{\bar{X}}{\bar{x}} + \frac{\bar{x}}{\bar{X}} \right) \right]^\alpha + d_2 (\bar{X} - \bar{x}) \right\} \exp \left( \frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right) \tag{3.16}
\]

where \(d_i (i = 1, 2)\) and \(\alpha\) are suitably chosen constants. We will consider two values of \(\alpha\) (\(\alpha = 1\) and \(\alpha = 2\)). Using error terms (3.1), this generalized mixture estimator can be written as:

\[
\hat{\mu}_{GM} = \left[ d_1 \bar{Y} (1 + e_y) \frac{1}{2\alpha} \left[ (1 + e_x)^{-1} + 1 + e_x \right]^\alpha - \bar{X} d_2 e_x \right] \exp \left[ \left( -\frac{e_x}{2} \right) \left( 1 + \frac{e_x}{2} \right)^{-1} \right]. \tag{3.17}
\]

Using first order approximation, this can be written as:

\[
\hat{\mu}_{GM} \approx \left[ d_1 \bar{Y} (1 + e_y) \frac{1}{2\alpha} \left( 2 + e_x^2 \right)^\alpha - d_2 \bar{X} e_x \right] \left( 1 - \frac{1}{2} e_x + \frac{3}{8} e_x^2 \right)
\approx \left[ d_1 \bar{Y} (1 + e_y) \left( 1 + \alpha \frac{e_x^2}{2} \right) - d_2 \bar{X} e_x \right] \left( 1 - \frac{1}{2} e_x + \frac{3}{8} e_x^2 \right)
\approx d_1 \bar{Y} + \frac{\alpha}{2} d_1 \bar{Y} e_x^2 + d_1 \bar{Y} e_y - d_2 \bar{X} e_x - \frac{1}{2} d_1 \bar{Y} e_x - \frac{1}{2} d_1 \bar{Y} e_y e_x + \frac{1}{2} d_2 \bar{X} e_x^2 + \frac{3}{8} d_1 \bar{Y} e_x^2.
\]
Thus it follows

\[ \hat{\mu}_{GM} - \bar{Y} \approx (d_1 - 1) \bar{Y} + d_1 \bar{Y} \left( e_y - \frac{1}{2} e_x - \frac{1}{2} e_x e_y + A e_x^2 \right) - d_2 \bar{X} \left( e_y - \frac{1}{2} e_x^2 \right), \quad (3.18) \]

where

\[ A = \frac{\alpha}{2} + \frac{3}{8}. \quad (3.19) \]

By taking expectation of (3.18), the bias of the proposed generalized mixture estimator, up to the first order of approximation, is given by:

\[ \text{Bias}(\hat{\mu}_{GM}) \approx (d_1 - 1) \bar{Y} + \lambda d_1 \bar{Y} \left( AC_x^2 - \frac{1}{2} \rho_{yx} C_y C_x \right) + \frac{d_2}{2} \lambda \bar{X} C_x^2. \quad (3.20) \]

By squaring equation (3.18) and keeping terms only up to the first order of approximation, we have:

\[ (\hat{\mu}_{GM} - \bar{Y})^2 \approx (d_1 - 1)^2 + d_1^2 \bar{Y}^2 \left( e_y - \frac{1}{2} e_x - \frac{1}{2} e_y e_x + A e_x^2 \right) + d_2^2 \bar{X}^2 \left( e_y - \frac{1}{2} e_x^2 \right)^2 + 2d_1 (d_1 - 1) \bar{Y}^2 \left( e_y - \frac{1}{2} e_x - \frac{1}{2} e_y e_x + A e_x^2 \right) - 2d_2 (d_1 - 1) \bar{X} \bar{Y} \left( e_y - \frac{1}{2} e_x^2 \right) 
- 2d_1 d_2 \bar{Y} \overline{X} \left( e_y - \frac{1}{2} e_x - \frac{1}{2} e_y e_x + A e_x^2 \right) \left( e_x - \frac{1}{2} e_x^2 \right) 
= (d_1 - 1)^2 \bar{Y}^2 + d_1^2 \bar{Y}^2 \left( e_y e_x + \frac{1}{4} e_x^2 \right) + d_2^2 \bar{X}^2 e_x^2 
+ 2d_2^2 \bar{Y}^2 \left( e_y - \frac{1}{2} e_x - \frac{1}{2} e_y e_x + A e_x^2 \right) - 2d_1 \bar{Y}^2 \left( e_y - \frac{1}{2} e_y - \frac{1}{2} e_x e_y + A e_x^2 \right) 
- 2d_1 d_2 \bar{Y} \left( e_y e_x - e_x^2 + e_y + d_2 \bar{X} \left( 2e_x - e_x^2 \right). \quad (3.21) \]
This can be further simplified to

\[(\hat{\mu}_{GM} - \bar{Y})^2 \approx (d_1 - 1)^2 + d_2^2 \bar{Y}^2 \left[ 2e_y - e_x - 2e_y e_x + e_y^2 + \left( 2A + \frac{1}{4} \right) e_x^2 \right] \]
\[+ d_2^2 \bar{X}^2 e_x^2 - d_1 \bar{Y}^2 \left( 2e_y - e_x - e_y e_x + 2Ae_x^2 \right) \]
\[- 2d_1d_2 \bar{X} \bar{Y} \left( e_x + e_y e_x - e_x^2 \right) + d_2 \bar{X} \bar{Y} \left( 2e_x - e_x^2 \right). \] (3.22)

Taking expectation of (3.22), the mean square error of the proposed estimator, up to the first order of approximation, is given as:

\[MSE(\hat{\mu}_{GM}) \approx (d_1 - 1)^2 \bar{Y}^2 + \lambda d_1^2 \bar{Y}^2 \left[ C_y^2 - 2\rho_{yx} C_y C_x + \left( 2A + \frac{1}{4} \right) C_x^2 \right] \]
\[+ \lambda d_2^2 \bar{X}^2 C_x^2 - \lambda d_1 \bar{Y}^2 \left( 2AC_x^2 - \rho_{yx} C_y C_x \right) \]
\[- 2\lambda d_1d_2 \bar{X} \bar{Y} \left( \rho_{yx} C_y C_x - C_x^2 \right) - \lambda d_2 \bar{X} \bar{Y} C_x^2 \] (3.23)

Taking partial derivatives with respect to \(d_1\) and \(d_2\), we have:

\[\frac{\partial MSE(\hat{\mu}_{GM})}{\partial d_1} = 2(d_1 - 1) \bar{Y}^2 + 2\lambda d_1 \bar{Y}^2 \left[ C_y^2 - 2\rho_{yx} C_y C_x + \left( 2A + \frac{1}{4} \right) C_x^2 \right] \]
\[- \lambda \bar{Y}^2 \left( 2AC_x^2 - \rho_{yx} C_y C_x \right) - 2\lambda d_2 \bar{X} \bar{Y} \left( \rho_{yx} C_y C_x - C_x^2 \right) \],

\[\frac{\partial MSE(\hat{\mu}_{GM})}{\partial d_2} = 2\lambda d_2 \bar{X}^2 C_x^2 - 2\lambda d_1 \bar{X} \bar{Y} \left( \rho_{yx} C_y C_x - C_x^2 \right) - \lambda \bar{X} \bar{Y} C_x^2. \] (3.24)
Setting these derivatives equal to zero, the optimum values of \( d_1 \) and \( d_2 \) are given as:

\[
d_{1(\text{opt})} = \frac{1 + \lambda \left( A - \frac{1}{2} \right) C_x^2}{1 + \lambda \left[ \left( 2A - \frac{3}{4} \right) C_x^2 + (1 - \rho_{yx}) C_y^2 \right]}, \quad \text{and} \quad (3.25)
\]

\[
d_{2(\text{opt})} = \frac{\bar{Y}}{X} \left[ \frac{1}{2} - d_{1(\text{opt})} \left( 1 - \rho_{yx} \frac{C_y}{C_x} \right) \right]. \quad (3.26)
\]

Substituting the optimum value of \( d_2 \) in (3.23) we get:

\[
MSE_{\text{min}}(\hat{\mu}_{GM}) \approx (d_1 - 1)^2 \bar{Y}^2 + \lambda d_1^2 \bar{Y}^2 \left[ \left( 2A + \frac{1}{4} \right) C_x^2 - 2\rho_{yx} C_x C_y + C_y^2 \right]
\]

\[
+ \lambda \bar{X}^2 C_x^2 \left[ \frac{1}{2} - d_{1(\text{opt})} \left( 1 - \rho_{yx} \frac{C_y}{C_x} \right) \right]^2 - \lambda d_1 \bar{Y}^2 \left[ 2AC_x^2 - \rho_{yx} C_x C_y \right]
\]

\[
- 2\lambda \bar{X} \bar{Y} \frac{d_1}{X} \left[ \frac{1}{2} - d_1 \left( 1 - \rho_{yx} \frac{C_y}{C_x} \right) \right] \left( \rho_{yx} C_x C_y - C_x^2 \right)
\]

\[
- \lambda \bar{X} \bar{Y} C_x^2 \frac{d_1}{X} \left[ \frac{1}{2} - d_1 \left( 1 - \rho_{yx} \frac{C_y}{C_x} \right) \right]
\]

\[
= d_1^2 \bar{Y}^2 - 2d_1 \bar{Y}^2 + \bar{Y}^2 + \lambda d_1^2 \left[ \left( 2A + \frac{1}{4} \right) C_x^2 - 2\rho_{yx} C_x C_y + C_y^2 \right]
\]

\[
+ \lambda \bar{Y}^2 C_x^2 \left[ \frac{1}{4} - d_1 \left( 1 - \rho_{yx} \frac{C_y}{C_x} \right) \right] + d_1^2 \left( 1 - \rho_{yx} \frac{C_y}{C_x} \right)^2
\]

\[
= d_1^2 \bar{Y}^2 \left\{ 1 + \lambda \left[ \left( 2A - \frac{3}{4} \right) C_x^2 + (1 - \rho_{yx}) C_y^2 \right] \right\}
\]

\[
- 2d_1 \bar{Y}^2 \left[ 1 + \lambda \left( A - \frac{1}{2} \right) C_x^2 \right] + \bar{Y}^2 - \frac{1}{4} \lambda \bar{Y}^2 C_x^2. \quad (3.27)
\]
Substituting the optimum value of $d_1$ in (3.27), the minimum mean square error is given by:

$$MSE_{\text{min}}(\hat{\mu}_{GM}) \approx \bar{Y}^2 \left\{ \left( 1 - \frac{1}{4} \lambda C_x^2 \right) - \frac{\left[ 1 + \lambda \left( A - \frac{1}{2} \right) C_x^2 \right]^2}{1 + \lambda \left[ (2A - \frac{3}{4}) C_x^2 + (1 - \rho_{xy}^2) C_y^2 \right]} \right\}.$$  

For $\alpha = 1$ this generalized mixture estimator becomes

$$\hat{\mu}_{GM1} = \left\{ d_1 \bar{y} \left[ \frac{1}{2} \left( \frac{\bar{X}}{\bar{x}} + \frac{\bar{x}}{\bar{X}} \right) \right] + d_2 (\bar{X} - \bar{x}) \right\} \exp \left( \frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right). \quad (3.28)$$

The optimum values of $d_1$ and $d_2$, for this estimator, are given by:

$$d_{1(\text{opt})} = \frac{1 + \frac{3}{8} \lambda C_x^2}{1 + \lambda \left[ C_x^2 + (1 - \rho_{xy}^2) C_y^2 \right]} , \quad \text{and} \quad (3.29)$$

$$d_{2(\text{opt})} = \frac{\bar{Y}}{\bar{X}} \left[ \frac{1}{2} - d_{1(\text{opt})} \left( 1 - \rho_{yx} \frac{C_y}{C_x} \right) \right]. \quad (3.30)$$

The minimum mean squared error, up to the first order of approximation, is given by:

$$MSE_{\text{min}}(\hat{\mu}_{GM1}) \approx \bar{Y}^2 \left\{ \left( 1 - \frac{1}{4} \lambda C_x^2 \right) - \frac{\left[ 1 + \lambda \left( A - \frac{1}{2} \right) C_x^2 \right]^2}{1 + \lambda \left[ C_x^2 + (1 - \rho_{xy}^2) C_y^2 \right]} \right\}. \quad (3.31)$$

When $\alpha = 2$, the generalized mixture estimator is given by:

$$\hat{\mu}_{GM2} = \left\{ d_1 \bar{y} \left[ \frac{1}{2} \left( \frac{\bar{X}}{\bar{x}} + \frac{\bar{x}}{\bar{X}} \right) \right]^2 + d_2 (\bar{X} - \bar{x}) \right\} \exp \left( \frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right). \quad (3.32)$$
The optimum values of \(d_1\) and \(d_2\) are given by:

\[
d_1^{(opt)} = \frac{1 + \frac{7}{8} \lambda C_x^2}{1 + \lambda [2C_x^2 + (1 - \rho_{yx}) C_y^2]}, \quad \text{and} \tag{3.33}
\]

\[
d_2^{(opt)} = \frac{Y}{X} \left[ \frac{1}{2} - d_1^{(opt)} \left( 1 - \rho_{yx} \frac{C_y}{C_x} \right) \right]. \tag{3.34}
\]

The minimum mean square of this estimator, up to the first order of approximation, is given by:

\[
MSE_{\min}(\hat{\mu}_{GM_2}) \approx Y^2 \left\{ \left( 1 - \frac{1}{4} \lambda C_x^2 \right) - \frac{\left[ 1 + \frac{7}{8} \lambda C_x^2 \right]^2}{1 + \lambda [2C_x^2 + (1 - \rho_{yx}) C_y^2]} \right\}. \tag{3.35}
\]

Since in most cases, the correlation coefficient between the study variable \(Y\) and the auxiliary variable \(X\) is positive, we also propose a generalized mixture estimator without the product term in the next section.

### 3.4 The Proposed Generalized Mixture Estimator II

By combining the ratio, regression, and the exponential estimators, we propose a second generalized mixture estimator without the product term:

\[
\hat{\mu}_{GMR} = \left[ k_1 \bar{y} \left( \frac{\bar{X}}{\bar{x}} \right)^\alpha + k_2 (\bar{X} - \bar{x}) \right] \exp \left( \frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right) \tag{3.36}
\]

where \(k_i\) (\(i = 1, 2\)) and \(\alpha\) are suitably chosen constants. We will consider two values of \(\alpha\) (\(\alpha = 1\) and \(\alpha = 2\)).
Rewriting $\hat{\mu}_{GMR}$ in terms $e_y$ and $e_x$, the estimator (3.36) can be written as:

$$\hat{\mu}_{GMR} = k_1 \bar{Y} \left[ (1 + e_y) (1 + e_x)^{-\alpha} - k_2 \bar{X} e_x \right] \exp \left[ \left(-\frac{e_x}{2}\right) \left(1 + \frac{e_x}{2}\right)^{-1} \right]$$  \hspace{1cm} (3.37)

Up to the first order of approximation, this estimator can be expressed as:

$$\hat{\mu}_{GMR} \approx k_1 \bar{Y} (e_y + 1) \left(1 - \alpha e_x + \frac{1}{2} \alpha (\alpha + 1) e_x^2 \right) - k_2 \bar{X} e_x$$  

$$= k_1 \bar{Y} - k_1 \alpha \bar{Y} e_x + \frac{1}{2} k_1 \bar{Y} \alpha (\alpha + 1) e_x^2 + k_1 \bar{Y} e_x - k_1 \alpha e_x e_x - k_2 \bar{X} e_x$$

$$\quad - \frac{1}{2} k_1 \bar{Y} e_x + \frac{1}{2} k_1 \alpha \bar{Y} e_x^2 - \frac{1}{2} k_1 \bar{Y} e_x e_x + \frac{1}{2} k_2 \bar{X} e_x^2 + \frac{3}{8} k_1 \bar{Y} e_x^2$$

$$= k_1 \bar{Y} - k_1 \alpha \bar{Y} e_x + \frac{1}{2} k_1 \alpha \bar{Y} e_x^2 - \frac{1}{2} k_1 \bar{Y} e_x e_x + \frac{1}{2} k_2 \bar{X} e_x^2 + \frac{3}{8} k_1 \bar{Y} e_x^2$$

$$\quad + k_1 \bar{Y} e_x - k_1 \alpha \bar{Y} e_x e_x - k_2 \bar{X} e_x$$

$$= \bar{Y} + (k_1 - 1) \bar{Y} + k_1 \bar{Y} \left\{ e_y - \left( \alpha + \frac{1}{2} \right) e_x - \left( \alpha + \frac{1}{2} \right) e_x e_y \right\}$$

$$\quad + \left\{ \frac{1}{2} \alpha (\alpha + 2) + \frac{3}{8} \right\} e_x^2 - k_2 \bar{X} e_x$$

$$= \bar{Y} + \left( k_1 - 1 \right) \bar{Y} + k_1 \bar{Y} \left\{ e_y - A e_x - A e_x e_y + B e_x^2 \right\} - k_2 \bar{X} \left( e_x - \frac{1}{2} e_x^2 \right)$$  \hspace{1cm} (3.38)

Using the substitutions

$$A = \alpha + \frac{1}{2} \quad \text{and} \quad B = \frac{1}{2} \alpha (\alpha + 2) + \frac{3}{8},$$  \hspace{1cm} (3.39)

we get

$$\hat{\mu}_{GMR} - \bar{Y} = (k_1 - 1) \bar{Y} + k_1 \bar{Y} \left( e_y - A e_x - A e_x e_y + B e_x^2 \right) - k_2 \bar{X} \left( e_x - \frac{1}{2} e_x^2 \right)$$  \hspace{1cm} (3.40)
By taking expectation on both sides of (3.40), the bias of the proposed estimator, up to the first order of approximation, is given by:

\[
\text{Bias}(\hat{\mu}_{GMR}) = (k_1 - 1) \bar{Y} + \lambda k_1 \bar{Y} \left( BC_x^2 - A\rho_{yx} C_x C_y \right) + \frac{1}{2} k_2 \lambda \bar{X} C_x^2. \tag{3.41}
\]

By squaring Equation (3.40), and keeping terms only up to the first order of approximation, we get:

\[
(\hat{\mu}_{GMR} - \bar{Y})^2 \approx (k_1 - 1)^2 \bar{Y}^2 + k_1^2 \bar{Y}^2 \left( e_y - Ae_x - Ae_x e_y + Be_x^2 \right)^2
+ k_2^2 \bar{X}^2 \left( e_x - \frac{1}{2} e_x^2 \right)^2 + 2k_1 (k_1 - 1) \bar{Y}^2 \left( e_y - Ae_x - Ae_x e_y + Be_x^2 \right)
- 2k_1 k_2 \bar{X} \bar{Y} \left( e_x - \frac{1}{2} e_x^2 \right) \left( e_y - Ae_x - Ae_x e_y + Be_x^2 \right)
- 2k_2 (k_1 - 1) \bar{X} \bar{Y} \left( e_x - \frac{1}{2} e_x^2 \right)
= (k_1 - 1)^2 \bar{Y}^2 + k_1^2 \bar{Y}^2 \left( e_y - 2Ae_x e_x + A^2 e_x^2 \right) + k_2^2 \bar{X}^2 e_x^2
+ 2k_1^3 \bar{Y}^2 \left( e_y - Ae_x - Ae_x e_y + Be_x^2 \right) - 2k_1 \bar{Y}^2 \left( e_y - Ae_x - Ae_x e_y + Be_x^2 \right)
- 2k_1 k_2 \bar{X} \bar{Y} \left( e_x e_y - Ae_x^2 \right) - 2k_1 k_2 \bar{X} \bar{Y} \left( e_x - \frac{1}{2} e_x^2 \right) + 2k_2 \bar{X} \bar{Y} \left( e_x - \frac{1}{2} e_x^2 \right)
= (k_1 - 1)^2 \bar{Y}^2 + k_1^2 \bar{Y}^2 \left[ 2e_y - 2Ae_x - 4Ae_x e_y + e_y^2 + (A^2 + 2B) e_x^2 \right]
+ k_2^2 \bar{X}^2 e_x^2 - 2k_1 \bar{Y}^2 \left( e_y - Ae_x - Ae_y e_x + Be_x^2 \right)
- 2k_1 k_2 \bar{X} \bar{Y} \left[ e_x + e_x e_y - \left( A + \frac{1}{2} \right) e_x^2 \right] + k_2 \bar{X} \bar{Y} \left( 2e_x - e_x^2 \right). \tag{3.42}
\]

Using the substitution:

\[
C = A^2 + 2B, \tag{3.43}
\]
we get

\[ (\hat{\mu}_{GMR} - \bar{Y})^2 \approx (k_1 - 1)^2 \bar{Y}^2 + k_1^2 \bar{Y}^2 \left( 2e_y - 2Ae_x - 4Ae_x e_y + e_y^2 + C e_x^2 \right) \]
\[ + k_2^2 \bar{X}^2 e_x^2 - 2k_2 \bar{Y} \left( e_y - Ae_x - Ae_y + Be_x^2 \right) \]
\[ - 2k_1 k_2 \bar{X} \left[ e_x + e_x e_y - \left( A + \frac{1}{2} \right) e_x^2 \right] + k_2 \bar{X} \bar{Y} \left( 2e_x - e_x^2 \right) \]

(3.44)

By taking expectation on both sides, the mean square error of the proposed estimator, up to the first order of approximation, is given by:

\[ \text{MSE}(\hat{\mu}_{GMR}) \approx (k_1 - 1)^2 \bar{Y}^2 + \lambda k_1^2 \bar{Y}^2 \left( CC_x^2 - 4A \rho_{xy} C_x C_y + C_y^2 \right) \]
\[ + \lambda k_2^2 \bar{X}^2 C_x^2 - \lambda k_1 \bar{Y} \left( 2BC_x^2 - 2A \rho_{xy} C_x C_y \right) \]
\[ - 2\lambda k_1 k_2 \bar{X} \bar{Y} \left[ \rho_{xy} C_x C_y - \left( A + \frac{1}{2} \right) C_x^2 \right] - \lambda k_2 \bar{X} \bar{Y} C_x^2. \]

(3.45)

Partially differentiating (3.45) with respect to \( k_1 \) and \( k_2 \), we get:

\[ \frac{\partial \text{MSE}(\hat{\mu}_{GMR})}{\partial k_1} = 2(k_1 - 1)Y^2 + 2\lambda k_1 Y^2 \left( CC_x^2 - 4A \rho_{xy} C_x C_y + C_y^2 \right) \]
\[ - \lambda Y^2 \left( 2BC_x^2 - 2A \rho_{xy} C_x C_y \right) - 2\lambda k_2 \bar{X} \bar{Y} \left[ \rho_{xy} C_x C_y - \left( A + \frac{1}{2} \right) C_x^2 \right] \]

(3.46)

\[ \frac{\partial \text{MSE}(\hat{\mu}_{GMR})}{\partial k_2} = 2\lambda k_2 \bar{X}^2 C_x^2 - 2\lambda k_1 \bar{X} \bar{Y} \left[ \rho_{xy} C_x C_y - \left( A + \frac{1}{2} \right) C_x^2 \right] - \lambda \bar{X} \bar{Y} C_x^2. \]
Setting the first derivatives equal to zero, the optimum values of \( k_1 \) and \( k_2 \) are given by:

\[
k_{1(\text{opt})} = \frac{1 + \lambda \left[ \left( B - \frac{1}{2} A - \frac{1}{4} \right) C_x^2 + \left( \frac{1}{2} - A \right) \rho_{yx} C_x C_y \right]}{1 + \lambda \left\{ \left[ C - \left( A + \frac{1}{2} \right)^2 \right] C_x^2 + \left( 1 - 2A \right) \rho_{yx} C_x C_y + \left( 1 - \rho_{yx}^2 \right) C_y^2 \right\}},
\]

(3.47)

and

\[
k_{2(\text{opt})} = \frac{\bar{Y}}{X} \left\{ \frac{1}{2} - k_{1(\text{opt})} \left[ \left( A + \frac{1}{2} \right) - \rho_{yx} \frac{C_y}{C_x} \right] \right\}.
\]

(3.48)

By substituting the optimum value of \( k_2 \) in the expression (3.45), we get:

\[
MSE_{\text{min}}(\hat{\mu}_{\text{GMR}}) \approx k_1 \bar{Y}^2 - 2k_1 \bar{Y}^2 + \bar{Y}^2 + \lambda k_1 \bar{Y}^2 \left( C_x^2 C_y^2 - 4A \rho_{yx} C_y C_x + C_y^2 \right)
\]

\[
+ \lambda \bar{X} \left[ \frac{1}{2} \left( A + \frac{1}{2} \right) - \rho_{yx} \frac{C_y}{C_x} \right] \bar{Y}^2 - 2\lambda \bar{X} \bar{Y} \left[ \frac{1}{2} - \frac{1}{2} \left( A + \frac{1}{2} \right) - \rho_{yx} \frac{C_y}{C_x} \right] \left[ \rho_{yx} C_y C_x - \left( A + \frac{1}{2} \right) C_x^2 \right]
\]

\[
- \lambda \bar{X} \frac{\bar{Y}}{X} \left[ \frac{1}{2} - \frac{1}{2} \left( A + \frac{1}{2} \right) - \rho_{yx} \frac{C_y}{C_x} \right] \bar{Y}^2 \left\{ 1 + \lambda \left\{ \left[ C - \left( A + \frac{1}{2} \right) \right]^2 \right\} + \left( 1 - 2A \right) \rho_{yx} C_x C_y + \left( 1 - \rho_{yx}^2 \right) C_x^2 \right\}
\]

(3.49)

Then substituting the optimum value of \( k_1 \) in (3.49), the minimum mean square error of the proposed estimator, up to the first order of approximation , is given by:
\[ MSE_{\min}(\hat{\mu}_{GMR}) \approx \bar{Y}^2 \left\{ \left( 1 - \frac{1}{4} \lambda C_x^2 \right) - \frac{\left[ 1 + \lambda \left( \frac{7}{8} C_x^2 - \rho_{xy} C_x C_y \right) \right]^2}{1 + \lambda \left[ 2 C_x^2 - 2 \rho_{xy} C_x C_y + (1 - \rho_{xy}^2) C_y^2 \right]} \right\} \] (3.50)

For \( \alpha = 1 \) this generalized mixture estimator is given by:

\[ \hat{\mu}_{GMR1} = \left[ k_1 \bar{y} \left( \frac{\bar{X}}{\bar{x}} \right) + k_2 (\bar{X} - \bar{x}) \right] \exp \left( \frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right). \] (3.51)

The optimum values of \( k_1 \) and \( k_2 \) are given by:

\[ k_{1(\text{opt})} = \frac{1 + \lambda \left[ \frac{7}{8} C_x^2 - \rho_{xy} C_x C_y \right]}{1 + \lambda \left[ 2 C_x^2 - 2 \rho_{xy} C_x C_y + (1 - \rho_{xy}^2) C_y^2 \right]}, \quad \text{and} \] (3.52)

\[ k_{2(\text{opt})} = \frac{\bar{Y}}{\bar{X}} \left[ \frac{1}{2} - k_{1(\text{opt})} \left( 2 - \rho_{xy} \frac{C_y}{C_x} \right) \right]. \] (3.53)

and the minimum mean square error, up to the first order of approximation, is given by:

\[ MSE_{\min}(\hat{\mu}_{GMR1}) = \bar{Y}^2 \left\{ \left( 1 - \frac{1}{4} \lambda C_x^2 \right) - \frac{\left[ 1 + \lambda \left( \frac{7}{8} C_x^2 - \rho_{xy} C_x C_y \right) \right]^2}{1 + \lambda \left[ 2 C_x^2 - 2 \rho_{xy} C_x C_y + (1 - \rho_{xy}^2) C_y^2 \right]} \right\}. \]

When \( \alpha = 2 \), this generalized mixture estimator is given by:

\[ \hat{\mu}_{GMR2} = \left[ k_1 \bar{y} \left( \frac{\bar{X}}{\bar{x}} \right)^2 + k_2 (\bar{X} - \bar{x}) \right] \exp \left( \frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right) \] (3.54)
The optimum values for \( k_1 \) and \( k_2 \) are given as:

\[
k_{1(\text{opt})} = \frac{1 + \lambda \left[ \frac{23}{8} C_x^2 - 2 \rho_{yx} C_x C_y \right]}{1 + \lambda \left[ 6C_x^2 - 4 \rho_{yx} C_x C_y + (1 - \rho_{yx}^2)C_y^2 \right]}, \quad \text{and} \quad (3.55)
\]

\[
k_{2(\text{opt})} = \frac{\bar{Y}}{\bar{X}} \left[ \frac{1}{2} - k_{1(\text{opt})} \left( 3 - \rho_{yx}, \frac{C_y}{C_x} \right) \right], \quad (3.56)
\]

and the minimum mean square error is given by:

\[
MSE_{\text{min}}(\hat{\mu}_{GMR2}) = \bar{Y}^2 \left\{ \left( 1 - \frac{1}{4} \lambda C_x^2 \right) - \frac{\left[ 1 + \lambda \left( \frac{23}{8} C_x^2 - 2 \rho_{yx} C_x C_y \right) \right]^2}{1 + \lambda \left[ 6C_x^2 - 4 \rho_{yx} C_x C_y + (1 - \rho_{yx}^2) C_y^2 \right]} \right\}.
\]

In the next section we will derive the conditions under which our estimators perform better than ordinary sample mean, ratio, exponential and regression type estimators.
3.5 Efficiency Comparisons

In this section efficiency of the second proposed estimator is compared with some of the commonly used estimators. We did not use the first proposed estimator because it was constantly less efficient than the second proposed estimator in our simulations and numerical examples presented in next sections. Conditions under which the proposed estimator is more efficient are given below:

\[
\text{MSE}(\hat{\mu}_{GMR}) < \text{MSE}(\bar{y}) \quad \text{if}
\]

\[
\lambda C_y^2 - \left\{ \left( 1 - \frac{1}{4} \lambda C_x^2 \right) - \frac{\left\{ 1 + \lambda \left[ (B - \frac{1}{2} A - \frac{1}{4}) C_x^2 + (-A + \frac{1}{2}) C_{xy} \right] \right\}^2}{1 + \lambda \left[ (2B - A - \frac{1}{4}) C_x^2 + (-2A + 1) \rho_{yx} C_x C_y + (1 - \rho_{xy}^2) C_y^2 \right]} \right\} > 0
\]

\[
\text{MSE}(\hat{\mu}_{GMR}) < \text{MSE}(\hat{\mu}_{R}) \quad \text{if}
\]

\[
\lambda (C_x - \rho_{xy} C_y)^2 + \lambda (1 - \rho_{xy}^2) C_y^2 - \left\{ \left( 1 - \frac{1}{4} \lambda C_x^2 \right) - \frac{\left\{ 1 + \lambda \left[ (B - \frac{1}{2} A - \frac{1}{4}) C_x^2 + (-A + \frac{1}{2}) C_{xy} \right] \right\}^2}{1 + \lambda \left[ (2B - A - \frac{1}{4}) C_x^2 + (-2A + 1) \rho_{yx} C_x C_y + (1 - \rho_{xy}^2) C_y^2 \right]} \right\} > 0
\]

\[
\text{MSE}(\hat{\mu}_{GMR}) < \text{MSE}(\hat{\mu}_{Reg}) \quad \text{if}
\]
the following inequality holds:

\[
\lambda \bar{Y}^2 C_y^2 (1 - \rho_{yx}^2) - \left\{ \left(1 - \frac{1}{4} \lambda C_x^2 \right) \right. \\
\left. - \frac{1 + \lambda \left[(B - \frac{1}{2}A - \frac{1}{4}) C_x^2 + (-A + \frac{1}{2}) C_{xy} \right]^2}{1 + \lambda \left\{ (2B - A - \frac{1}{4}) C_x^2 + (-2A + 1) \rho_{yx} C_x C_y + (1 - \rho_{xy}^2) C_y^2 \right\}} \right\} > 0
\]

\[
MSE(\hat{\mu}_{GMR}) < MSE(\hat{\mu}_{ER}) \quad \text{if}
\]

\[
\lambda \left( \frac{1}{2} C_x - \rho_{xy} C_y \right)^2 + \lambda \left(1 - \rho_{xy}^2\right) C_y^2 - \left\{ \left(1 - \frac{1}{4} \lambda C_x^2 \right) \right. \\
\left. - \frac{1 + \lambda \left[(B - \frac{1}{2}A - \frac{1}{4}) C_x^2 + (-A + \frac{1}{2}) C_{xy} \right]^2}{1 + \lambda \left\{ (2B - A - \frac{1}{4}) C_x^2 + (-2A + 1) \rho_{yx} C_x C_y + (1 - \rho_{xy}^2) C_y^2 \right\}} \right\} > 0
\]

Numerical examples and simulation results show that these conditions are generally true, and hence the proposed estimator may be preferred over the existing estimators when these conditions hold true.

### 3.6 Numerical Examples

In this section we compare the performances of different estimators with the proposed generalized mixture estimators using some real data sets whose summary statistics are in Table 8. Table 9 shows the Theoretical Percent Relative Efficiencies of the estimators as compared to the ordinary sample mean which is calculated from the following expression:

\[
PRET(\hat{\mu}_i) = \frac{MSET(\bar{y})}{MSET(\hat{\mu}_i)} \times 100
\]
where \( i = R; Reg; ER; S, RP; R, Reg; GK \), and \( MSET \) is the theoretical mean square error. From Table 9 we can confirm that all the percent relative efficiencies are greater than 100 indicating that all estimators are better than the sample mean estimator. The proposed generalized estimators are more efficient than other estimators given in Table 9.

Table 8. Summary Statistics for the Real Populations Used in Comparing \( \hat{\mu}_{GM} \) and \( \hat{\mu}_{GMR} \) with other Mean Estimators

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Population 1</th>
<th>Population 2</th>
<th>Population 3</th>
<th>Population 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N )</td>
<td>70</td>
<td>34</td>
<td>256</td>
<td>204</td>
</tr>
<tr>
<td>( n )</td>
<td>25</td>
<td>20</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>( \rho_{yx} )</td>
<td>0.7293</td>
<td>0.4491</td>
<td>0.887</td>
<td>0.71</td>
</tr>
<tr>
<td>( Y )</td>
<td>96.7</td>
<td>856.4118</td>
<td>56.47</td>
<td>966</td>
</tr>
<tr>
<td>( X )</td>
<td>175.2671</td>
<td>208.8824</td>
<td>44.45</td>
<td>26441</td>
</tr>
<tr>
<td>( C_y )</td>
<td>0.6254</td>
<td>0.8561</td>
<td>1.42</td>
<td>2.4739</td>
</tr>
<tr>
<td>( C_x )</td>
<td>0.8037</td>
<td>0.7205</td>
<td>1.40</td>
<td>1.7171</td>
</tr>
<tr>
<td>( f )</td>
<td>0.3571</td>
<td>0.5882</td>
<td>0.3906</td>
<td>0.2451</td>
</tr>
</tbody>
</table>

(1) Population 1 [Source: Singh and Chaudhary (1986), pp.108]

(2) Population 2 [Source: Singh and Chaudhary(1986), pp. 177]

(3) Population 3 [Source: Cochran (1977), pp. 196]

(4) Population 4 [Source: Kadilar & Cingi (2005)]
Table 9. The Theoretical Percent Relative Efficiency for Various Mean Estimators

<table>
<thead>
<tr>
<th>Estimators</th>
<th>Population 1</th>
<th>Population 2</th>
<th>Population 3</th>
<th>Population 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\mu}_Y$</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>$\hat{\mu}_R$</td>
<td>128.6948</td>
<td>105.0011</td>
<td>448.3998</td>
<td>201.5302</td>
</tr>
<tr>
<td>$\hat{\mu}_{Reg}$</td>
<td>213.6198</td>
<td>125.2647</td>
<td>468.975</td>
<td>201.6536</td>
</tr>
<tr>
<td>$\hat{\mu}_{ER}$</td>
<td>210.2398</td>
<td>125.1392</td>
<td>271.3702</td>
<td>159.1883</td>
</tr>
<tr>
<td>$\hat{\mu}_{S,RP}$</td>
<td>213.6198</td>
<td>125.2647</td>
<td>468.975</td>
<td>201.6536</td>
</tr>
<tr>
<td>$\hat{\mu}_{R,Reg}$</td>
<td>214.6255</td>
<td>126.7737</td>
<td>470.2037</td>
<td>210.9342</td>
</tr>
<tr>
<td>$\hat{\mu}_{GK}$</td>
<td>215.7188</td>
<td>127.1322</td>
<td>472.0147</td>
<td>213.4533</td>
</tr>
<tr>
<td>$\hat{\mu}_{GM1}$</td>
<td>224.2375</td>
<td>128.9346</td>
<td>487.6321</td>
<td>226.2338</td>
</tr>
<tr>
<td>$\hat{\mu}_{GM2}$</td>
<td>240.5612</td>
<td>131.3791</td>
<td>518.698</td>
<td>226.2338</td>
</tr>
<tr>
<td>$\hat{\mu}_{GMR1}$</td>
<td>222.6746</td>
<td>128.7937</td>
<td>474.0536</td>
<td>215.9437</td>
</tr>
<tr>
<td>$\hat{\mu}_{GMR2}$</td>
<td>298.3243</td>
<td>137.8524</td>
<td>536.5996</td>
<td>242.7579</td>
</tr>
</tbody>
</table>
3.7 Conclusion

In this chapter, we have proposed two generalized mixture estimators in simple random sampling without replacement by using information on an auxiliary variable. The proposed estimators are a mixture of some of the commonly known estimators. We have derived the minimum mean square errors up to the first order of approximation. Also we discussed two special cases $\alpha = 1$ and $\alpha = 2$. Numerical examples with real data show that both of the proposed estimators are more efficient than other estimators considered here. Also the estimators for $\alpha = 2$ perform better than the estimators with $\alpha = 1$.

We may note that at a theoretical level, one may be tempted to optimize $\alpha$. Our goal though was to have a general family of estimators where many of the existing estimators become special cases of the proposed estimator with specific choice of $\alpha$. For example, with $\alpha = 0$ our generalized mixture estimator II becomes combination of the regression and exponential ratio type estimators. For $\alpha = 1$, it involves the ratio term also. For $\alpha = -1$, it involves the product term.

In the next chapter we will use the proposed estimator $\hat{\mu}_{GMR}$ in the situation when the study variable is sensitive in nature and can not be observed directly, and a non-sensitive auxiliary variable is available.
CHAPTER IV

IMPROVED GENERALIZED MIXTURE ESTIMATORS WITH RRT MODELS

4.1 Introduction

Randomized response technique (RRT) is used to estimate the mean of a sensitive variable $Y$ when direct observation is not possible. In this chapter, our focus is on estimation of the mean of a sensitive variable $Y$ which cannot be observed directly using a non-sensitive auxiliary variable $X$. Sousa et al. (2010) [69] introduced the ratio type estimator and Gupta et al. (2012) [19] proposed the ordinary regression and a generalized regression-cum-ratio estimators based on RRT models. Following Bahl & Tuteja (1991) [2], Koyucu et al. (2014) [32] proposed the generalized exponential ratio type estimators to improve the efficiency of the mean estimator based on RRT models. In this chapter we propose an ordinary exponential ratio type estimator and a generalized mixture estimators where RRT estimators of the population mean of the study variable is further improved by using information about an auxiliary variable. We will use the following notations.

Let $Y$ be the study sensitive variable which cannot be observed directly. Let $X$ be a non sensitive auxiliary variable which has a strong positively correlation with $Y$, and let $S$ be a scrambling variable. Assume that $S$ is independent of $Y$ and $X$. Also, assume that the population mean and the population variance of the scrambling variable are known and given as $\mu_s = 0$ and $\sigma_S^2$. The population mean and the population variance of the non-sensitive auxiliary variable are known and given as $\bar{X}$ and $S_x^2$. The population mean and the population variance of the study variable are
unknown and given as $\bar{Y}$ and $S^2_Y$. Let a random sample of size $n$ be drawn without replacement from a finite population $U = (U_1, U_2, \ldots, U_N)$. For $i$th unit, let $y_i$ and $x_i$, respectively be the values of the study variable $Y$ and the auxiliary variable $X$. The respondent is asked to report a scrambled response for $Y$ given by:

$$Z = Y + S,$$  \hspace{1cm} (4.1)

but is asked to provide the true response for $X$. Note that from (4.1) $\bar{Z} = \bar{Y}$ and,

$$C^2_z = C^2_y + \frac{\sigma^2_s}{\bar{Y}^2}$$  \hspace{1cm} (4.2)

where $C_z$ and $C_y$ are the coefficients of the variation of the reported variable $Z$ and the study variable $Y$, respectively. We will use the same error terms as in Sukhatme and Sukhatme (1970), given as:

$$e_z = \frac{\bar{z} - Z}{Z} \text{ and } e_x = \frac{\bar{x} - \bar{X}}{\bar{X}},$$  \hspace{1cm} (4.3)

for which the following holds true:

$$E(e_z) = E(e_x) = 0, \quad E(e_z^2) = \lambda C_z^2, \quad E(e_x^2) = \lambda C_x^2$$

$$E(e_z e_x) = \lambda C_{zz} = \lambda \rho_{zx} C_z C_x, \text{ where } \lambda = \left(\frac{1}{n} - \frac{1}{N}\right).$$  \hspace{1cm} (4.4)

### 4.2 The Proposed Generalized Mixture Estimator in RRT

Following Bahl & Tuteja (1991) we propose the ordinary exponential ratio type estimator for estimating the population mean of the sensitive characteristic $Y$ when
non sensitive auxiliary variable $X$ is used. This estimator is given by:

$$
\hat{\mu}_{ER} = \bar{z} \exp \left( \frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right),
$$

(4.5)

where $\bar{z}$ and $\bar{x}$ are the sample means of the reported responses and an auxiliary variable, respectively. Using error terms (4.3), this estimator can be written as:

$$
\hat{\mu}_{ER} = \bar{Z} \left( 1 + e_z \right) \exp \left[ -e_x (e_x + 2)^{-1} \right].
$$

(4.6)

Up to the first order of approximation, this estimator can be written as:

$$
\hat{\mu}_{ER} - \bar{Z} \approx \bar{Z} \left( e_z - \frac{1}{2} e_x - \frac{1}{2} e_z e_x + \frac{3}{8} e_z^2 \right).
$$

(4.7)

Recognizing that $\bar{Z} = \bar{Y}$ in (4.7), the bias of this estimator up to the first order of approximation, is given as:

$$
\text{Bias}(\hat{\mu}_{ER}) \approx \lambda \bar{Y} \left( \frac{3}{8} C_x^2 - \frac{1}{2} \rho_{zx} C_z C_x \right)
$$

(4.8)

By squaring equation (4.7) and using first order of approximation, we get:

$$
(\hat{\mu}_{ER} - \bar{Z})^2 \approx \bar{Z}^2 \left( e_z^2 - e_z e_x + \frac{1}{4} e_x^2 \right).
$$

(4.9)
Taking expectation, the mean square error of the proposed estimator, up to the first order of approximation, is given by:

\[
MSE(\hat{\mu}_{ER}) \approx \lambda Y^2 \frac{1}{4} \left(4C_2^2 - 4\rho_{xz}C_zC_x + C_x^2 \right) .
\] (4.10)

The conditions under which the proposed estimator is more efficient than the ordinary sample mean and RRT ratio estimators are given below:

a) \[ MSE(\hat{\mu}_{ER}) < MSE(\hat{\mu}_Y) \quad \text{if} \quad \rho_{zx} > \frac{1C_x}{4C_z} , \] (4.11)

b) \[ MSE(\hat{\mu}_{ER}) < MSE(\hat{\mu}_R) \quad \text{if} \quad \rho_{zx} < \frac{3C_x}{4C_z} . \] (4.12)

By combining the regression, ratio and exponential estimators we further generalized the estimator in (4.5) and propose a generalized mixture estimator given by:

\[
\hat{\mu}_{GRR} = \left[ d_1 \tilde{z} \left( \frac{\bar{X}}{\bar{x}} \right) + d_2 \left( \bar{X} - \bar{x} \right) \right] \exp \left( \frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right) ,
\] (4.13)

where \( d_i (i = 1, 2) \) and \( \alpha \) are suitably chosen constants. We will consider two values for \( \alpha \) (\( \alpha = 1 \) and \( \alpha = 2 \)). Using error terms (4.3), this generalized mixture estimator \( \hat{\mu}_{GRR} \) can be written as:

\[
\hat{\mu}_{GRR} = \left[ d_1 \tilde{Z} \left( 1 + e_z \right) \left( 1 + e_x \right)^{-\alpha} - d_2 \bar{X}e_x \right] \exp \left[ -\frac{e_x}{2} \left( 1 + \frac{e_x}{2} \right)^{-1} \right] .
\] (4.14)
Up to the first order of approximation, this estimator can be written as:

$$\hat{\mu}_{\text{GRR}} \approx \left[ d_1 \bar{Z} (e_z + 1) \left( 1 - \alpha e_x + \frac{1}{2} \alpha (\alpha + 1) e_x^2 \right) - d_2 \bar{X} e_x \right] \left( 1 - \frac{1}{2} e_x + \frac{3}{8} e_x^2 \right)$$

Using the substitutions

$$A = \alpha + \frac{1}{2} \quad \text{and} \quad B = \frac{1}{2} \alpha (\alpha + 2) + \frac{3}{8},$$

we get

$$\hat{\mu}_{\text{GRR}} - \bar{Z} \approx (d_1 - 1) \bar{Z} + d_1 \bar{Z} \left( e_z - A e_x - A e_x e_x + B e_x^2 \right) - d_2 \bar{X} \left( e_x - \frac{1}{2} e_x^2 \right). \quad (4.17)$$

By taking expectation on both sides of (4.17) and recognizing that $\bar{Z} = \bar{Y}$, the bias of the proposed estimator, up to the first order of approximation, is given by:

$$\text{Bias}(\hat{\mu}_{\text{GRR}}) \approx (d_1 - 1)\bar{Y} + \lambda d_1 \bar{Y} \left( B C_x^2 - A p_{zz} C_{zz} C_x \right) + \lambda d_2 \bar{X} \frac{1}{2} C_x^2. \quad (4.18)$$
By squaring Equation (4.17), and keeping terms only up to the first order of approximation, we get:

\[
(\hat{\mu}_{GRR} - Z)^2 \approx (d_1 - 1)^2 Z^2 + d_1^2 Z^2 (e_x - Ae_x - Ae_x e_x + Be_x^2)^2 + d_2^2 X^2 \left( e_x - \frac{1}{2} e_x^2 \right)^2 \\
+ 2d_1 (d_1 - 1) Z^2 (e_x - Ae_x - Ae_x e_x + Be_x^2) \\
- 2d_1 d_2 \bar{X} \bar{Z} \left( e_x - \frac{1}{2} e_x^2 \right) (e_x - Ae_x - Ae_x e_x + Be_x^2) - 2d_2 (d_1 - 1) \bar{X} \bar{Z} \left( e_x - \frac{1}{2} e_x^2 \right)
\]

\[
= (d_1 - 1)^2 Z^2 + d_1^2 Z^2 \left( e_x^2 - 2Ae_x e_x + A^2 e_x^2 \right) + d_2^2 X^2 e_x^2 \\
+ 2d_1 \bar{Y}^2 (e_x - Ae_x - Ae_x e_x + Be_x^2) - 2d_1 \bar{Z}^2 (e_x - Ae_x - Ae_x e_x + Be_x^2) \\
- 2d_1 d_2 \bar{X} \bar{Z} (e_x e_x - Ae_x e_x) - 2d_1 d_2 \bar{X} \bar{Z} \left( e_x - \frac{1}{2} e_x^2 \right) + 2d_2 \bar{X} \bar{Z} \left( e_x - \frac{1}{2} e_x^2 \right).
\]

This can be further simplified to:

\[
(\hat{\mu}_{GRR} - \bar{Z})^2 \approx (d_1 - 1)^2 \bar{Z}^2 + d_1^2 \bar{Z}^2 \left[ 2e_x^2 - 2Ae_x e_x - 4Ae_x e_x + e_x^2 + (A^2 + 2B) e_x^2 \right] \\
+ d_2^2 X^2 e_x^2 - 2d_1 \bar{Z}^2 \left[ e_x e_x - Ae_x e_x + Be_x^2 \right] \\
- 2d_1 d_2 \bar{X} \bar{Z} \left[ e_x + e_x e_x - \left( A + \frac{1}{2} \right) e_x^2 \right] + 2d_2 \bar{X} \bar{Z} \left( e_x - \frac{1}{2} e_x^2 \right) .
\]

By taking expectation of (4.20) and recognizing that $\bar{Z} = \bar{Y}$, the mean square error of the proposed estimator, up to the first order of approximation, is given by:

\[
MSE(\hat{\mu}_{GRR}) \approx (d_1 - 1)^2 \bar{Y}^2 + \lambda d_1^2 \bar{Y}^2 \left[ (A^2 + 2B) C_x^2 - 4A \rho_{xx} C_x C_x + C_x^2 \right] \\
+ \lambda d_2^2 \bar{X}^2 C_x^2 - 2\lambda d_1 \bar{Y}^2 \left( BC_x^2 - A \rho_{xx} C_x C_x \right) \\
- 2\lambda d_1 d_2 \bar{X} \bar{Y} \left[ \rho_{xx} C_x C_x - \left( A + \frac{1}{2} \right) C_x^2 \right] - \lambda d_2 \bar{X} \bar{Y} C_x^2 . \tag{4.21}
\]
Partially differentiating (4.21) with respect to $d_1$ and $d_2$, we get:

$$\frac{\partial \text{MSE}(\hat{\mu}_{GRR})}{\partial d_1} = 2 (d_1 - 1) \bar{Y}^2 + 2 \lambda d_1 \bar{Y}^2 \left[ (A^2 + 2B)C^2_x - 4A \rho_{zx} C_z C_x + C^2_x \right]$$

$$- 2 \lambda \bar{Y}^2 \left[ BC^2_x - A \rho_{zx} C_z C_x \right] - 2 \lambda d_2 \bar{X} \bar{Y} \left[ \rho_{zx} C_z C_x - \left( A + \frac{1}{2} \right) C^2_x \right],$$

(4.22)

$$\frac{\partial \text{MSE}(\hat{\mu}_{GRR})}{\partial d_2} = 2 \lambda d_2 \bar{X}^2 C^2_x - 2 \lambda d_1 \bar{X} \bar{Y} \left[ \rho_{zx} C_z C_x - \left( A + \frac{1}{2} \right) \right] - \lambda \bar{X} \bar{Y} C^2_z.$$ 

(4.23)

Setting the first derivatives equal to zero, the optimum value of $d_1$ and $d_2$ are given by:

$$d_1(\text{opt}) = \frac{1 + \lambda \left[ (B - \frac{1}{2} A - \frac{1}{4}) C^2_x + \left( \frac{1}{2} - A \right) \rho_{zx} C_z C_x \right]}{1 + \lambda \left[ (2B - A - \frac{1}{4}) C^2_z + (1 - 2A) \rho_{zx} C_z C_x + (1 - \rho^2_{zx}) C^2_z \right]},$$

(4.24)

and

$$d_2(\text{opt}) = \frac{\bar{Y}}{\bar{X}} \left\{ \frac{1}{2} - d_1(\text{opt}) \left[ \left( A + \frac{1}{2} \right) - \rho_{zx} \frac{C_z}{C_x} \right] \right\}.$$ 

(4.25)
By substituting the optimum value of $d_2$ in the expression (4.21), we get:

\[
MSE_{\min}(\hat{\mu}_{GRR}) \approx d_1^2 \bar{Y}^2 - 2d_1 \bar{Y}^2 + \bar{Y}^2 + \lambda d_1^2 \bar{Y}^2 \left[ (A^2 + 2B) C_x^2 + 4A \rho_{zx} C_z C_x + C_x^2 \right] \\
+ \lambda \bar{X}^2 C_x^2 \bar{Y}^2 \left\{ \frac{1}{2} - d_1 \left[ \left( A + \frac{1}{2} \right) - \rho_{xx} \frac{C_z}{C_x} \right] \right\}^2 \\
- 2\lambda d_1 \bar{Y}^2 \left( BC_x^2 - A \rho_{zx} C_z C_x \right) \\
- 2\lambda \bar{X} \bar{y} d_1 X \left\{ \frac{1}{2} - d_1 \left[ \left( A + \frac{1}{2} \right) - \rho_{xx} \frac{C_z}{C_x} \right] \right\} \left[ \rho_{zx} C_x^2 C_x - \left( A + \frac{1}{2} \right) C_x^2 \right] \\
- \lambda \bar{X} \bar{y} C_x^2 X \left\{ \frac{1}{2} - d_1 \left[ \left( A + \frac{1}{2} \right) - \rho_{yx} \frac{C_y}{C_x} \right] \right\} \\
= d_1^2 \bar{Y}^2 \left\{ 1 + \left[ \left( 2B - A - \frac{1}{4} \right) C_x^2 + (1 - 2A) \rho_{zx} C_x C_x + (1 - \rho_{zx}^2) C_x^2 \right] \right\} \\
+ \bar{Y}^2 - \frac{1}{4} \lambda \bar{Y}^2 C_x^2 - 2d_1 \bar{Y}^2 \left\{ 1 + \lambda \left[ \left( B - \frac{1}{2} A - \frac{1}{4} \right) C_x^2 + \left( \frac{1}{2} - A \right) \rho_{zx} C_x C_x \right] \right\}.
\]

Then substituting the optimum value of $d_2$ in (4.26), the minimum mean square error of the proposed estimator, up to the first order of approximation, is given by:

\[
MSE_{\min}(\hat{\mu}_{GRR}) \approx \bar{Y}^2 \left\{ 1 - \frac{1}{4} \lambda C_x^2 \right\} \\
- \frac{\left\{ 1 + \lambda \left[ \left( B - \frac{1}{2} A - \frac{1}{4} \right) C_x^2 + (1 - A) \rho_{zx} C_x C_x \right] \right\}^2}{1 + \lambda \left[ \left( 2B - A - \frac{1}{4} \right) C_x^2 + (1 - 2A) \rho_{zx} C_x C_x + (1 - \rho_{zx}^2) C_x^2 \right]}.
\] (4.26)

For $\alpha = 1$ the generalized mixture estimator is given by:

\[
\hat{\mu}_{GRR1} = \left[ d_1 \bar{z} \left( \frac{\bar{X}}{\bar{x}} \right) + d_2 (\bar{X} - \bar{x}) \right] \exp \left( \frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right).
\] (4.27)
The optimum values of $d_1$ and $d_2$ are given by:

$$d_{1\text{GRR1}(\text{opt})} = \frac{1 + \left[ \frac{7}{8} C_x^2 - \rho_{xx} C_z C_x \right]}{1 + \lambda \left[ 2 C_x^2 - 2 \rho_{xx} C_z C_x + (1 - \rho_{xx}^2) C_z^2 \right]}, \quad \text{and} \quad (4.28)$$

$$d_{2\text{GRR1}(\text{opt})} = \frac{\bar{Y}}{\bar{X}} \left[ \frac{1}{2} - d_1(\text{opt}) \left( 2 - \rho_{xx} \frac{C_z}{C_x} \right) \right], \quad (4.29)$$

and the minimum mean square error, up to the first order of approximation, is given by:

$$MSE_{\text{min}}(\hat{\mu}_{\text{GRR1}}) \approx \bar{Y}^2 \left\{ \left( 1 - \frac{1}{4} \lambda C_x^2 \right) - \frac{\left[ 1 + \lambda \left( \frac{7}{8} C_x^2 - \rho_{xx} C_z C_x \right) \right]^2}{1 + \lambda \left[ 2 C_x^2 - 2 \rho_{xx} C_z C_x + (1 - \rho_{xx}^2) C_z^2 \right]} \right\}. \quad (4.30)$$

For $\alpha = 2$ the generalized mixture estimator is given by:

$$\hat{\mu}_{\text{GRR2}} = \left[ d_1 \bar{z} \left( \frac{X}{\bar{x}} \right)^2 + d_2 \left( \bar{X} - \bar{x} \right) \right] \exp \left( \frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right). \quad (4.30)$$

The optimum values of $d_1$ and $d_2$ are given by:

$$d_{1\text{GRR2}(\text{opt})} = \frac{1 + \lambda \left[ \frac{23}{8} C_x^2 - 2 \rho_{xx} C_z C_x \right]}{1 + \lambda \left[ 6 C_x^2 - 4 \rho_{xx} C_z C_x + (1 - \rho_{xx}^2) C_z^2 \right]}, \quad \text{and} \quad (4.31)$$

$$d_{2\text{GRR2}(\text{opt})} = \frac{\bar{Y}}{\bar{X}} \left[ \frac{1}{2} - d_1(\text{opt}) \left( 3 - \rho_{xx} \frac{C_z}{C_x} \right) \right], \quad (4.32)$$
and the minimum mean square error up to the first order of approximation, is given as:

\[
MSE_{\min}(\hat{\mu}_{GRR}) \approx \gamma^2 \left\{ \left( 1 - \frac{1}{4} \lambda C_y^2 \right) - \frac{\left[ 1 + \lambda \left( \frac{23}{8} C_x^2 - 2 \rho_{zz} C_y C_x \right) \right]^2}{1 + \lambda \left[ 6 C_x^2 - 4 \rho_{zz} C_z C_x + (1 - \rho_{zz}^2) C_y^2 \right]} \right\}.
\]

### 4.3 Efficiency Comparisons

In this section efficiency of the proposed estimator is compared with some of the commonly used RRT estimators. Conditions under which the generalized mixture RRT estimator is more efficient are given below:

\[
MSE(\hat{\mu}_{GRR}) < MSE(\mu_Y) \quad \text{if} \quad \lambda C_z^2 - \left\{ \left( 1 - \frac{1}{4} \lambda C_x^2 \right) \right\} > 0 \tag{4.33}
\]

\[
MSE(\hat{\mu}_{GRR}) < MSE(\hat{\mu}_R) \quad \text{if} \quad \lambda (C_x - \rho_{xz} C_y)^2 + \lambda (1 - \rho_{xz}^2) C_y^2 - \left\{ \left( 1 - \frac{1}{4} \lambda C_y^2 \right) \right\} > 0 \tag{4.34}
\]
It also holds:

\[ MSE(\hat{\mu}_{GRR}) < MSE(\hat{\mu}_{Reg}) \quad \text{if} \]

\[
\lambda \bar{Y}^2 C_x^2 (1 - \rho_{zx}^2) - \left\{ 1 - \frac{1}{4} \lambda C_x^2 \right\} > 0 \quad (4.35)
\]

\[ MSE(\hat{\mu}_{GRR}) < MSE(\hat{\mu}_{ER}) \quad \text{if} \]

\[
\lambda \left( \frac{1}{2} C_x - \rho_{zx} C_z \right)^2 + \lambda (1 - \rho_{zx}^2) C_z^2 - \left\{ 1 - \frac{1}{4} \lambda C_x^2 \right\} > 0 \quad (4.36)
\]

We will use the real data and simulated data to show that these conditions are generally true, and hence the proposed estimator may be preferred over the existing estimators.

### 4.4 Numerical Examples

In this section, we compare the efficiency of proposed estimators with other existing RRT mean estimators using real data. The Population Statistics for the real data are given in Table 10. The scrambling variable \( S \) is assume to be a normal distribution with mean zero and standard deviation equal to 2. The reported response is given by \( Z = Y + S \). Table 11 gives Theoretical Percent Relative Efficiency for various RRT estimators based on the first order of approximation. The Theoretical Percent
Relative Efficiency of the estimators as compared to the ordinary RRT sample mean are calculated from the following equation:

\[
PRET(\hat{\mu}_i) = 100 \times \frac{MSET(\hat{\mu}_y)}{MSET(\hat{\mu}_i)}
\]  

(4.37)

where \( i = R, \text{Reg}, \text{ER}, \text{GRR}, \text{GER}, \text{GRR1}, \text{and GRR2}. \)

Table 10. Summary Statistics for the Real Populations Used in Comparing \( \hat{\mu}_{\text{GRR}} \) with other RRT Mean Estimators

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Population 1</th>
<th>Population 2</th>
<th>Population 3</th>
<th>Population 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N )</td>
<td>70</td>
<td>34</td>
<td>256</td>
<td>204</td>
</tr>
<tr>
<td>( n )</td>
<td>25</td>
<td>20</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>( \rho_{yx} )</td>
<td>0.7293</td>
<td>0.4491</td>
<td>0.887</td>
<td>0.71</td>
</tr>
<tr>
<td>( \rho_{zx} )</td>
<td>0.81079</td>
<td>0.44909</td>
<td>0.8867</td>
<td>0.7099</td>
</tr>
<tr>
<td>( \bar{Y} )</td>
<td>96.7</td>
<td>856.4118</td>
<td>56.47</td>
<td>966</td>
</tr>
<tr>
<td>( \bar{X} )</td>
<td>175.2671</td>
<td>208.8824</td>
<td>44.45</td>
<td>26441</td>
</tr>
<tr>
<td>( S_{x}^{2} )</td>
<td>19842.15</td>
<td>22650.18</td>
<td>3872.573</td>
<td>2061327175</td>
</tr>
<tr>
<td>( S_{y}^{2} )</td>
<td>3657.368</td>
<td>537544.3</td>
<td>6430.019</td>
<td>5711084</td>
</tr>
<tr>
<td>( \sigma_{s}^{2} )</td>
<td>3.67395</td>
<td>3.67395</td>
<td>3.67395</td>
<td>3.67395</td>
</tr>
<tr>
<td>( C_{y} )</td>
<td>0.6254</td>
<td>0.8561</td>
<td>1.42</td>
<td>2.4739</td>
</tr>
<tr>
<td>( C_{x} )</td>
<td>0.8037</td>
<td>0.7205</td>
<td>1.40</td>
<td>1.7171</td>
</tr>
<tr>
<td>( C_{z} )</td>
<td>0.6257</td>
<td>0.8561</td>
<td>1.4204</td>
<td>2.4739</td>
</tr>
<tr>
<td>( f )</td>
<td>0.3571</td>
<td>0.5882</td>
<td>0.3906</td>
<td>0.2451</td>
</tr>
</tbody>
</table>
Population 1 [Source: Singh and Chaudhary (1986), pp.108]
Population 2 [Source: Singh and Chaudhary(1986), pp. 177]
Population 3 [Source: Cochran (1977), pp. 196]
Population 4 [Source: Kadilar & Cingi (2005)]

Table 11. The Theoretical Percent Relative Efficiency for RRT Mean Estimators

<table>
<thead>
<tr>
<th>Estimators</th>
<th>Population 1</th>
<th>Population 2</th>
<th>Population 3</th>
<th>Population 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\mu}_Y$</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>$\hat{\mu}_R$</td>
<td>76.3753</td>
<td>105.001</td>
<td>447.5094</td>
<td>201.5505</td>
</tr>
<tr>
<td>$\hat{\mu}_{Reg}$</td>
<td>291.875</td>
<td>125.2645</td>
<td>467.9889</td>
<td>201.6534</td>
</tr>
<tr>
<td>$\hat{\mu}_{ER}$</td>
<td>269.5187</td>
<td>125.1390</td>
<td>449.1049</td>
<td>159.3275</td>
</tr>
<tr>
<td>$\hat{\mu}_{GRR}$</td>
<td>292.8943</td>
<td>126.7898</td>
<td>472.3173</td>
<td>211.3242</td>
</tr>
<tr>
<td>$\hat{\mu}_{GER}$</td>
<td>294.468</td>
<td>127.1320</td>
<td>478.3395</td>
<td>213.413</td>
</tr>
<tr>
<td>$\hat{\mu}_{GRR1}$</td>
<td>303.6344</td>
<td>128.7935</td>
<td>485.3493</td>
<td>212.9479</td>
</tr>
<tr>
<td>$\hat{\mu}_{GRR2}$</td>
<td>431.1358</td>
<td>137.8521</td>
<td>775.2617</td>
<td>242.964</td>
</tr>
</tbody>
</table>
4.5 Conclusion

In this chapter, we proposed the generalized mixture estimator and the ordinary exponential ratio type estimator for the mean of a sensitive variable in simple random sampling without replacement by using information about a non-sensitive auxiliary variable. The proposed generalized mixture estimator is a mixture of some of the commonly known RRT estimators. Numerical examples show that for the proposed estimators all the percent relative efficiencies are greater than 100 indicating that all these estimators are better than the RRT ordinary mean estimator and RRT ratio type estimator. We also note that the proposed generalized mixture estimator is more efficient than the other estimators considered here. The estimator for $\alpha = 2$ performs better than the one with $\alpha = 1$ for the numerical examples.
CHAPTER V
SIMULATION RESULTS

5.1 Introduction

In this chapter we compare the performance of different estimators with the proposed generalized mixture estimators in the situations when the study variable is non-sensitive, and when the study variable is sensitive and cannot be observed directly. In both cases a non-sensitive auxiliary variable is available. The simulated populations we use have the same characteristics as the real data sets considered in the previous chapters. We will consider three finite populations of size \( N = 5000 \) each with summary statistics as presented in Table 12. The scrambling variable \( S \) is taken to be a normal variable with mean zero and standard deviation equal to 2. The reported response is given by \( Z = Y + S \). For each population, we consider the sample sizes \( n = 100, 200 \) and \( 500 \). The empirical mean square error is estimated based on 10,000 samples selected from each population. The R-code for the simulation study is given at the end of this dissertation.

Included in our comparisons will be the Sousa et al. (2010) transformed ratio type estimator given by:

\[
\hat{\mu}_{TR} = \bar{z} \left( \frac{c\bar{X} + d}{c\bar{x} + d} \right),
\]  

(5.1)
where $c$ and $d$ are the unit-free parameters. The mean square error of this estimator, up to the first order of approximation, is given by:

$$MSE(\hat{\mu}_{TR}) \approx \lambda \bar{Y}^2 \left( \eta^2 C_x^2 - 2\eta \rho_z \bar{Z} C_x + C_z^2 \right), \quad (5.2)$$

where $\eta = \frac{cX}{cX+d}$.

We will consider the four special cases given by:

\[
\begin{align*}
\hat{\mu}_{TR1} &= \frac{\bar{Z}}{\bar{x} + \beta_1(x)} \bar{X} + \beta_1(x), \\
\hat{\mu}_{TR2} &= \frac{\bar{Z}}{\bar{x} + \beta_2(x)} \bar{X} + \beta_2(x), \\
\hat{\mu}_{TR3} &= \frac{\bar{Z}}{\bar{Z} \beta_1(x) \bar{x} + \beta_2(x)} \beta_1(x) \bar{X} + \beta_2(x), \\
\hat{\mu}_{TR4} &= \frac{\bar{Z}}{\bar{Z} \beta_2(x) \bar{x} + \beta_1(x)} \beta_2(x) \bar{X} + \beta_1(x).
\end{align*}
\]

where $\beta_1(x)$ is the coefficient of skewness and $\beta_2(x)$ is the coefficient of kurtosis.

### 5.2 Simulation Results

First we will show the Empirical Percent Relative Efficiencies and the Theoretical Percent Relative Efficiencies (in bold) for the non-RRT mean estimators based on the first order of approximation. The results are given in the Table 13. For the RRT mean estimators, the results are given in Table 14.

We can confirm that all the percent relative efficiencies (except for the ordinary ratio estimator $\hat{\mu}_R$) are greater than 100 indicating that all these estimators are better than the ordinary sample mean estimator. The Theoretical Percent Relative Efficiencies suggest that estimators with $\alpha = 2$ perform better than the proposed estimators with $\alpha = 1$. 

84
For the RRT mean estimators the proposed generalized mixture estimators perform just like the non-RRT mean estimators.

Table 12. Summary Statistics Used in Comparing $\hat{\mu}_{GM}$, $\hat{\mu}_{GMR}$ and $\hat{\mu}_{GRR}$ with other Mean Estimators

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Population 1</th>
<th>Population 2</th>
<th>Population 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N)</td>
<td>5000</td>
<td>5000</td>
<td>5000</td>
</tr>
<tr>
<td>(\rho_{yx})</td>
<td>0.306173</td>
<td>0.6043378</td>
<td>0.8518795</td>
</tr>
<tr>
<td>(\rho_{zx})</td>
<td>0.3063796</td>
<td>0.6043924</td>
<td>0.8517596</td>
</tr>
<tr>
<td>(\bar{Y})</td>
<td>94.13349</td>
<td>94.54217</td>
<td>95.00613</td>
</tr>
<tr>
<td>(\bar{X})</td>
<td>61.48666</td>
<td>61.27170</td>
<td>60.99362</td>
</tr>
<tr>
<td>(S^2_y)</td>
<td>6295.845</td>
<td>6301.445</td>
<td>6299.421</td>
</tr>
<tr>
<td>(S^2_x)</td>
<td>3491.867</td>
<td>3503.012</td>
<td>3516.064</td>
</tr>
<tr>
<td>(\sigma^2_s)</td>
<td>4.040761</td>
<td>4.040761</td>
<td>4.040761</td>
</tr>
<tr>
<td>(C_y)</td>
<td>0.84291</td>
<td>0.83964</td>
<td>0.83541</td>
</tr>
<tr>
<td>(C_x)</td>
<td>0.96105</td>
<td>0.96596</td>
<td>0.97217</td>
</tr>
<tr>
<td>(C_z)</td>
<td>0.84294</td>
<td>0.8397219</td>
<td>0.8355349</td>
</tr>
<tr>
<td>(\beta_{1(x)})</td>
<td>0.00977</td>
<td>0.00896</td>
<td>0.00582</td>
</tr>
<tr>
<td>(\beta_{1(y)})</td>
<td>-0.01039</td>
<td>-0.01044</td>
<td>-0.00816482</td>
</tr>
<tr>
<td>(\beta_{2(x)})</td>
<td>2.92869</td>
<td>2.92334</td>
<td>2.925049</td>
</tr>
<tr>
<td>(\beta_{2(y)})</td>
<td>3.073645</td>
<td>3.049695</td>
<td>3.010963</td>
</tr>
</tbody>
</table>
The Theoretical and Empirical Percent Relative Efficiency for the non-RRT Mean Estimators:

Table 13. The Theoretical Percent Relative Efficiency (PRET) and the Empirical Percent Relative Efficiency (PREE) for the non-RRT Mean Estimators

<table>
<thead>
<tr>
<th>Estimators</th>
<th>n</th>
<th>Population 1</th>
<th></th>
<th>Population 2</th>
<th></th>
<th>Population 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>PREE</td>
<td>PRET</td>
<td></td>
<td>PREE</td>
<td>PRET</td>
<td></td>
</tr>
<tr>
<td>$\hat{\mu}_Y$</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>$\hat{\mu}_R$</td>
<td>58.0319</td>
<td>62.43019</td>
<td>103.8177</td>
<td>107.18</td>
<td>258.5247</td>
<td>269.1498</td>
<td></td>
</tr>
<tr>
<td>$\hat{\mu}_{Reg}$</td>
<td>108.9269</td>
<td>110.3438</td>
<td>156.777</td>
<td>157.5359</td>
<td>361.5658</td>
<td>364.5627</td>
<td></td>
</tr>
<tr>
<td>$\hat{\mu}_{ER}$</td>
<td>100.4883</td>
<td>102.469</td>
<td>156.7632</td>
<td>157.3259</td>
<td>284.8589</td>
<td>288.0067</td>
<td></td>
</tr>
<tr>
<td>$\hat{\mu}_{S,RP}$</td>
<td>108.2584</td>
<td>110.3438</td>
<td>154.1957</td>
<td>157.5359</td>
<td>342.6528</td>
<td>364.5627</td>
<td></td>
</tr>
<tr>
<td>$\hat{\mu}_{R,Reg}$</td>
<td>107.1271</td>
<td>111.0401</td>
<td>152.4903</td>
<td>158.2268</td>
<td>329.7562</td>
<td>365.2466</td>
<td></td>
</tr>
<tr>
<td>$\hat{\mu}_GK$</td>
<td>107.0532</td>
<td>111.3146</td>
<td>153.5348</td>
<td>158.6368</td>
<td>345.8642</td>
<td>366.3567</td>
<td></td>
</tr>
<tr>
<td>$\hat{\mu}_{GM1}$</td>
<td>106.7733</td>
<td>112.8668</td>
<td>149.5855</td>
<td>161.2368</td>
<td>302.5733</td>
<td>376.2145</td>
<td></td>
</tr>
<tr>
<td>$\hat{\mu}_{GM2}$</td>
<td>103.7155</td>
<td>115.1912</td>
<td>142.0656</td>
<td>165.4793</td>
<td>266.2331</td>
<td>395.8243</td>
<td></td>
</tr>
<tr>
<td>$\hat{\mu}_{GMR1}$</td>
<td>107.6861</td>
<td>113.8431</td>
<td>155.9745</td>
<td>161.0671</td>
<td>339.9154</td>
<td>370.5307</td>
<td></td>
</tr>
<tr>
<td>$\hat{\mu}_{GMR2}$</td>
<td>100.60872</td>
<td>127.3194</td>
<td>149.3095</td>
<td>178.9323</td>
<td>271.671</td>
<td>430.9847</td>
<td></td>
</tr>
</tbody>
</table>

| $\hat{\mu}_Y$ | 200 | 100  | 100  | 100  | 100  | 100  | 100  | 100  | 100  |
| $\hat{\mu}_R$ | 60.08251 | 62.43019 | 105.7923 | 107.18 | 258.7645 | 269.1498 |
| $\hat{\mu}_{Reg}$ | 109.8632 | 110.3438 | 157.5753 | 157.5359 | 357.2337 | 364.5627 |
| $\hat{\mu}_{ER}$ | 101.4283 | 102.469 | 157.2676 | 157.3259 | 284.2063 | 288.0067 |
| $\hat{\mu}_{S,RP}$ | 109.5522 | 110.3438 | 156.3128 | 157.5359 | 349.7985 | 364.5627 |
| $\hat{\mu}_{R,Reg}$ | 108.8401 | 110.6849 | 155.5144 | 158.8743 | 343.7958 | 364.8977 |
| $\hat{\mu}_GK$ | 108.7607 | 110.8187 | 155.9329 | 158.0744 | 351.5238 | 365.4400 |
| $\hat{\mu}_{GM1}$ | 108.6089 | 111.5749 | 154.1839 | 159.3396 | 329.7909 | 370.2103 |
| $\hat{\mu}_{GM2}$ | 107.2027 | 112.6952 | 150.2984 | 161.3735 | 307.7655 | 379.39
<table>
<thead>
<tr>
<th>Estimators</th>
<th>n</th>
<th>Population 1</th>
<th></th>
<th>Population 2</th>
<th></th>
<th>Population 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>PREE</td>
<td>PRET</td>
<td>PREE</td>
<td>PRET</td>
<td>PREE</td>
<td>PRET</td>
</tr>
<tr>
<td>$\hat{\mu}_{GMR1}$</td>
<td>108.9538</td>
<td>112.0474</td>
<td>157.6905</td>
<td>159.2574</td>
<td>350.1518</td>
<td>367.4717</td>
<td></td>
</tr>
<tr>
<td>$\hat{\mu}_{GMR2}$</td>
<td>108.7825</td>
<td>118.2626</td>
<td>143.0767</td>
<td>167.5455</td>
<td>274.3496</td>
<td>394.8204</td>
<td></td>
</tr>
<tr>
<td>$\hat{\mu}_Y$</td>
<td>500</td>
<td>61.0523</td>
<td>62.4302</td>
<td>107.1060</td>
<td>107.180</td>
<td>269.9781</td>
<td>269.1498</td>
</tr>
<tr>
<td>$\hat{\mu}_R$</td>
<td></td>
<td>110.1817</td>
<td>110.3438</td>
<td>156.1066</td>
<td>157.5359</td>
<td>365.7146</td>
<td>364.5627</td>
</tr>
<tr>
<td>$\hat{\mu}_{Reg}$</td>
<td></td>
<td>101.9275</td>
<td>102.469</td>
<td>155.4934</td>
<td>157.3259</td>
<td>288.2839</td>
<td>288.0067</td>
</tr>
<tr>
<td>$\hat{\mu}_{ER}$</td>
<td></td>
<td>109.8801</td>
<td>110.3438</td>
<td>155.3295</td>
<td>157.5359</td>
<td>361.9954</td>
<td>364.5627</td>
</tr>
<tr>
<td>$\hat{\mu}_{S,RP}$</td>
<td></td>
<td>109.6644</td>
<td>110.4717</td>
<td>155.1292</td>
<td>157.6628</td>
<td>359.6135</td>
<td>364.6883</td>
</tr>
<tr>
<td>$\hat{\mu}_{R,Reg}$</td>
<td></td>
<td>109.6368</td>
<td>110.5218</td>
<td>155.4442</td>
<td>157.7377</td>
<td>362.7923</td>
<td>364.8914</td>
</tr>
<tr>
<td>$\hat{\mu}_{GK}$</td>
<td></td>
<td>109.6038</td>
<td>110.8045</td>
<td>154.5626</td>
<td>158.2103</td>
<td>353.8349</td>
<td>366.6674</td>
</tr>
<tr>
<td>$\hat{\mu}_{GM1}$</td>
<td></td>
<td>109.2878</td>
<td>111.2206</td>
<td>152.8622</td>
<td>158.9633</td>
<td>343.9008</td>
<td>370.0193</td>
</tr>
<tr>
<td>$\hat{\mu}_{GM2}$</td>
<td></td>
<td>109.8011</td>
<td>110.9804</td>
<td>156.2002</td>
<td>158.1797</td>
<td>362.6607</td>
<td>365.6504</td>
</tr>
<tr>
<td>$\hat{\mu}_{GMR1}$</td>
<td></td>
<td>109.0028</td>
<td>113.2309</td>
<td>150.7239</td>
<td>161.1911</td>
<td>331.9076</td>
<td>375.4523</td>
</tr>
</tbody>
</table>
### Table 14. The Theoretical Percent Relative Efficiency (PRET) and the Empirical Percent Relative Efficiency (PREE) for the RRT Mean Estimators

| Estimators | $n$ | Population 1 | | Population 2 | | Population 3 | | PREE | PRET | PREE | PRET | PREE | PRET |
|------------|-----|--------------||--------------||--------------||--------|--------|--------|--------|--------|--------|
| $\hat{\mu}_Y$ | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| $\hat{\mu}_R$ | | 58.71567 | 62.48659 | 98.51078 | 107.2565 | 255.0575 | 269.1183 |
| $\hat{\mu}_{TR1}$ | | 58.72834 | 62.49837 | 98.53273 | 107.2776 | 255.1092 | 269.1685 |
| $\hat{\mu}_{TR2}$ | | 62.42388 | 65.93512 | 105.443 | 113.9063 | 279.7681 | 293.0594 |
| $\hat{\mu}_{TR3}$ | | 108.7746 | 108.9033 | 122.5912 | 123.0173 | 124.6077 | 124.8461 |
| $\hat{\mu}_{TR4}$ | | 58.7200 | 62.49061 | 98.51829 | 107.2637 | 255.0752 | 269.1355 |
| $\hat{\mu}_{Reg}$ | | 109.1013 | 110.4212 | 153.724 | 157.6236 | 357.1331 | 364.4144 |
| $\hat{\mu}_{ER}$ | | 100.9350 | 102.5526 | 153.4705 | 157.4118 | 283.6987 | 287.9487 |
| $\hat{\mu}_{GRR}$ | | 106.6441 | 111.1243 | 154.1921 | 158.3213 | 342.5125 | 365.1052 |
| $\hat{\mu}_{GER}$ | | 107.852 | 111.3926 | 151.8842 | 158.7251 | 329.1788 | 366.2081 |
| $\hat{\mu}_{GRR1}$ | | 107.8473 | 113.9221 | 153.2379 | 161.156 | 336.7513 | 370.3777 |
| $\hat{\mu}_{GRR2}$ | | 101.6116 | 127.4049 | 136.203 | 179.0270 | 273.0817 | 430.7435 |
| $\hat{\mu}_Y$ | 200 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| $\hat{\mu}_R$ | | 62.4458 | 62.48659 | 104.9488 | 107.2565 | 258.9979 | 269.1183 |
| $\hat{\mu}_{TR1}$ | | 62.45694 | 62.49837 | 104.9705 | 107.2776 | 259.0479 | 269.1685 |
| $\hat{\mu}_{TR2}$ | | 66.05773 | 65.93512 | 111.7769 | 113.9063 | 282.9415 | 293.0594 |
| $\hat{\mu}_{TR3}$ | | 109.1664 | 108.9033 | 122.9115 | 123.0173 | 124.6378 | 124.8461 |
| $\hat{\mu}_{TR4}$ | | 62.4488 | 62.49061 | 104.9563 | 107.2637 | 259.015 | 269.1355 |
| $\hat{\mu}_{Reg}$ | | 110.5068 | 110.4212 | 156.5281 | 157.6236 | 355.9265 | 364.4144 |
| $\hat{\mu}_{ER}$ | | 103.2897 | 102.5526 | 156.6336 | 157.4118 | 284.161 | 287.9487 |
| $\hat{\mu}_{GRR}$ | | 109.0753 | 110.764 | 157.9637 | 161.156 | 336.7513 | 370.3777 |
| $\hat{\mu}_{GER}$ | | 109.6907 | 110.8964 | 155.9092 | 158.1624 | 343.4442 | 365.2916 |

88
<table>
<thead>
<tr>
<th>Estimators</th>
<th>n</th>
<th>Population 1</th>
<th>Population 2</th>
<th>Population 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>PREE</td>
<td>PRET</td>
<td>PREE</td>
</tr>
<tr>
<td>$\hat{\mu}_{GRR1}$</td>
<td>109.4142</td>
<td>112.1255</td>
<td>156.6565</td>
<td>159.3457</td>
</tr>
<tr>
<td>$\hat{\mu}_{GRR2}$</td>
<td>109.9778</td>
<td>118.3439</td>
<td>150.6148</td>
<td>167.6366</td>
</tr>
<tr>
<td>$\hat{\mu}_Y$</td>
<td>500</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>$\hat{\mu}_R$</td>
<td>62.28899</td>
<td>62.48659</td>
<td>103.5708</td>
<td>107.2565</td>
</tr>
<tr>
<td>$\hat{\mu}_{TR1}$</td>
<td>62.3010</td>
<td>62.49837</td>
<td>103.5917</td>
<td>107.2776</td>
</tr>
<tr>
<td>$\hat{\mu}_{TR2}$</td>
<td>65.80321</td>
<td>65.93512</td>
<td>110.155</td>
<td>113.9063</td>
</tr>
<tr>
<td>$\hat{\mu}_{TR3}$</td>
<td>108.9663</td>
<td>108.9033</td>
<td>122.5472</td>
<td>123.0173</td>
</tr>
<tr>
<td>$\hat{\mu}_{TR4}$</td>
<td>62.29309</td>
<td>62.49061</td>
<td>103.5779</td>
<td>107.2637</td>
</tr>
<tr>
<td>$\hat{\mu}_{Reg}$</td>
<td>110.5143</td>
<td>110.4212</td>
<td>154.6682</td>
<td>157.6236</td>
</tr>
<tr>
<td>$\hat{\mu}_{ER}$</td>
<td>102.6987</td>
<td>102.5526</td>
<td>154.7359</td>
<td>157.4118</td>
</tr>
<tr>
<td>$\hat{\mu}_{GRR}$</td>
<td>110.3119</td>
<td>110.5494</td>
<td>154.9472</td>
<td>157.7508</td>
</tr>
<tr>
<td>$\hat{\mu}_{GER}$</td>
<td>110.4029</td>
<td>110.5993</td>
<td>154.5534</td>
<td>157.8255</td>
</tr>
<tr>
<td>$\hat{\mu}_{GRR1}$</td>
<td>110.4402</td>
<td>111.058</td>
<td>155.0007</td>
<td>158.2676</td>
</tr>
<tr>
<td>$\hat{\mu}_{GRR2}$</td>
<td>109.636</td>
<td>113.3097</td>
<td>153.0042</td>
<td>161.2800</td>
</tr>
</tbody>
</table>
CHAPTER VI
OVERALL CONCLUSIONS AND FUTURE WORK

In this dissertation we proposed some generalized mixture estimators which are mixtures of the commonly used mean estimators. We also extended the proposed estimators to the situation when the study variable is sensitive and a non-sensitive auxiliary variable is available.

One may extend the proposed estimators to the more complex optional RRT situations. The parameters one need to estimate for optimal models are the population mean of the sensitive study variable and the proportion of respondents who consider the question sensitive (the sensitivity level $W$ of the study variable). Since in the optional RRT models two parameters need to be estimated, a larger sample size is needed.

We have examined adequacy of the first order approximation for the theoretical mean square errors in the ratio type mean estimators and the ratio type variance estimators. One can examine the adequacy of second order of approximation and robustness in the ratio and regression type estimators as well.

The proposed estimators may also be extended to more complex sampling designs as well such stratified sampling and multi-stage sampling.

One could also consider the multivariate case where more than one non-sensitive auxiliary variables are available.
REFERENCES


library("copula")

# Generate N=5,000 observation of bivariate normal, bivariate Poisson and
# bivariate gamma distributions with the population mean (X,Y)=(4,6) and
# the population variance (4,9)
distributions with mean(x)=4 and mean(y)=6.

# mv.BN<- mvdc(normalCopula(0.5), c("norm", "norm"),list(list(mean=4,sd=2),
# list(mean=6,sd=3)))
# lambda1<-4; lambda2<-6
# mv.PP<- mvdc(normalCopula(-0.5), c("pois", "pois"),
# list(list(lambda=-lambda1), list(lambda<-lambda2)))
# alpha1<-4; beta1<-1; alpha2<-4; beta2<-1.5
# mg.GG <- mvdc(normalCopula(0.2), c("gamma", "gamma"),list(list(shape=
# alpha1, scale=beta1), list(shape=alpha2,scale=beta2)))

# population.data<-rMvdc(5,000,mv.BN); population.data<-rMvdc(5,000,mv.PP)
# and population.data<-rMvdc(5,000,mv.GG);
# population.data<-read.table("population.data.N.0.8",header=TRUE)
X<-population.data$X ; Y<-population.data$Y

N<-nrow(population.data); rhoxy=cor(X,Y); mux<-mean(X); muy<-mean(Y);
varx<-var(X); vary<-var(Y); stdevx<-sd(X); stdevy<-sd(Y);
library("moments") #install packages for skewness and kurtosis
beta1x<-skewness(X); beta2x<-kurtosis(X)# skewness and kurtosis for X
beta1y<-skewness(Y); beta2y<-kurtosis(Y)# skewness and kurtosis for Y
cx<-stdevx/mux; cy<-stdevy/muy; md<-quantile(X,0.5) # coeff. of var.

##### THE SIMULATION PART ######

xbar<-numeric(b); ybar<-numeric(b); tR<-numeric(b); tP<-numeric(b);
tER<-numeric(b); tEP<-numeric(b); t1<-numeric(b); t2<-numeric(b);
t3<-numeric(b); t4<-numeric(b); t5<-numeric(b); t6<-numeric(b);
t7<-numeric(b); t8<-numeric(b); t9<-numeric(b); t10<-numeric(b);
t11<-numeric(b); t12<-numeric(b); t13<-numeric(b); t14<-numeric(b);
t15<-numeric(b); t16<-numeric(b); t17<-numeric(b); t18<-numeric(b);
t19<-numeric(b); t20<-numeric(b); t21<-numeric(b)
mse.tR<-numeric(b); mse.tP<-numeric(b); mse.tEP<-numeric(b);
mse.tER<-numeric(b); mse.t1<-numeric(b); mse.t2<-numeric(b);
mse.t3<-numeric(b); mse.t4<-numeric(b); mse.t5<-numeric(b);
mse.t6<-numeric(b); mse.t7<-numeric(b); mse.t8<-numeric(b);
mse.t9<-numeric(b); mse.t10<-numeric(b); mse.t11<-numeric(b);
mse.t12<-numeric(b); mse.t13<-numeric(b); mse.t14<-numeric(b);
mse.t15<-numeric(b); mse.t16<-numeric(b); mse.t17<-numeric(b);
mse.t18<-numeric(b); mse.t19<-numeric(b); mse.t20<-numeric(b);
mse.t21<-numeric(b); n<-100; b<-10000; f<-n/N # n=100, 200 and 500
for (k in 1:b){s<-sample(N,n,replace=TRUE);
xbar[k]<-mean(X[s])
ybar[k]<-mean(Y[s])
tR[k]<-ybar[k]*(mux/xbar[k])}
\[ t_P[k] <- ybar[k] * (xbar[k] / \mu_x) \]
\[ t_{ER}[k] <- ybar[k] * \exp((\mu_x - xbar[k]) / (\mu_x + xbar[k])) \]
\[ t_{EP}[k] <- ybar[k] * \exp((xbar[k] - \mu_x) / (xbar[k] + \mu_x)) \]
\[ t_1[k] <- ybar[k] * ((\mu_x + c_x) / (xbar[k] + c_x)) \]
\[ t_2[k] <- ybar[k] * ((xbar[k] + c_x) / (\mu_x + c_x)) \]
\[ t_3[k] <- ybar[k] * ((beta_2x * xbar[k] + c_x) / (beta_2x * \mu_x + c_x)) \]
\[ t_4[k] <- ybar[k] * ((c_x * xbar[k] + beta_2x) / (c_x * \mu_x + beta_2x)) \]
\[ t_5[k] <- ybar[k] * ((xbar[k] + \sigma_{x}) / (\mu_x + \sigma_{x})) \]
\[ t_6[k] <- ybar[k] * ((\mu_x + \sigma_{x}) / (xbar[k] + \sigma_{x})) \]
\[ t_7[k] <- ybar[k] * ((beta_1x * xbar[k] + \sigma_{x}) / (beta_1x * \mu_x + \sigma_{x})) \]
\[ t_8[k] <- ybar[k] * ((beta_2x * xbar[k] + \sigma_{x}) / (beta_2x * \mu_x + \sigma_{x})) \]
\[ t_9[k] <- ybar[k] * ((\mu_x + \rho_{yx}) / (xbar[k] + \rho_{yx})) \]
\[ t_{10}[k] <- ybar[k] * ((\mu_x + beta_2x) / (xbar[k] + beta_2x)) \]
\[ t_{11}[k] <- ybar[k] * ((xbar[k] + beta_2x) / (\mu_x + beta_2x)) \]
\[ t_{12}[k] <- ybar[k] * ((beta_2x * \mu_x + beta_1x) / (beta_2x * xbar[k] + beta_1x)) \]
\[ t_{13}[k] <- ybar[k] * ((\mu_x + \sigma_{y}) / (xbar[k] + \sigma_{y})) \]
\[ t_{14}[k] <- ybar[k] * ((\mu_x + \mu_d) / (xbar[k] + \mu_d)) \]
\[ t_{15}[k] <- ybar[k] * ((beta_2x * \mu_x + c_x) / (beta_2x * xbar[k] + c_x)) \]
\[ t_{16}[k] <- ybar[k] * ((c_x * \mu_x + \rho_{yx}) / (c_x * xbar[k] + \rho_{yx})) \]
\[ t_{17}[k] <- ybar[k] * ((c_x * \mu_y + \mu_d) / (c_x * xbar[k] + \mu_d)) \]
\[ t_{18}[k] <- ybar[k] * ((\rho_{yx} * \mu_y + \rho_{yx}) / (\rho_{yx} * xbar[k] + \rho_{yx})) \]
\[ t_{19}[k] <- ybar[k] * ((\rho_{yx} * \mu_x + \mu_y) / (\rho_{yx} * xbar[k] + \mu_y)) \]
\[ t_{20}[k] <- ybar[k] * ((\rho_{yx} * \mu_x + \mu_d) / (\rho_{yx} * xbar[k] + \mu_d)) \]
\[ t_{21}[k] <- ybar[k] * ((\sigma_{y} * \mu_x + \mu_d) / (\sigma_{y} * xbar[k] + \mu_d)) \]
\[ mse.tR[k] <- (tR[k] - \mu_y)^2; mse.tR.E <- mean(mse.tR) \]
\[ mse.tP[k] <- (tP[k] - \mu_y)^2; mse.tP.E <- mean(mse.tP) \]
mse.tER[k]<-(tER[k]-muy)^2;mse.tER.E<-mean(mse.tER)
mse.tEP[k]<-(tEP[k]-muy)^2;mse.tEP.E<-mean(mse.tEP)
mse.t1[k]<-(t1[k]-muy)^2;mse.t1.E<-mean(mse.t1)
mse.t2[k]<-(t2[k]-muy)^2;mse.t2.E<-mean(mse.t2)
mse.t3[k]<-(t3[k]-muy)^2;mse.t3.E<-mean(mse.t3)
mse.t4[k]<-(t4[k]-muy)^2;mse.t4.E<-mean(mse.t4)
mse.t5[k]<-(t5[k]-muy)^2;mse.t5.E<-mean(mse.t5)
mse.t6[k]<-(t6[k]-muy)^2;mse.t6.E<-mean(mse.t6)
mse.t7[k]<-(t7[k]-muy)^2;mse.t7.E<-mean(mse.t7)
mse.t8[k]<-(t8[k]-muy)^2;mse.t8.E<-mean(mse.t8)
mse.t9[k]<-(t9[k]-muy)^2;mse.t9.E<-mean(mse.t9)
mse.t10[k]<-(t10[k]-muy)^2;mse.t10.E<-mean(mse.t10)
mse.t11[k]<-(t11[k]-muy)^2;mse.t11.E<-mean(mse.t11)
mse.t12[k]<-(t12[k]-muy)^2;mse.t12.E<-mean(mse.t12)
mse.t13[k]<-(t13[k]-muy)^2;mse.t13.E<-mean(mse.t13)
mse.t14[k]<-(t14[k]-muy)^2;mse.t14.E<-mean(mse.t14)
mse.t15[k]<-(t15[k]-muy)^2;mse.t15.E<-mean(mse.t15)
mse.t16[k]<-(t16[k]-muy)^2;mse.t16.E<-mean(mse.t16)
mse.t17[k]<-(t17[k]-muy)^2;mse.t17.E<-mean(mse.t17)
mse.t18[k]<-(t18[k]-muy)^2;mse.t18.E<-mean(mse.t18)
mse.t19[k]<-(t19[k]-muy)^2;mse.t19.E<-mean(mse.t19)
mse.t20[k]<-(t20[k]-muy)^2;mse.t20.E<-mean(mse.t20)
mse.t21[k]<-(t21[k]-muy)^2;mse.t21.E<-mean(mse.t21)
}

#THE MEANS OVER 10,000 TRIALS FOR THE RATIO MEAN ESTIMATORS
tR.bar<-mean(tR); tP.bar<-mean(tP); tER.bar<-mean(tER);
tEP.bar<-mean(tEP); t1.bar<-mean(t1); t2.bar<-mean(t2);
t3.bar<-mean(t3); t4.bar<-mean(t4); t5.bar<-mean(t5);
t6.bar<-mean(t6); t7.bar<-mean(t7); t8.bar<-mean(t8);
t9.bar<-mean(t9); t10.bar<-mean(t10); t11.bar<-mean(t11);
t12.bar<-mean(t12); t13.bar<-mean(t13); t14.bar<-mean(t14);
t15.bar<-mean(t15); t16.bar<-mean(t16); t17.bar<-mean(t17);
t18.bar<-mean(t18); t19.bar<-mean(t19); t20.bar<-mean(t20);
t21.bar<-mean(t21)

mean.E<-cbind(tR.bar,tP.bar,tER.bar,tEP.bar,t1.bar,t2.bar,t3.bar,
t4.bar,t5.bar,t6.bar,t7.bar,t8.bar,t9.bar,t10.bar,t11.bar,t12.bar,
t13.bar, t14.bar,t15.bar,t16.bar,t17.bar,t18.bar,t19.bar,t20.bar,
t21.bar)

#vector with the empirical mean square errors for the mean estimators
mse.emp.table<-c(mse.tR.E, mse.tP.E, mse.tER.E, mse.tEP.E, mse.t1.E,
mse.t21.E)

# # # # T H E   T H E O R E T I C A L   P A R T   # # # # #

theta1<-theta2<-mux/(mux+cx)
theta3<-theta15<-(beta2x*mux)/(beta2x*mux+cx)
theta4<-(cx*mux)/(cx*mux+beta2x)
theta5<-theta13<-mux/(mux+stdevx)
theta6<-(beta1x*mux)/(beta1x*mux+stdevx)
\[
\theta_7 = \frac{(\beta_2x \cdot \mu_x)}{(\beta_2x \cdot \mu_x + \text{std}_x)}
\]
\[
\theta_8 = \theta_9 = \frac{\mu_x}{\mu_x + \rho_{y|x}}
\]
\[
\theta_{10} = \theta_{11} = \frac{\mu_x}{\mu_x + \beta_2x}
\]
\[
\theta_{12} = \frac{(\beta_2x \cdot \mu_x)}{(\beta_2x \cdot \mu_x + \beta_{1x})}
\]
\[
\theta_{14} = \frac{\mu_x}{\mu_x + \text{md}}
\]
\[
\theta_{16} = \frac{(\text{cx} \cdot \mu_x)}{(\text{cx} \cdot \mu_x + \rho_{y|x}}
\]
\[
\theta_{17} = \frac{(\text{cx} \cdot \mu_x)}{(\text{cx} \cdot \mu_x + \text{md}}
\]
\[
\theta_{18} = \frac{(\beta_{1x} \cdot \mu_x)}{(\beta_{1x} \cdot \mu_x + \rho_{y|x})}
\]
\[
\theta_{19} = \frac{(\rho_{y|x} \cdot \mu_x)}{(\rho_{y|x} \cdot \mu_x + \text{std}_x)}
\]
\[
\theta_{20} = \frac{(\rho_{y|x} \cdot \mu_x)}{(\rho_{y|x} \cdot \mu_x + \text{md})}
\]
\[
\theta_{21} = \frac{(\text{std}_x \cdot \mu_x)}{(\text{std}_x \cdot \mu_x + \text{md}}
\]

\[
\text{mse.ratio} = \text{function(\theta)}\{
(\frac{(1-f)}{n}) \cdot \mu_y^2 \cdot (\text{cy}^2 + \theta \cdot \text{cx}^2 \cdot (\theta - 2 \cdot \text{co}))
\}
\]
\[
\text{mse.product} = \text{function(\theta)}\{
(\frac{(1-f)}{n}) \cdot \mu_y^2 \cdot (\text{cy}^2 + \theta \cdot \text{cx}^2 \cdot (\theta + 2 \cdot \text{co}))
\}
\]
\[
\text{co} = \rho_{y|x} \cdot (\text{cy}/\text{cx}); \text{f} = n/N
\]

# THE THEORETICAL MEAN SQUARE ERRORS FOR THE MEAN ESTIMATORS #
\[
\text{mse.tr.th} = (\frac{(1-f)}{n}) \cdot \mu_y^2 \cdot (\text{cx}^2 - 2 \cdot \rho_{y|x} \cdot \text{cx} \cdot \text{cy} + \text{cy}^2)
\]
\[
\text{mse.tp.th} = (\frac{(1-f)}{n}) \cdot \mu_y^2 \cdot (\text{cx}^2 + 2 \cdot \rho_{y|x} \cdot \text{cx} \cdot \text{cy} + \text{cy}^2)
\]
\[
\text{mse.tER.th} = (\frac{(1-f)}{n}) \cdot (\frac{1}{4}) \cdot \mu_y^2 \cdot (\text{cx}^2 - 4 \cdot \rho_{y|x} \cdot \text{cx} \cdot \text{cy} + 4 \cdot \text{cy}^2)
\]
\[
\text{mse.tEP.th} = (\frac{(1-f)}{n}) \cdot (\frac{1}{4}) \cdot \mu_y^2 \cdot (\text{cx}^2 + 4 \cdot \rho_{y|x} \cdot \text{cx} \cdot \text{cy} + 4 \cdot \text{cy}^2)
\]
\[
\text{mse.t1.th} = \text{mse.ratio(\theta_{1})}; \text{mse.t2.th} = \text{mse.product(\theta_{2})}
\]
\[
\text{mse.t3.th} = \text{mse.product(\theta_{3})}; \text{mse.t4.th} = \text{mse.product(\theta_{4})}
\]
mse.t5.th<-mse.product(theta5); mse.t6.th<-mse.product(theta6)
mse.t7.th<-mse.product(theta7); mse.t8.th<-mse.ratio(theta8)
mse.t9.th<-mse.product(theta9); mse.t10.th<-mse.ratio(theta10)
mse.t11.th<-mse.product(theta11); mse.t12.th<-mse.ratio(theta12)
mse.t13.th<-mse.ratio(theta13); mse.t14.th<-mse.ratio(theta14)
mse.t15.th<-mse.ratio(theta15); mse.t16.th<-mse.ratio(theta16)
mse.t17.th<-mse.ratio(theta17); mse.t18.th<-mse.ratio(theta18)
mse.t19.th<-mse.ratio(theta19); mse.t20.th<-mse.ratio(theta20)
mse.t21.th<-mse.ratio(theta21)

# vector with the theoretical mean square errors for the mean estimators
mse.th.table<-c(mse.tr.th,mse.tp.th,mse.tER.th,mse.tEP.th,mse.t1.th,
mse.t2.th,mse.t3.th,mse.t4.th,mse.t5.th,mse.t6.th,mse.t7.th,mse.t8.th,
mse.t9.th, mse.t10.th, mse.t11.th, mse.t12.th, mse.t13.th, mse.t14.th,
mse.t15.th, mse.t16.th, mse.t17.th, mse.t18.th,mse.t19.th,mse.t20.th,
mse.t21.th)

# the ratio of the mean estimators
mse.ratio<-100*(mse.emp.table/mse.th.table)

# the ratio of the empirical and corresponding theoretical mean
# square errors for the ratio type mean estimators
ratio.mean<-cbind(r.R=mse.ratio[1],r.tER=mse.ratio[3],
r.t1.R=mse.ratio[5],
r.t8.R=mse.ratio[12],r.t10.R=mse.ratio[14],r.t12.R=mse.ratio[16],
r.t13.R=mse.ratio[17],r.t14.R=mse.ratio[18],r.t15.R=mse.ratio[19],
r.t16.R=mse.ratio[20],r.t17.R=mse.ratio[21],r.t18.R=mse.ratio[22],
r.t19.R=mse.ratio[23],r.t20.R=mse.ratio[24],r.t21.R=mse.ratio[25])

# the ratio of the empirical and corresponding theoretical mean
# square errors for the product type mean estimators

```
product.mean<-cbind(r.P=mse.ratio[2], r.tEP=mse.ratio[4],
                    r.t2.P=mse.ratio[6],
                    r.t3.P=mse.ratio[7], r.t4.P=mse.ratio[8], r.t5.P=mse.ratio[9],
                    r.t6.P=mse.ratio[10], r.t7.P=mse.ratio[11], r.t9.P=mse.ratio[13],
                    r.t11.P=mse.ratio[15])
```

# the summary statistics for the variance part

```
mu40<-(1/N)*sum((Y-muy)^4); mu20<-(1/N)*sum((Y-muy)^2);
Beta2.y<-mu40/(mu20^2);
mu04<-(1/N)*sum((X-mux)^4); mu02<-(1/N)*sum((X-mux)^2);
Beta2.x<-mu04/(mu02^2)
mu22<-(1/N)*sum((Y-muy)^2*(X-mux)^2); lambda22<-mu22/(mu20*mu02)
q2<-quantile(X,0.5); md<-q2; q1<-quantile(X,0.25,names=FALSE);
q3<-quantile(X,0.75,names=FALSE); qa<-(q3+q1)/2; Lambda<-1/n
```

### THE SIMULATION PART ###

```
s.x<-numeric(b); s.y<-numeric(b); Sr<-numeric(b); Sexp<-numeric(b);
S1<-numeric(b); S2<-numeric(b); S3<-numeric(b); S4<-numeric(b);
S5<-numeric(b); S6<-numeric(b); mse.Sr<-numeric(b);
mse.Sexp<-numeric(b);mse.S1<-numeric(b); mse.S2<-numeric(b);
mse.S3<-numeric(b);mse.S4<-numeric(b);
mse.S5<-numeric(b); mse.S6<-numeric(b)
for (k in 1:b){s<-sample(N,n,replace=TRUE);
s.x[k]<-var(X[s])
s.y[k]<-var(Y[s])
```
\[ Sr[k] = -s.y[k] \times \frac{\text{varx}}{s.x[k]} \] # Isaki (1983)

\[ Sexp[k] = -s.y[k] \times \exp\left(\frac{\text{varx} - s.x[k]}{\text{varx} + s.x[k]}\right) \]

\[ S1[k] = -s.y[k] \times \frac{\text{varx} + cx}{s.x[k] + cx} \]

\[ S2[k] = -s.y[k] \times \frac{\text{varx} + \beta2x}{s.x[k] + \beta2x} \]

\[ S3[k] = -s.y[k] \times \frac{\text{varx} \times \beta2x + cx}{s.x[k] \times \beta2x + cx} \]

\[ S4[k] = -s.y[k] \times \frac{\text{varx} \times cx + \beta2x}{s.x[k] \times cx + \beta2x} \]

\[ S5[k] = -s.y[k] \times \frac{\text{varx} + qa}{s.x[k] + qa} \]

\[ S6[k] = -s.y[k] \times \frac{\text{varx} \times cx + md}{s.x[k] \times cx + md} \]

\[ \text{mse.Sr}[k] = (Sr[k] - \text{vary})^2; \text{mse.Sr.E} = \text{mean(mse.Sr)} \]

\[ \text{mse.Sexp}[k] = (Sexp[k] - \text{vary})^2; \text{mse.Sexp.E} = \text{mean(mse.Sexp)} \]

\[ \text{mse.S1}[k] = (S1[k] - \text{vary})^2; \text{mse.S1.E} = \text{mean(mse.S1)} \]

\[ \text{mse.S2}[k] = (S2[k] - \text{vary})^2; \text{mse.S2.E} = \text{mean(mse.S2)} \]

\[ \text{mse.S3}[k] = (S3[k] - \text{vary})^2; \text{mse.S3.E} = \text{mean(mse.S3)} \]

\[ \text{mse.S4}[k] = (S4[k] - \text{vary})^2; \text{mse.S4.E} = \text{mean(mse.S4)} \]

\[ \text{mse.S5}[k] = (S5[k] - \text{vary})^2; \text{mse.S5.E} = \text{mean(mse.S5)} \]

\[ \text{mse.S6}[k] = (S6[k] - \text{vary})^2; \text{mse.S6.E} = \text{mean(mse.S6)} \]

\}

# vector contains the empirical mean square errors for the var. est.


### # # # THEORETICAL PART # # # #

\[ R0 = 1 \]

\[ R1 = -\text{varx}/(\text{varx} + \text{cx}) \]

\[ R2 = -\text{varx}/(\text{varx} + \beta2x) \]

\[ R3 = -\text{varx} \times \beta2x/(\text{varx} \times \beta2x + \text{cx}) \]
R4 <- varx * cx / (varx * cx + beta2x)
R5 <- varx / (varx + qa)
R6 <- varx / (varx + md)

mse.var <- function(c) {
    Lambda * (vary)^2 * ((beta2y - 1) + c^2 * (beta2x - 1) - 2 * c * (lambda22 - 1))
}

# THEORETICAL MEAN SQUARE ERRORS FOR THE VARIANCE ESTIMATORS
mse.Sexp.th <- Lambda * (vary)^2 * ((beta2y - 1) + (1/4) * (beta2x - 1)
    - (lambda22 - 1))
ms.Sr.th <- mse.var(R0)
ms.S1.th <- mse.var(R1)
ms.S2.th <- mse.var(R2)
ms.S3.th <- mse.var(R3)
ms.S4.th <- mse.var(R4)
ms.S5.th <- mse.var(R5)
ms.S6.th <- mse.var(R6)

# vector with the theoretical mean square errors for the ratio
# type variance estimators
mse.var.th <- c(ms.Sr.th, mse.Sexp.th, ms.S1.th, ms.S2.th, ms.S3.th,
    ms.S4.th, ms.S5.th, ms.S6.th)

# the ratio of the empirical and corresponding theoretical mean square
# errors for the ratio type variance estimators
mse.var.ratio <- -100 * (mse.var.emp / mse.var.th)
ratio.var <- cbind(r.Sr = mse.var.ratio[1], r.Sexp = mse.var.ratio[2],
    r.S1 = mse.var.ratio[3], r.S2 = mse.var.ratio[4], r.S3 = mse.var.ratio[5],
    r.S4 = mse.var.ratio[6], r.S5 = mse.var.ratio[7], r.S6 = mse.var.ratio[8])
APPENDIX B

R CODE II

population.data<-read.table("population.data",header=TRUE)
X<-population.data$X; Y<-population.data$Y

# # # T H E S U M M A R Y S T A T I S T I C S # # #

N<-nrow(population.data); rhoyx=cor(X,Y); mux<-mean(X); muy<-mean(Y);
varx<-var(X); vary<-var(Y); stdevx<-sd(X); stdevy<-sd(Y);
beta1x<-skewness(X); betaly<-skewness(Y); beta2x<-kurtosis(X);
beta2y<-kurtosis(Y); cx<-stdevx/mux; cy<-stdevy/muy

# # # T H E S I M U L A T I O N P A R T # # #

xbar<-numeric(b);ybar<-numeric(b);var.samplex<-numeric(b);
muER<-numeric(b);
var.sampley<-numeric(b); rhoyx.sample<-numeric(b); muR<-numeric(b);
muReg<-numeric(b); muSRP<-numeric(b); muRRReg<-numeric(b);
muGK<-numeric(b);
uuGM1<-numeric(b); muGM2<-numeric(b); muGMR1<-numeric(b);
muGMR2<-numeric(b);
mse.ybar<-numeric(b); mse.muER<-numeric(b); mse.muR<-numeric(b);
mse.muReg<-numeric(b); mse.muSRP<-numeric(b); mse.muRRReg<-numeric(b);
mse.muGK<-numeric(b); mse.muGM1<-numeric(b); mse.muGM2<-numeric(b);
mse.muGMR1<-numeric(b); mse.muGMR2<-numeric(b); sd.samplex<-numeric(b);
sd.sampley<-numeric(b); sd.samplexy<-numeric(b); beta.sample<-numeric(b);
alpha.opt<-numeric(b); cy.sample<-numeric(b); k1opt<-numeric(b);
k2opt<-numeric(b); l1opt<-numeric(b); l2optn<-numeric(b);
l2optd<-numeric(b);
l2opt<-numeric(b); d1optGM1<-numeric(b); d2optGM1<-numeric(b);
d1optGM2<-numeric(b); d2optGM2<-numeric(b); k1optGMR1<-numeric(b);
k2optGMR1<-numeric(b); k1optGMR2<-numeric(b); k2optGMR2<-numeric(b);
n<-100; b<-10000; f<-n/N; lambda<-(1-f)/n # n=100, 200 and 500
for (k in 1:b){s<-sample(N,n,replace=TRUE); # print(s)
xbar[k]<-mean(X[s]); ybar[k]<-mean(Y[s]); rhoyx.sample[k]<-cor(X[s],Y[s]);
var.samplex[k]<-var(X[s]); var.sampley[k]<-var(Y[s]);
sd.sampley[k]<-sd(Y[s]); sd.samplexy[k]<-cov(X[s],Y[s]);
cy.sample[k]<-sd.sampley[k]/ybar[k];
beta.sample[k]<-sd.samplexy[k]/var.samplex[k];
alpha.opt[k]<-(0.5)+(rhoux.sample[k]*(cy.sample[k]/cx))
k1opt[k]<-1/(1+lambda*(1-rhoux.sample[k]^2)*cy.sample[k]^2)
k2opt[k]<-(ybar[k]/xbar[k])*((rhoux.sample[k]*cy.sample[k])/cx)
l1opt[k]<-((-8)+lambda*cx^2)(8*(-1+lambda*(rhoux.sample[k]^2-1)*cy.sample[k]^2))
l2optn[k]<-ybar[k]*((-8)*rhoux.sample[k]*cy.sample[k]+cx*(4-lambda*cx^2)
-lambda*rhoux.sample[k]*cx*cy.sample[k]+4*lambda*(rhoux.sample[k]^2-1)*
cy.sample[k]^2))
l2optd[k]<-ybar[k]*((-8)*rhoux.sample[k]*cy.sample[k]+cx*(4-lambda*cx^2)
-lambda*rhoux.sample[k]*cx*cy.sample[k]+4*lambda*(rhoux.sample[k]^2-1)*
cy.sample[k]^2))
l2optn[k]<-ybar[k]*(-1+lambda*(rhoux.sample[k]^2-1)*cy.sample[k]^2)
l2opt[k]<-l2optn[k]/l2optd[k]
d1optGM1[k]<-((1+(3/8)*lambda*cx^2)/(1+lambda*cy.sample[k]^2)*
(1-rhoux.sample[k]^2)+lambda*cx^2)
\[ d_{optGM1}[k] <- \frac{ybar[k]}{xbar[k]} \times (0.5 - d_{1optGM1}[k] \times (1 - \rho_{oyx.sample}[k] \times (cy.sample[k] / cx)) \]

\[ d_{1optGM2}[k] <- \frac{1 + (7/8) \lambda \times cx^2}{1 + \lambda \times (1 - \rho_{oyx.sample}[k]^2) \times cy.sample[k]^2 + 2 \lambda \times cx^2} \]

\[ d_{2optGM2}[k] <- \frac{ybar[k]}{xbar[k]} \times (0.5 - d_{1optGM2}[k] \times (1 - \rho_{oyx.sample}[k] \times (cy.sample[k] / cx)) \]

\[ k_{1optGMR1}[k] <- \frac{1 + (7/8) \lambda \times cx^2 - \lambda \rho_{oyx.sample}[k] \times cx \times cy.sample[k]}{1 + \lambda \times cy.sample[k]^2 \times (1 - \rho_{oyx.sample}[k]^2) + 2 \lambda \times cx^2 - 2 \lambda \rho_{oyx.sample}[k] \times cx \times cy.sample[k]} \]

\[ k_{2optGMR1}[k] <- \frac{ybar[k]}{xbar[k]} \times (0.5 - k_{1optGMR1}[k] \times (2 - \rho_{oyx.sample}[k] \times (cy.sample[k] / cx)) \]

\[ k_{1optGMR2}[k] <- \frac{1 + (23/8) \lambda \times (cx^2 - 2 \rho_{oyx.sample}[k] \times cx \times cy.sample[k])}{1 + \lambda \times (cy.sample[k]^2 \times (1 - \rho_{oyx.sample}[k]^2) + 6 \times cx^2 - 4 \times \rho_{oyx.sample}[k] \times cx \times cy.sample[k])} \]

\[ k_{2optGMR2}[k] <- \frac{ybar[k]}{xbar[k]} \times (0.5 - k_{1optGMR2}[k] \times (3 - \rho_{oyx.sample}[k] \times (cy.sample[k] / cx)) \]

# The non-RRT Mean Estimators

\[ \mu_R[k] <- ybar[k] \times (mu/xbar[k]) \]

\[ \mu_{Reg}[k] <- ybar[k] + beta.sample[k] \times (mu - xbar[k]) \]

\[ \mu_{ER}[k] <- ybar[k] \times \exp((mu - xbar[k]) / (mu + xbar[k])) \]

\[ \mu_{SRP}[k] <- ybar[k] \times (alpha.opt[k] \times \exp((mu - xbar[k]) / (mu + xbar[k])) + (1 - alpha.opt[k]) \times \exp((xbar[k] - mu) / (xbar[k] + mu)) \]

\[ \mu_{RReg}[k] <- k_{1opt}[k] \times ybar[k] + k_{2opt}[k] \times (mu - xbar[k]) \]

\[ \mu_{GK}[k] <- \frac{\lambda_{opt}[k] \times ybar[k] + 12 \times opt[k] \times (mu - xbar[k]) \times \exp((mu - xbar[k]) / (mu + xbar[k]))}{(mu + xbar[k])} \] # Grover & Kaur (2011)

\[ \mu_{GM1}[k] <- (d_{1optGM1}[k] \times ybar[k] \times (0.5 \times (mu/xbar[k] + xbar[k]/mu)) + \]

110
\[ d_{\text{optGM1}}[k] \cdot \exp\left(\frac{(\mu - x_{\text{bar}}[k])}{\mu + x_{\text{bar}}[k]}\right) \]

\[ \mu_{\text{GM2}}[k] \leftarrow \left( d_{\text{1optGM2}}[k] \cdot y_{\text{bar}}[k] \cdot \left(0.5 \cdot \frac{\mu}{x_{\text{bar}}[k]} + \frac{x_{\text{bar}}[k]}{\mu} \right)^2 \right) + \left( d_{\text{2optGM2}}[k] \cdot (\mu - x_{\text{bar}}[k]) \right) \cdot \exp\left(\frac{(\mu - x_{\text{bar}}[k])}{\mu + x_{\text{bar}}[k]}\right) \]

\[ \mu_{\text{GMR1}}[k] \leftarrow \left( k_{\text{1optGMR1}}[k] \cdot y_{\text{bar}}[k] \cdot \left(\frac{\mu}{x_{\text{bar}}[k]}\right) + k_{\text{2optGMR1}}[k] \right) \cdot \exp\left(\frac{(\mu - x_{\text{bar}}[k])}{\mu + x_{\text{bar}}[k]}\right) \]

\[ \mu_{\text{GMR2}}[k] \leftarrow \left( k_{\text{1optGMR2}}[k] \cdot y_{\text{bar}}[k] \cdot \left(\frac{\mu}{x_{\text{bar}}[k]}\right)^2 + k_{\text{2optGMR2}}[k] \right) \cdot \exp\left(\frac{(\mu - x_{\text{bar}}[k])}{\mu + x_{\text{bar}}[k]}\right) \]

# the empirical mean square errors over 10,000 trials

\[ \text{mse.ybar}[k] \leftarrow (y_{\text{bar}}[k] - \mu_{\text{y}})^2; \text{mse.ybar.E} \leftarrow \text{mean(mse.ybar)} \]

\[ \text{mse.muR}[k] \leftarrow (\mu_{\text{R}}[k] - \mu_{\text{y}})^2; \text{mse.muR.E} \leftarrow \text{mean(mse.muR)} \]

\[ \text{mse.muER}[k] \leftarrow (\mu_{\text{ER}}[k] - \mu_{\text{y}})^2; \text{mse.muER.E} \leftarrow \text{mean(mse.muER)} \]

\[ \text{mse.muReg}[k] \leftarrow (\mu_{\text{Reg}}[k] - \mu_{\text{y}})^2; \text{mse.muReg.E} \leftarrow \text{mean(mse.muReg)} \]

\[ \text{mse.muSRP}[k] \leftarrow (\mu_{\text{SRP}}[k] - \mu_{\text{y}})^2; \text{mse.muSRP.E} \leftarrow \text{mean(mse.muSRP)} \]

\[ \text{mse.muRReg}[k] \leftarrow (\mu_{\text{RReg}}[k] - \mu_{\text{y}})^2; \text{mse.muRReg.E} \leftarrow \text{mean(mse.muRReg)} \]

\[ \text{mse.muGK}[k] \leftarrow (\mu_{\text{GK}}[k] - \mu_{\text{y}})^2; \text{mse.muGK.E} \leftarrow \text{mean(mse.muGK)} \]

\[ \text{mse.muGM1}[k] \leftarrow (\mu_{\text{GM1}}[k] - \mu_{\text{y}})^2; \text{mse.muGM1.E} \leftarrow \text{mean(mse.muGM1)} \]

\[ \text{mse.muGM2}[k] \leftarrow (\mu_{\text{GM2}}[k] - \mu_{\text{y}})^2; \text{mse.muGM2.E} \leftarrow \text{mean(mse.muGM2)} \]

\[ \text{mse.muGMR1}[k] \leftarrow (\mu_{\text{GMR1}}[k] - \mu_{\text{y}})^2; \text{mse.muGMR1.E} \leftarrow \text{mean(mse.muGMR1)} \]

\[ \text{mse.muGMR2}[k] \leftarrow (\mu_{\text{GMR2}}[k] - \mu_{\text{y}})^2; \text{mse.muGMR2.E} \leftarrow \text{mean(mse.muGMR2)} \]

# The Empirical Mean Square Error for the non-RRT Mean Estimators

\[ \text{mse.E} \leftarrow \text{cbind(ybar=mse.ybar.E, muR=mse.muR.E, muReg=mse.muReg.E, muER=mse.muER.E, muSRP=mse.muSRP.E, muRReg=mse.muRReg.E, muGK=mse.muGK.E, muGM1=mse.muGM1.E, muGM2=mse.muGM2.E, muGMR1=mse.muGMR1.E, muGMR2=mse.muGMR2.E)} \]

# The Empirical Percent Relative Efficiency for the Mean estimators
pre.E<-100*(mse.ybar.E/mse.E)

# # # # T H E T H E O R E T I C A L P A R T # # # # #

mse.ybar<-lambda*muy^2*cy^2
mse.muR.th<-lambda*muy^2*(cy^2+cx^2-2*rhoyx*cy*cx)
mse.muReg.th<-lambda*muy^2*cy^2*(1-rhoyx^2)
mse.muER.th<-lambda*muy^2*(cy^2+(1/4)*cx^2-rhoyx*cx*cy)
mse.muSRP.th<-lambda*muy^2*(1-rhoyx^2)*cy^2
mse.muRReg.th<-muy^2*(1+(1/(-1+lambda*(rhoyx^2-1)*cy^2)))
tgkn<-lambda*muy^2*(lambda*(cx^4)-16*(rhoyx^2-1)*(lambda*(cx^2)-4)*cy^2)
tgkd<-64*(-1+lambda*(rhoyx^2-1)*cy^2)
mse.muGK.th<-tgkn/tgkd
mse.muGM1.th<-muy^2*(((1-0.25*lambda*cx^2)-(1+(3/8)*lambda*cx^2)^2)/
(1+lambda*cy^2*(1-rhoyx^2)+lambda*cx^2)))
mse.muGM2.th<-muy^2*(((1-0.25*lambda*cx^2)-(1+(7/8)*lambda*cx^2)^2)/
(1+lambda*cy^2*(1-rhoyx^2)+2*lambda*cx^2)))
mse.muGMR1.th<-muy^2*(((1-0.25*lambda*cx^2)-(1+lambda*((7/8)*cx^2
-rhoyx*cx*cy))^2)/(1+lambda*(2*cx^2+(1-rhoyx^2)*cy^2-2*rhoyx*cx*cy))))
mse.muGMR2.th<-muy^2*(((1-0.25*lambda*cx^2)-(1+lambda*((23/8)*cx^2
-2*rhoyx*cx*cy))^2)/(1+lambda*(6*cx^2+(1-rhoyx^2)*cy^2-4*rhoyx*cx*cy))))
# The Theoretical Mean Square Errors for the non-RRT Mean Estimators
mse.th.tb<-cbind(ybar=mse.ybar,MuR.th=mse.muR.th, MuReg.th=mse.muReg.th,
muER.th=mse.muER.th, muSRP.th=mse.muSRP.th, muRReg.th=mse.muRReg.th,
muGK=mse.muGK.th,muGM1=mse.muGM1.th,muGM2=mse.muGM2.th,
muGMR1=mse.muGMR1.th,muGMR2=mse.muGMR2.th)
# The Theort. Per. Rel. Eff. for the non-RRT Mean Estimators

\[
\text{pre.th.table}<-100*(\text{mse.ybar/mse.th.tb})
\]

# RANDOMIZED RESPONSE TECHNIQUE #

\[
\text{population.data}<-\text{read.table("population.data",header=TRUE)}
\]

\[
\text{scrambled.data}<-\text{read.table("scrambled.data",header=TRUE)}
\]

\[
X<-\text{population.data}\$X; \ Y<-\text{population.data}\$Y; \ S<-\text{scrambled.data}\$S
\]

\[
Z=Y+S
\]

\[
N<-\text{nrow(population.data)}; \ \rho_{zx}=\text{cor(X,Z)}; \ \rho_{yx}=\text{cor(X,Y)}
\]

\[
mux<-\text{mean(X)}; \ \muy<-\text{mean(Y)}; \ \muz<-\text{mean(Z)}; \ \text{varx}=\text{var(X)}; \ \text{vary}=\text{var(Y)};
\]

\[
\text{vars}=\text{var(S)}; \ \text{stdevx}=\text{sd(X)}; \ \text{stdevy}=\text{sd(Y)}; \ \text{stdevz}=\text{sd(Z)};
\]

\[
\beta_1x<-\text{skewness(X)}; \ \beta_2x<-\text{kurtosis(X)}
\]

### THE SIMULATION PART ###

\[
n<-100; \ b<-10000 \ # \ n=100, \ 200 \ \text{and} \ 500
\]

\[
f<-n/N; \ \lambda=(1-f)/n
\]

\[
xbar<-\text{numeric(b)}; \ zbar<-\text{numeric(b)}; \ \rho_{zx}.\text{sample}<-\text{numeric(b)};
\]

\[
\text{var.samplex}<-\text{numeric(b)}; \ \text{var.samplez}<-\text{numeric(b)}; \ \text{sd.samplex}<-\text{numeric(b)};
\]

\[
\text{sd.samplez}<-\text{numeric(b)}; \ \text{sd.samplezx}<-\text{numeric(b)}; \ \text{cz.sample}<-\text{numeric(b)};
\]

\[
\beta_1z.x<-\text{numeric(b)}; \ \muR<-\text{numeric(b)}; \ \muTR1<-\text{numeric(b)};
\]

\[
\muTR2<-\text{numeric(b)}; \ \muTR3<-\text{numeric(b)}; \ \muTR4<-\text{numeric(b)};
\]

\[
\muERR<-\text{numeric(b)}; \ \muReg<-\text{numeric(b)}; \ \muGRR<-\text{numeric(b)};
\]

\[
\muGER<-\text{numeric(b)}; \ \muGRR1<-\text{numeric(b)}; \ \muGRR2<-\text{numeric(b)};
\]
mse.zbar<-numeric(b); mse.muR<-numeric(b); mse.muTR1<-numeric(b);
mse.muTR2<-numeric(b); mse.muTR3<-numeric(b); mse.muTR4<-numeric(b);
mse.muERR<-numeric(b); mse.muReg<-numeric(b); mse.muGRR<-numeric(b);
mse.muGER<-numeric(b); mse.muGRR1<-numeric(b); mse.muGRR2<-numeric(b);
w1opt<-numeric(b); w2opt<-numeric(b); k1opt<-numeric(b);
k2opt<-numeric(b); d1optGRR1<-numeric(b); d2optGRR1<-numeric(b);
d1optGRR2<-numeric(b); d2optGRR2<-numeric(b);
for (k in 1:b){s<-sample(N,n,replace=TRUE); #print(s)
xbar[k]<-mean(X[s]); zbar[k]<-mean(Z[s]); var.samplez[k]<-var(Z[s])
var.samplex[k]<-var(X[s]); sd.samplez[k]<-sd(Z[s])
sd.samplezx[k]<-cov(X[s],Z[s]); cz.sample[k]<-sd.samplez[k]/zbar[k];
beta.sample.zx[k]<-sd.samplezx[k]/var.samplex[k]
k1opt[k]<-(1-lambda*cx^2)/(1-lambda*(cx^2-cz.sample[k]^2*(1-
rhozx.sample[k]^2)))
k2opt[k]<-(zbar[k]/xbar[k])*(1+k1opt[k]*((rhozx.sample[k]*
(cx*(cz.sample[k]/cx))-2))
w1opt[k]<-(1-lambda*(1/8)*cx^2)/((1+lambda*(cz.sample[k]^2
*(1-rhozx.sample[k]^2)))
w2opt[k]<-(zbar[k]/xbar[k])*(0.5-w1opt[k]*(1-rhozx.sample[k]
*(cz.sample[k]/cx)))
d1optGRR1[k]<-(1+(7/8)*lambda*cx^2-lambda*rhozx.sample[k]*cx
*(cz.sample[k])/(1+lambda*cz.sample[k]^2*(1-rhozx.sample[k]^2)
+2*lambda*cx^2-2*lambda*rhozx.sample[k]*cx*cz.sample[k])
d2optGRR1[k]<-(zbar[k]/xbar[k])*(0.5-d1optGRR1[k]*(2-rhozx.sample[k]
*(cz.sample[k]/cx)))
d1optGRR2[k]<-(1+(23/8)*lambda*cx^2-2*lambda*rhozx.sample[k]*cx
\[\frac{cz\text{. sample}[k]}{(1+\lambda \cdot \text{cz\text{. sample}[k]}^2(1-\rho\text{zx}\text{. sample}[k])^2 + 6 \lambda \cdot cx^2 - 4 \lambda \cdot \rho\text{zx}\text{. sample}[k] \cdot cx \cdot \text{cz\text{. sample}[k]})}
\]

d2optGRR2[k] \leftarrow \frac{zbar[k]}{xbar[k]} \cdot (0.5 - d1optGRR2[k] \cdot (3 - \rho\text{zx}\text{. sample}[k]) \cdot \frac{cz\text{. sample}[k]}{cx}))

# RRT Mean Estimators

\[
\mu_{R}[k] \leftarrow \frac{zbar[k]}{xbar[k]} \cdot (\mu_{x}/xbar[k])
\]

\[
\mu_{TR1}[k] \leftarrow \frac{zbar[k]}{xbar[k]} \cdot (\frac{\mu_{x} + \beta_{1}x}{xbar[k] + \beta_{1}x})
\]

\[
\mu_{TR2}[k] \leftarrow \frac{zbar[k]}{xbar[k]} \cdot (\frac{\mu_{x} + \beta_{2}x}{xbar[k] + \beta_{2}x})
\]

\[
\mu_{TR3}[k] \leftarrow \frac{zbar[k]}{xbar[k]} \cdot (\frac{\beta_{1}x \cdot \mu_{x} + \beta_{2}x}{\beta_{1}x \cdot xbar[k] + \beta_{2}x})
\]

\[
\mu_{TR4}[k] \leftarrow \frac{zbar[k]}{xbar[k]} \cdot (\frac{\beta_{2}x \cdot \mu_{x} + \beta_{1}x}{\beta_{2}x \cdot xbar[k] + \beta_{1}x})
\]

\[
\mu_{Reg}[k] \leftarrow \frac{zbar[k]}{xbar[k]} + \beta_{\text{sample.zx}[k]} \cdot (\mu_{x} - xbar[k])
\]

\[
\mu_{ERR}[k] \leftarrow \frac{zbar[k]}{xbar[k]} \cdot \exp\left(\frac{\mu_{x} - xbar[k]}{\mu_{x} + xbar[k]}\right)
\]

\[
\mu_{GRR}[k] \leftarrow \frac{k_{1\text{opt}}[k] \cdot zbar[k] + k_{2\text{opt}}[k]}{xbar[k]} \cdot \exp\left(\frac{\mu_{x} - xbar[k]}{\mu_{x} + xbar[k]}\right)\quad \text{# Zatezalo et al. (2016)}
\]

\[
\mu_{GRR1}[k] \leftarrow \frac{d_{1\text{optGRR1}[k]} \cdot zbar[k] + d_{2\text{optGRR1}[k]} \cdot \mu_{x}}{xbar[k]} \cdot \exp\left(\frac{\mu_{x} - xbar[k]}{\mu_{x} + xbar[k]}\right)
\]

\[
\mu_{GRR2}[k] \leftarrow \frac{d_{1\text{optGRR2}[k]} \cdot zbar[k] + d_{2\text{optGRR2}[k]} \cdot (\mu_{x} - xbar[k])}{xbar[k]} \cdot \exp\left(\frac{\mu_{x} - xbar[k]}{\mu_{x} + xbar[k]}\right)
\]

\[
\text{mse.zbar}[k] \leftarrow (zbar[k] - \mu_{y})^2; \text{mse.zbar.E} \leftarrow \text{mean}(\text{mse.zbar})
\]

\[
\text{mse.muR}[k] \leftarrow (\mu_{R}[k] - \mu_{y})^2; \text{mse.muR.E} \leftarrow \text{mean}(\text{mse.muR})
\]

\[
\text{mse.muTR1}[k] \leftarrow (\mu_{TR1}[k] - \mu_{y})^2; \text{mse.muTR1.E} \leftarrow \text{mean}(\text{mse.muTR1})
\]

\[
\text{mse.muTR2}[k] \leftarrow (\mu_{TR2}[k] - \mu_{y})^2; \text{mse.muTR2.E} \leftarrow \text{mean}(\text{mse.muTR2})
\]

\[
\text{mse.muTR3}[k] \leftarrow (\mu_{TR3}[k] - \mu_{y})^2; \text{mse.muTR3.E} \leftarrow \text{mean}(\text{mse.muTR3})
\]

\[
\text{mse.muTR4}[k] \leftarrow (\mu_{TR4}[k] - \mu_{y})^2; \text{mse.muTR4.E} \leftarrow \text{mean}(\text{mse.muTR4})
\]

\[
\text{mse.muReg}[k] \leftarrow (\mu_{Reg}[k] - \mu_{y})^2; \text{mse.muReg.E} \leftarrow \text{mean}(\text{mse.muReg})
\]
mse.muERR[k]<-(mu[ERR[k]]-muy)^2; mse.muERR.E<-mean(mse.muERR)
mse.muGRR[k]<-(mu[GRR[k]]-muy)^2; mse.muGRR.E<-mean(mse.muGRR)
mse.muGER[k]<-(mu[GER[k]]-muy)^2; mse.muGER.E<-mean(mse.muGER)
mse.muGRR[k]<-(muGRR[k]-muy)^2; mse.muGRR1.E<-mean(mse.muGRR)
mse.muGRR2[k]<-(muGRR2[k]-muy)^2; mse.muGRR2.E<-mean(mse.muGRR2)
}

#The Empirical Mean Square Errors for the RRT Mean Estimators
mse.E<-cbind(zbar=mse.zbar.E,muR=mse.muR.E,muTR1=mse.muTR1.E,
muTR2=mse.muTR2.E,muTR3=mse.muTR3.E,muTR4=mse.muTR4.E,
muReg=mse.muReg.E,muER=mse.muER.E,muGRR=mse.muGRR.E,
muGER=mse.muGER.E,muGRR1=mse.muGRR1.E,muGRR2=mse.muGRR2.E)

# The Emp. Perc. Rel. Efficiency for the RRT Mean Estimators
pre.E<-100*(mse.zbar.E/mse.E)

# T H E T H E O R E T I C A L P A R T #

mse.zbar.th<-lambda*(vary+vars)
mse.muR.th<-lambda*muy^2*(cz^2+cx^2-2*rhozx*cz*cx)
mse.muTR1.th<-lambda*muy^2*(cz^2+(mux/(mux+betax))^2*cx^2
-2*(mux/(mux+betax))*rhozx*cz*cx)
mse.muTR2.th<-lambda*muy^2*(cz^2+(mux/(mux+beta2x))^2*cx^2
-2*(mux/(mux+beta2x))*rhozx*cz*cx)
mse.muTR3.th<-lambda*muy^2*(cz^2+((betax*mux)/(betax*mux+beta2x))^2
*cz^2-2*((betax*mux)/(betax*mux+beta2x))*rhozx*cz*cx)
mse.muTR4.th<-lambda*muy^2*(cz^2+((beta2x*mux)/(beta2x*mux+betax))^2
*cx^2-2*((beta2x*mux)/(beta2x*mux+betax))*rhozx*cz*cx)
mse.muReg.th<-lambda*muy^2*cz^2*(1-rhozx^2)
mse.muER.th<-lambda*muy^2*(cz^2-rhozx*cz*cx+0.25*cx^2)
mse.muGRR.th<-muy^2*(lambda*cz^2*(1-rhozx^2)*(1-lambda*cx^2))/((lambda *cz^2*(1-rhozx^2)+(1-lambda*cx^2)))
mse.muGER.th<-muy^2*((1-0.25*lambda*cx^2)-((1-lambda*(1/8)*cx^2)^2/(1+lambda*(1-rhozx^2)*cz^2)))
mse.muGRR1.th<-muy^2*(((1-0.25*lambda*cx^2)-((1-lambda*(1/8)*cx^2)^2/(1+lambda*(2*cx^2+(1-rhozx^2)*cz^2-2*rhozx*cz*cx)))))
mse.muGRR2.th<-muy^2*(((1-0.25*lambda*cx^2)-((1-lambda*drhozx2*cz^2)+(1-rhozx*cz*cx))^2/(1+lambda*(6*cx^2+(1-rhozx^2)*cz^2-4*rhozx*cz*cx))))

# the theoretical mean square for the RRT mean estimators
mse.th<-cbind(zbar.th=mse.zbar.th,muR.th=mse.muR.th,
muTR1.th=mse.muTR1.th,muTR2.th=mse.muTR2.th,muTR3.th=mse.muTR3.th,
muTR4.th=mse.muTR4.th,muReg.th=mse.muReg.th,muER.th=mse.muER.th,
muGRR.th=mse.muGRR.th,muGER.th=mse.muGER.th,muGRR1.th=mse.muGRR1.th,
muGRR2.th=mse.muGRR2.th)

# The Theor. Perc. Rel. Eff. for the RRT Mean Estimators
pre.th.table<-100*(mse.zbar.th/mse.th)