National College of Business Administration & Economics Lahore

GENERALIZED EXPONENTIAL ESTIMATORS FOR POPULATION VARIANCE USING RANDOMIZED RESPONSE MODEL

BY

YASIR NAWAZ

MASTER OF PHILOSOPHY
IN
STATISTICS

FEBRUARY, 2022

NATIONAL COLLEGE OF BUSINESS ADMINISTRATION & ECONOMICS

GENERALIZED EXPONENTIAL ESTIMATORS FOR POPULATION VARIANCE USING RANDOMIZED RESPONSE MODEL

BY

YASIR NAWAZ

A dissertation submitted to
Faculty of Social Sciences

In Partial Fulfillment oft he
Requirements for the Degree of

MASTER OF PHILOSOPHY
IN
STATISTICS

FEBRUARY, 2022

بسم الله الرحمن الرحيم

In the name of ALLAH,
The Most Beneficial,
The Most Merciful,

NATIONAL COLLEGE OF BUSINESS ADMINISTRATION & ECONOMICS LAHORE

GENERALIZED EXPONENTIAL ESTIMATORS FOR POPULATION VARIANCE USING RANDOMIZED RESPONSE MODEL

BY
YASIR NAWAZ

A dissertation submitted to Faculty of Social Sciences, in partial fulfillment of the requirements for the degree of

MASTER OF PHILOSOPHY IN STATISTICS

Dissertation Committee:

____
Chairman

____
Member

____
Member

____
Director Research
National College of Business
Administration and Economics

DECLARATION

It is to declare that this research work has not been submitted for obtaining similar degree from any other university/college.

YASIR NAWAZ
FEBRUARY, 2022

DEDICATION

Dedicated
To

ACKNOWLEDGEMENT

Unrestrained humblest thanks to Allah Almighty alone, the creator of entire universe who is worthy of all the praises and can convert the impossible to possible. All love and respect to our beloved holy prophet Hazrat Muhammad (Peace be Upon Him), whose revelations are the source of guidance for us in both the worlds.

I would like to extend my gratitude to the National College of Business Administration and Economics for providing me such environment and resources, which enabled me to complete this research successfully.

I feel great honor and pleasure in expressing my profound and cordial gratitude to my supervisor ***Prof. Dr. Muhammad Hanif*** for academic support in whole process from the beginning to the end of this thesis. Thanks to all my teachers in the college for giving me the different concepts throughout my studies.

I am incredibly grateful to my Co-supervisor ***Dr. Muhammad Nouman Qureshi***, for his kind advices and valuable help whenever I needed throughout my M. Phil period. I am indeed thankful for the support I have received from all my class fellows.

Especially, I have no words to pay gratitude to my parents, siblings and my family whose affection, guidance and continuous encouragement enabled me to complete this research work. In the end, I pray to Almighty Allah to give me wisdom and strength to use this knowledge the way He wants.

RESEARCH COMPLETION CERTIFICATE

Certified that the research work contained in this thesis entitled **"Generalized Exponential Estimators for Population Variance using Randomized Response Model"** has been carried out and completed by **Yasir Nawaz** under my supervision during his **M.Phil. Statistics** Programme.

(Dr. Muhammad Hanif)
Supervisor

SUMMARY

In this Dissertation, we have proposed generalized exponential estimators using single and two auxiliary variables for variance estimation of sensitive study variable. The first Chapter includes the introduction and background of the survey sampling. Later on we described the use of the auxiliary information, and variance estimation. We also mentioned the statement of the problem, objectives and scope of the study. Chapter 2 comprises the literature review about the use of the auxiliary information for sensitive study variable for variance estimation. Randomized response models and some classical estimators for sensitive study variable with single and two auxiliary variables are discussed in Chapter 3. The proposed generalized exponential estimator with single and two auxiliary variables are proposed in Chapter 4 and Chapter 5 respectively. Mathematical expressions for approximate bias and mean square error of proposed estimators are also derived and discussed in the same Chapters. We also obtained many families of the proposed estimators using different choices of the constants. Mathematical comparisons of the generalized exponential estimators are made with the sample variance, ratio and exponential ratio estimators based on sensitive study variable. Extensive numerical study is conducted to support the mathematical results using real and artificial data-sets to evaluate the performance of the proposed estimators. Conclusion and closing remarks on the research are given in Chapter 6.

LIST OF TABLES

LIST OF FIGURES

TABLE OF CONTENTS

CHAPTER 1:

INTRODUCTION

1.1 SENSITIVE SURVEYS

Survey sampling helps the researcher to arrive at the generalization of results by studying the small part of population. The basic aim of a survey is to collect the information on the relevant issues of the population under study using appropriate sampling techniques. So, the researchers always inclined to gather the information through distinctive approaches of data collection from some or all the elements chosen in the sample."

Sometimes, data having the characteristics which are personal or illegal in nature such as drunken driving, criminal behavior, reuse of infected syringes, carrying an HIV virus, etc., and the queries on such appearances are known to be threatening or sensitive questions."

In many surveys based on sensitive issues, direct questioning is often not beneficial to get the accurate response and may cause the refusals or wrong reporting from the respondent. Asking questions on sensitive issues is actually aboutrequesting the respondent to provide information they are not willing to answer."

In many convoluted surveys, two kinds of behavior may be expected from the respondent to answer the sensitive questions; non-response or scrambled response. Non-response usually occurs due to the high rate of refusals from the respondents. On the contrary, scrambled response or misreporting arises when the respondent decides to answer but deliver incorrect or false feedback. Such response errors normally cause a large amount of bias to estimate the parameters on sensitive questions.

1.2 VARIANCE ESTIMATION

Variance estimation of the targeted population is of the major importance in survey sampling. It is assumed to be a key indicator to monitor the process in control. It is considered to be helpful to estimate the overall variability and to reach on valid conclusion by using the variation in the samples. Variance estimation is commonly used to measure the variation in climate change. Different qualities of seeds and soil fertility in agricultural research generally used variance estimation. At the same time, variance estimator is also useful to build a valid confidence interval for the estimated parameter of the study.

A general class of quadratic unbiased estimators for the population variance was first proposed by Liu (1974). Later on, Das and Tripathi (1978) gave a class of estimators to estimate population variance. Ratio and regression estimators for population variance were first discussed by Isaki (1983). So far, many researchers have been worked out on variance estimation in different modes of study and the development is continue."

1.3 AUXILIARY INFORMATION

Additional information is usually called the prior information or the auxiliary information which may be obtained from the previous experiments such as past records, surveys, censuses etc. In sampling theory, supportingvariable is widely used to increase the efficiency and decrease the error when the study variable is correlated with the supporting variable(s). Nowadays, almost all the researchers inclined to obtain the information of many variables instead of just study variable. The use of the benchmark variablealong with the primary variable of interest is very usual to obtain precise estimates for the population parameters such as the mean, median, total, variance, proportion etc. The ratio, product and regression type estimators are the most commonly used estimation methods. Sami-ud-Din and Hanif (2007) presented a variety of procedures to use the auxiliary information in survey sampling.

The use of supplementary information in survey sampling was well described by Tripathi (1970, 1973, 1976). According to him:

i) *"It may be used for planning of a survey or at the time of sample design e.g., from information about variable, the population could be stratified etc.*

ii) *It may be utilized to estimate the population parameters e.g., the ratio, regression, difference, product estimators etc.*

iii) *It may be utilized for the selection of sample e.g., for calculation of proportional to size probabilities.*

iv) *It may also be utilized combining any of above two."*

1.4 SENSITIVE STUDY VARIABLE

There are many situations in surveys when our main research variable is sensitive in nature and it is difficult to get the true response from the respondents directly. Some sensitive study variables are Illegal income, Alcoholism, Homosexual activities, Drug addition,Doping usage, Political views, Spot fixing, Mobbing, Gambling, Abortion, Tax evasion andSexual and Physical abuse. To estimate the sensitive study variable, randomized response technique is considered to be very useful and efficient."

1.5 RANDOMIZED RESPONSE TECHNIQUE

It is a very common practice that almost all the respondents do not provide a true answer whenever the sensitive questions asked by the researchers. Such situations create many problems to reach a valid conclusion about the problem. "Warner (1965) first gave the notion of Randomized Response Technique (*RRT*) to get the response while protecting the privacy of the respondent. With the help of *RRT* scramble responses become unscrambled. The consent of *RRT* is for respondents to answer sensitive issues while maintaining privacy. It generally delivers privacy to the respondents with the help of anunplanned relation between the character's questions and response."

1.6 RESEARCH QUESTIONS

- Will the proposed estimators beat the existing estimators mathematically and numerically?
- Will the proposed estimators increase the precision and decrease the error of estimation?

1.7 STATEMENT OF PROBLEM

From the existing literature, it is observed that a lot of work has been done on the estimation of population mean for sensitive study variable using non-sensitive auxiliary variable(s) but few researchers work is available on the variance estimation using randomized response technique. It is also notice that no contribution has done using exponential estimators along with the randomized response models (RRMs) having one or more than one auxiliary variable for the estimation of population variance having sensitive study variable. In this research, our main concern is to propose some exponential estimators for population variance having sensitive research variable using of one and two benchmark variables under *SRS* design.

1.8 LIMITATION OF THE STUDY

- The proposed estimators are only suitable for populationvariance estimation of a sensitive study variable in the presence of non-sensitive auxiliary variable(s) having reasonable amount of correlation coefficient.
- The proposed estimators would be better than the existing estimators under certain mathematical conditions.
- The proposed estimators will only perform well for sensitive study variable if the RRMs will be used for simulation and real-data application.

1.9 OBJECTIVE OF THE STUDY

The major objectives of the study are to:

- Suggest the exponential estimators for variance estimation of sensitive study variable in the presence of oneor two auxiliary variables using randomized response techniques.
- Derive the expressions of approximate bias and mean square error (*MSE*) of the aforementioned estimators by using Taylor and exponent series.
- Obtained the conditions in which our proposed exponential estimators will perform better than the competing estimators available in literature.
- Assess the performance of proposed estimators using simulation and real data application.

1.10 RESEARCH METHODOLOGY

- In this research we have proposed some exponential estimators for sensitive study variable using auxiliary variables for population variance using well-known Taylor and exponential series."
- A widespread simulation and real-data illustration are conducted using R-software to judge the implementation of the proposed exponential estimators over the competing estimators.

CHAPTER 2:

LITERATURE REVIEW

2.1 INTRODUCTION

"The use of auxiliary information is as old as the history of survey sampling. The auxiliary information is commonly used to obtain the improved designs and to get the high precision for the estimation of unknown population parameters. "In this chapter, we have discussed the use of supplementary information for variance estimation of sensitive study variable at estimation stage. A bulk of literature about the use of auxiliary information is available but we mentioned the most relevant to our study."

Warner (1965) first gave the notion of quantitative randomized response model to estimate the mean of sensitive study variable. "Later on, a verity of research has been carried out in this area to get the response on sensitive questions. After the initial work on randomized response model, Greenberg *et al.* (1969, 1971) and Warner (1971) further enhanced the technique."

"Many researchers worked on the randomized response models and decided that the deliberate recording of misleading answers is the first and foremost cause of measurement error resulting in serious response bias. Randomized response models are used as a device to reduce the response bias by providing the confidentiality defense to the respondents. Eichhron and Hayre (1983), Gupta *et al.* (2002, 2010), Gupta and Shabbir (2004), Arnab and Dorffer (2006), Wu et al. (2008), Saha (2008) and Perri (2008) considered to the authors who estimated mean of sensitive study variable without using the supporting variable. Various authors have developed many efficient estimators using different randomized response models for different situations under different sampling procedures for the estimation of sensitive study variable using auxiliary information. The most prominent are: Sousa *et al.* (2010) used the auxiliary information to estimate the sensitive study variable and proposed classical ratio estimator. Empirical and theoretical study showed better performance of the proposed estimator".

Gupta *et al.* (2012) "used the auxiliary variable to estimate the scrambling responses and proposed regression-type estimators. Mathematical properties of the properties were also derived to assess the performance of the proposed estimator.

An extensive numerical study was also conducted using the additive model for randomized responses".

Koyuncu*et al.* (2014) used the RRT to proposed exponential-type estimators with single and two auxiliary variables to improve the efficiency of the mean estimator. "The expressions of MSE and bias were obtained up the first order for both proposed estimators. The gain in efficiency revealed that the proposed estimators were more efficient than the existing estimators.

"Kalucha *et al.* (2014) and Gupta *et al.* (2015) respectively proposed additive ratio and regression estimators for finite population mean using optional *RRT*. The comparisons, both theoretically and numerically were carried out with the ordinary mean estimator proposed by Gupta *et al.* (2010)."

"Mushtaq *et al.* (2017) proposed a family of estimators of a sensitive study variable using auxiliary information under stratified random sampling design. The equations of MSE and bias are derived up to the first order. Empirical and theoretical results via simulation study show that the proposed class of estimators is more efficient than the existing estimators given by Sousa *et al.*(2014)."

"Ozgul and Cingi (2017) developed an improved estimator for mean estimation of the sensitive study variable using partial additive randomized response models in the presence of two non-sensitive auxiliary variables. The MSE of the developed estimator is derived and compared with the estimators suggested by Sousa *et al.* (2010), Diana and Perri (2011) and Gupta *et al.* (2012)."

"Khalil *et al.* (2018) proposed generalized estimator for sensitive study variable using benchmark variable with measurement error under simple and stratified sampling. The proposed estimators were compared with the mean and ratio estimators mathematically and numerically. Numerical results showed that the proposed estimators are more efficient than the competing estimator for both simple and stratified sampling schemes".

"Grover and Kaur (2018) proposed an efficient scrambled estimator for population mean using general linear transformation of non-sensitive study variable. The efficiency comparisons with the exiting estimators were also carried out both numerically and theoretically. It is found that the proposed scrambled estimator is more efficient than the existing estimators under certain conditions."

“Khalil *et al.* (2019) improve the generalized estimator proposed by Khalil et al. (2018) by using the optional RRT models for the estimation of sensitive study variable. The results obtained both theoretically and empirically showed the better performance of the proposed estimator.”

Zhang *et al.* (2019) introduced a geometric mean ratio estimator for finite population mean using optional RRT models. The results showed that the geometric mean estimator is more efficient than the ordinary RRT mean estimator for the situation in which the coefficient of correlation between the study variable and the auxiliary variable is greater than 0.5. Authors also showed that the efficiency of the proposed geometric mean ratio estimator and additive ratio estimator is same but has least bias.”

“Saleem *et al.* (2019) used two-phase sampling to proposed regression-cum-exponential estimator for the estimation of mean of sensitive study variable. The mathematical expressions of bias and MSE are obtained and conducted a numerical study to judge the performance of the proposed estimator.”

Sanaullah *et al.* (2019) “used the non-sensitive concomitant variable to propose difference-cum-exponential and generalized difference-cum-exponential estimators for mean estimation of scrambled responses. Authors proved that the proposed estimators were more efficient than the ratio estimator proposed by Sousa *et al.* (2010), regression-type proposed by Gupta *et al.* (2012) and exponential-type estimator proposed by Koyuncu*et al.* (2014).”

Shahzad *et al.* (2019) “considered the family of estimators proposed by Koyuncu and Kadilar (2009) and suggested a regression-type estimator for population mean of the sensitive study variable using auxiliary information under ranked set sampling.An extensive simulation study was also conducted to demonstrate the PRE of the proposed estimators over the challenging estimators.

“Gupta *et al.* (2020) used the auxiliary information and randomized response model having two scrambling variables to propose the ratio and generalized ratio estimator for population variance. Mathematical and numerical comparisons showed the superiority of the proposed estimators over the sample variance estimator.”

“Waseem *et al.* (2020, 2021) used single and two auxiliary variables to suggest the generalized exponential-type estimator for the estimation of mean of scrambling responses in *SRS* design. A new optional randomized response model has been used for scrambled responses. Authors also discussed the special cases of the suggested estimator with privacy protection. Numerical studies were also conducted to support the performance of the proposed estimators using simulation and real data illustration.

Sajida Balqees (2020) “proposed classical estimators, modified forms of classical estimators, a class of exponential estimators and generalized estimator for mean estimation of sensitive study variable in the presence of non-sensitive study variable using systematic sampling scheme. An extensive numerical study was also carried out using real and artificial populations.”

Muhammad Yousaf (2020) “suggested a class of ln-type mean estimators for the estimation of sensitive study variable using the information of variance of the auxiliary variable under simple and stratified sampling. The proposed classes of estimators were compared with the ordinary mean and Sousa *et al.* (2010 and 2014) estimators empirically and theoretically. The results showed better performance of the proposed class of estimators than the existing estimators for both real and artificial populations.”

Mehwish Anwar (2020) suggested the generalized estimators for sensitive study variable using the non-sensitive auxiliary attribute in simple and stratified random sampling. The expressions of bias and MSE of the proposed estimator were also derived and compared with the classical estimators (ordinary mean, classical ratio, regression, exponential ratio and exponential product estimators). The results of simulation and empirical study showed that the generalized estimators are more efficient than the classical estimators under both simple and stratified random sampling design.

Quratulain Rana (2021) proposed the generalized estimators for confidential study variable using auxiliary attribute in two-phase simple and stratified sampling. The expressions of bias and MSE of the proposed estimator were also derived and compared with the mean, ratio, and exponential estimators. The results of numerical study revealed that the generalized estimators performedwell than the existing estimators under two-phase simple and stratified sampling designs”.

CHAPTER 3:

RANDOMIZED RESPONSE MODELS AND ASSOCIATED ESTIMATORS

3.1 INTRODUCTION

In this chapter, we have discussed randomized response techniques and randomized response models along with their basic properties (means and variances). Some exponential estimators are also mentioned for non-sensitive study and the auxiliary variables for population mean and variance. The expressions of bias and MSE are also given with their respective estimators.

3.2 RANDOMIZED RESPONSE TECHNIQUE

The randomized response technique was first introduced by Warner (1965) in a very simple and significant way. Many variations of that original technique have since been introduced by different authors that include Greenberg *et al.* (1969, 1971), Eichhorn and Hayre (1983), Gupta *et al.* (2002), Gupta *et al.* (2006), Gupta *et al.* (2010) and Huang (2010). "A quick search will reveal that even for surveys dealing with non-sensitive questions, a response rate of around 40 - 50 percent is considered excellent. Warner's (1965) original approach dealt with a binary response situation. This was done by asking the research question directly to some of the randomly selected respondents. Other respondents are asked the indirect version of the same question. For example, some of the respondents face the question, "Did you file a correct income tax return last year?" while others face the question, "Did you file an incorrect income tax return last year?". The researcher does not know which respondent faced which question. Essentially the responses are randomized (or scrambled). These scrambled responses can later be unscrambled at aggregate level but not at the individual level. That is, one can estimate what proportion of tax payers cheated on taxes last year but there is no way of knowing who cheated and who did not. 5 The RRT models can be characterized in various ways." An RRT model is called quantitative, binary or multi-category depending on the type of response the question carries. For quantitative RRT models, interest is in estimating the mean of the variable of interest. For binary RRT models we are interested in estimating the prevalence of certain behaviors in the population. An RRT model can also be called Full RRT, Partial RRT depending on whether all or only some of the respondents provide scrambled responses, and depending on whether a respondent has been forced to provide a true response, as in Mangat and Singh (1990) two-stage model, or the choice is left to the respondent as in Gupta (2002) optional model. The RRT models can also be classified as additive, multiplicative or generalized depending on whether we use additive scrambling, multiplicative scrambling, or more general scrambling.

3.3 RANDOMIZED RESPONSE MODELS

"The RRT is an advance technique that aims to get the almost accurate answers to a sensitive question that respondents might be hesitant to answer truthfully, such as use of drugs, sexual practice, illegal earning, incidence of acts of domestic violence, or in collusion with terrorists or extremists nationally or internationally and many others are included in surveys of human populations. Warner (1965) was the first to suggest an ingenious method of counteracting fears in response to sensitive questions and called it RRT. He developed an interviewing procedure designed to reduce or eliminate the bias. Later, several other authors have suggested various alternative randomized response strategies."

Some well-known models are discussed below:

3.3.1 Additive Model

The additive model first given by Pollock and Beck (1976) where the related response is obtained by adding the scrambling variable in the main survey variable

$$Z = Y + S, \tag{1}$$

"where Z is the reported response, Y be the variable of interest and S is the scrambling variable. Further, it is assumed $S \sim N\,(0,\ 1\%{*}\,S_x)$.

The mean and variance of sensitive study variable is:

$$E(Z) = E(Y).$$

or,

$$\mu_z = \mu_x.$$

and

$$Var(Z) = Var(Y) + Var(S).$$

or,

$$S_z^2 = S_y^2 + S_s^2.$$

3.3.2 Multiplicative Model

Eichhron and Hayre (1983) first gave the idea of multiplicative model to get the scrambled response. The model is defined as

$$Z = Y * T, \tag{2}$$

where T is the scrambling variable and Z and Y are as usual sensitive study and the non-sensitive auxiliary variables respectively. It is assumed that $E(T) = 1$

The mean and variance of sensitive study variable is

$$E(Z) = E(Y).$$

or,

$$\mu_z = \mu_x.$$

and

$$Var(Z) = Var(Y).Var(T).$$

or,

$$S_z^2 = S_y^2.S_T^2.$$

3.3.3 General Linear Model

Diana and Perri (2011) proposed a more general linear model using two scrambling variables. The model is defined by adding and multiplying two different scrambling variables simultaneously with the confidential study variable as

$$Z = TY + S. \tag{3}$$

It is common to assume that the scrambling variables S and T follow normal distribution within means zero and one and variances S_s^2 and S_T^2 and respectively. Note that the models presented in (1) and (2) are special cases of the model given in (3).

The mean and variance of general linear model are respectively given by

$$E(Z)=E(Y).$$

or,

$$\mu_z=\mu_x.$$

and

$$S_z^2=S_y^2+S_s^2\left(\mu_y^2+S_y^2\right)+S_T^2.$$

3.4 SOME EXISTING EXPONENTIAL ESTIMATORS

In the past, many authors have proposed exponential estimators for the estimation of population parametershaving non-sensitive study variable. In this section, we have mentioned some existing estimators for population mean and variance of sensitive study variable in *SRS* design.The usual unbiased variance estimator along with the variance is defined as

$$t_y=s_z^2-S_s^2=\frac{1}{(n-1)}\sum_{i=1}^{n}\left(z_i-\bar{z}\right)^2-S_s^2.$$

and

$$\operatorname{var}\left(t_y\right)=\theta\lambda_{40}^{*}.$$

Classical ratio estimator for the estimation of population variance of confidential variable is defined as

$$t_r=\left(s_z^2-S_s^2\right)\left(\frac{S_x^2}{s_x^2}\right).$$

The expressions of approximate Bias and *MSE* of t_r are

$$Bias\left(t_r\right)=\theta S_z^2\left[\lambda_{04}^{*}-\lambda_{22}^{*}\right].$$

and

$$MSE\left(t_r\right)=\theta S_z^4\left[\lambda_{40}^{*}+\lambda_{04}^{*}-2\lambda_{22}^{*}\right].$$

The ratio estimator having two auxiliary variables for estimating the mean of population variance of sensitive study variable is defined as

$$t_r^* = \left(s_z^2 - S_s^2\right)\left(\frac{S_x^2}{s_x^2}\right)\left(\frac{S_u^2}{s_u^2}\right).$$

The expressions of approximate Bias and *MSE* of t_r are

$$Bias\left(t_r^*\right) \approx \theta S_y^2\left[\lambda_{040}^* + \lambda_{004}^*\right] - \theta S_z^2\left[\lambda_{022}^* + \lambda_{202}^*\right].$$

and

$$MSE\left(t_r^*\right) = \theta S_z^4 \lambda_{400}^* + \theta S_y^4\left[\lambda_{040}^* + \lambda_{004}^* + 2\lambda_{022}^* - 2\Re_{zy}\left(\lambda_{220}^* + \lambda_{202}^*\right)\right].$$

Mehwish Omer (2020) suggested exponential type ratio estimator for estimating the population variance of a confidential variable using single benchmark variable is defined as

$$t_{e11} = s_z^2 exp\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right).$$

The equations of bias and *MSE* of $t_{er}$are respectively given by

$$Bias\left(t_{e11}\right) = \frac{1}{8}\theta S_z^2\left[3\lambda_{04}^* - 4\lambda_{22}^*\right].$$

and

$$MSE\left(t_{e11}\right) = \theta S_z^4\left[\lambda_{40}^* + \frac{1}{4}\lambda_{04}^* - \lambda_{22}^*\right].$$

The exponential ratio estimator for estimating the population variance of a sensitive study variable using two auxiliary variables is defined as

$$t_{e11}^* = s_z^2 exp\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2} + \frac{S_u^2 - s_u^2}{S_u^2 + s_u^2}\right).$$

The expressions of bias and *MSE* of $t_{er}$are respectively given by

$$Bias\left(t_{e11}^{*}\right)=\frac{1}{4}\theta S_{y}^{2}\left[\lambda_{040}^{*}+\lambda_{004}^{*}+\lambda_{022}^{*}\right]-\frac{1}{2}\theta S_{z}^{2}\left[\lambda_{220}^{*}+\lambda_{202}^{*}\right].$$

and

$$MSE\left(t_{e11}^{*}\right)=\theta S_{z}^{4}\lambda_{400}^{*}+\theta\frac{1}{4}S_{y}^{4}\left[\lambda_{040}^{*}+\lambda_{004}^{*}+2\lambda_{022}^{*}-4\Re_{zy}\left(\lambda_{220}^{*}+\lambda_{202}^{*}\right)\right].$$

The exponential ratio estimatorfirst proposed by Bahl and Tuteja (1991) are mentioned for sensitive study variable using single auxiliary variable for the estimation of population mean under *SRS* design. The estimator with the expressions of bias and *MSE* is defined as

$$t_{er}=\bar{z}exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right).$$

$$Bias\left(t_{er}\right)=\frac{1}{8}\theta\bar{Y}\left[3C_{x}^{2}-4\rho_{zx}C_{y}C_{x}\right].$$

and

$$MSE\left(t_{er}\right)=\theta\bar{Y}^{2}\left[C_{z}^{2}+\frac{1}{4}C_{x}^{2}-\rho_{zx}C_{y}C_{x}\right].$$

Waseem *et al.* (2020) proposed generalized exponential estimator for estimating the population mean of confidential study variable using supporting variable. The generalized exponential estimator along with the expressions of bias and MSE are defined as

$$t_{erh}=\bar{z}exp\left(\alpha\frac{\bar{X}^{\left(\frac{1}{h}\right)}-\bar{x}^{\left(\frac{1}{h}\right)}}{\bar{X}^{\left(\frac{1}{h}\right)}+(a-1)\bar{x}^{\left(\frac{1}{h}\right)}}\right).$$

$$Bias\left(t_{erh}\right)=\theta\bar{Y}\left[\frac{\alpha\bar{Z}C_{x}^{2}}{ah^{2}}-\frac{\alpha\bar{Z}\rho_{yx}C_{y}C_{x}}{ah}-\frac{\alpha\bar{Z}C_{x}^{2}}{a^{2}h^{2}}+\frac{\alpha^{2}\bar{Z}C_{x}^{2}}{2a^{2}h^{2}}\right].$$

and

$$MSE\left(t_{erh}\right)=\theta\bar{Z}^{2}\left[C_{z}^{2}+\frac{\alpha^{2}C_{x}^{2}}{a^{2}h^{2}}-\frac{2\alpha\rho_{yx}C_{y}C_{x}}{ah}\right].$$

The optimum value of "*a*" is

$$MSE\left(t_{erh}\right)=\theta\bar{Z}^2\left[C_z^2+\frac{\alpha^2C_x^2}{a^2h^2}-\frac{2\alpha\rho_{yx}C_yC_x}{ah}\right].$$

$$\hat{a}=\frac{\alpha C_x^2}{h\rho_{zx}C_zC_x}.$$

The minimum *MSE* of $t_{erh}$is given by

$$MSE_{\min}\left(t_{erh}\right)=\theta\bar{Y}^2C_z^2\left[1-\rho_{zx}^2\right].$$

Waseem *et al.* (2021) proposed generalized exponential estimator for estimating the population mean of sensitive study variable using twobenchmark variables. The generalized exponential estimator along with the expressions of bias and *MSE* are defined as

$$t_{errh}=\bar{z}exp\left(\alpha\left(\frac{\bar{X}^{\left(\frac{1}{h}\right)}-\bar{x}^{\left(\frac{1}{h}\right)}}{\bar{X}^{\left(\frac{1}{h}\right)}+\left(a-1\right)\bar{x}^{\left(\frac{1}{h}\right)}}+\frac{\bar{U}^{\left(\frac{1}{h}\right)}-\bar{u}^{\left(\frac{1}{h}\right)}}{\bar{U}^{\left(\frac{1}{h}\right)}+\left(a-1\right)\bar{u}^{\left(\frac{1}{h}\right)}}\right)\right).$$

$$Bias\left(t_{errh}\right)=\theta\bar{Y}\left[\frac{\alpha C_x^2}{ah^2}\left(1-\frac{1}{a}\right)+\frac{\alpha C_u^2}{bh^2}\left(1-\frac{1}{b}\right)-\frac{\alpha}{h}\left(\frac{\rho_{zx}C_zC_x}{a}-\frac{\rho_{zu}C_zC_u}{b}\right)-\frac{\alpha^2}{2h^2}\left(\frac{C_x^2}{a^2}+\frac{C_u^2}{b^2}-\frac{2\rho_{xu}C_xC_u}{ab}\right)\right].$$

and

$$MSE\left(t_{errh}\right)=\theta\bar{Y}^2\left[C_z^2+\frac{\alpha^2C_x^2}{a^2h^2}+\frac{\alpha^2C_u^2}{b^2h^2}-\frac{2\alpha\rho_{zx}C_zC_x}{ah}-\frac{2\alpha\rho_{zu}C_zC_u}{bh}+\frac{2\alpha^2\rho_{xu}C_xC_u}{abh}\right].$$

The optimum value of "*a*"and "*b*" are

$$\hat{a} = \frac{\alpha}{\rho_{zu} h}\left[\frac{\rho_{xu} C_x C_u - C_x^2 C_u^2}{\rho_{zx} C_z C_x C_u^2 - C_u^2 C_z C_u}\right]$$

and

$$\hat{b} = \frac{\alpha}{\rho_{zx} h}\left[\frac{\rho_{xu} C_x C_u - C_x^2 C_u^2}{\rho_{xu} C_u C_z C_x^2 - C_x^2 C_z C_x}\right].$$

The minimum MSE of $t_{errh}$is given by

$$MSE_{\min}\left(t_{errh}\right) = \theta \bar{Y}^2 \left[C_z^2 + \frac{C_x^2 \rho_{zu} C_z C_u + C_u^2 \rho_{zx} C_z C_x - 2\rho_{zx}\rho_{zu}\rho_{xu} C_z^2 C_x^2 C_u^2}{\rho_{xu} C_x C_u - C_x^2 C_u^2}\right].$$

Using the exponential estimators for mean estimation given by Waseem *et al.* (2020, 2021), we have proposed the two exponential estimators for estimating the variance of the finite population of sensitive study variable using auxiliary information. We have used additive *RRM* to get the reported response.

CHAPTER 4:

GENERALIZED EXPONENTIAL ESTIMATOR FOR VARIANCE ESTIMATION OF SENSITIVE STUD VARIABLE USING AUXILIARY INFORMATION

4.1 INTRODUCTION

In this chapter, we have proposed an exponential estimator for sensitive study variable using supplementary information operating*RRM*. Mathematical terminologies of bias and *MSE* of the proposed exponential estimator are derived using Taylor and exponential series. The expression of minimum *MSE* of the proposed exponential estimator is also derived using the optimization of the constant. Few special cases are also obtained as the family of the proposed estimator using different values of the scalar constants.

4.2 SAMPLING METHODOLOGY, NOTATIONS AND FORMULAS

"Consider a finite population, say $U = \{U_1, U_2, U_3, \ldots, U_N\}$ of N units label from $i = 1, 2, 3, \ldots, N$. Let Y be a sensitive survey variable and X is a non-sensitive supporting variable which is correlated with the main sensitive study variable Y. Let S be a scrambling variable and assumed to be independent from both the sensitive study variable Y and the non-sensitive auxiliary variable X. Further assumed that the scrambling variable follows normal distribution with zero mean and constant variance. The reported response, say Z, is obtained using the additive model $Z = Y+S$, given by Pollock and Beck (1976). Let y_i, z_i and x_i, be the experimental values of Y, Z and X obtained during the surveys. Suppose a random sample of size n units is drawn from the populations of Y and X using *SRSWOR*.

Let the means and variances of actual response Y, reported response Z and auxiliary variable X are denoted as

$$\mu_y\left(=N^{-1}\sum_{i=1}^{N} y_i\right), \mu_z\left(=N^{-1}\sum_{i=1}^{N} z_i\right), \mu_x\left(=N^{-1}\sum_{i=1}^{N} x_i\right)$$

and

$$S_y^2\left(=(N-1)^{-1}\sum_{i=1}^{N}\left(y_i-\bar{Y}\right)^2\right), S_z^2\left(=(N-1)^{-1}\sum_{i=1}^{N}\left(z_i-\bar{Z}\right)^2\right),$$

and

$$S_x^2\left(=(N-1)^{-1}\sum_{i=1}^{N}\left(x_i-\bar{X}\right)^2\right)S_x^2 \text{ respectively".}$$

Let ξ_y, ξ_z and ξ_x sampling error terms of actual response Y, reported response Z and the auxiliary variable X for population variances are defined as

$$\left.\begin{aligned} s_y^2=S_y^2\left(1+\xi_y\right), s_z^2=S_z^2\left(1+\xi_z\right), s_x^2=S_x^2\left(1+\xi_x\right) \\ E\left(\xi_y\right)=E\left(\xi_z\right)=E\left(\xi_x\right)=0 \\ E\left(\xi_z^2\right)=\left(\lambda_{40}-1\right)=\lambda_{40}^*, E\left(\xi_x^2\right)=\left(\lambda_{04}-1\right) \\ =\lambda_{04}^*, E\left(\xi_z\xi_x\right)=\left(\lambda_{22}-1\right)=\lambda_{22}^* \end{aligned}\right\rangle \tag{4.2.1}$$

where

$$\lambda_{pq}=\frac{\mu_{pq}}{\mu_{20}^{p/2}\mu_{02}^{p/2}}, \mu_{pq}=\frac{1}{N-1}\sum_{i=1}^{N}\left(z_i-\bar{Z}\right)^p\left(x_i-\bar{X}\right)^q, \text{ and } \theta=1/n.$$

4.3 PROPOSED GENERALIZED EXPONENTIAL ESTIMATOR

In this section, we have proposed a generalized exponential estimator for the estimation of population variance of confidential variable using benchmark auxiliary information in SRS design. The estimator is defined as

$$t_{eji}=\left(s_z^2-S_s^2\right)exp\left(F\frac{\left(\alpha S_x^{2b^{-1}}+\beta\right)-\left(\alpha s_x^{2b^{-1}}+\beta\right)}{\left(\alpha S_x^{2b^{-1}}+\beta\right)+(f-1)\left(\alpha s_x^{2b^{-1}}+\beta\right)}\right). \tag{4.3.1}$$

where b (b> 0) and$F$($\infty < F < -\infty$) are two real constants and supposed to be known. Further, α $(\alpha \neq 0)$ and β be the scalar quantities which can take distinctive groupings of real number and known parameters of benchmark variable to modify the form of proposed exponential estimators. The optimized constant "*F*"is used to catch the minimum *MSE* of the proposed generalized exponential estimator.

In order to derive the bias and *MSE* of the proposed estimator given in Eq. (4.3.1), we rewrite it in the form of error terms defined in Eq. (4.2.1).

$$t_{eji} = \left(S_z^2\left(1+\xi_z\right) - S_s^2\right)$$
$$exp\left(F\frac{\left(\alpha S_x^{2b^{-1}}+\beta\right)-\left(\alpha S_x^{2b^{-1}}\left(1+\xi_x\right)^{b^{-1}}+\beta\right)}{\left(\alpha S_x^{2b^{-1}}+\beta\right)+\left(f-1\right)\left(\alpha S_x^{2b^{-1}}\left(1+\xi_x\right)^{b^{-1}}+\beta\right)}\right). \tag{4.3.2}$$

On simplification of Eq. (4.3.2), we have

$$t_{eji} = \left(S_z^2 + S_z^2\xi_z - S_s^2\right)$$
$$exp\left(F\frac{\left(\alpha S_x^{2b^{-1}}+\beta\right)-\left(\left(\alpha S_x^{2b^{-1}}+\beta\right)-\alpha b^{-1}S_x^{2b^{-1}}\xi_x\right)}{\left(\alpha S_x^{2b^{-1}}+\beta\right)+\left(f-1\right)\left(\left(\alpha S_x^{2b^{-1}}+\beta\right)+\alpha b^{-1}S_x^{2b^{-1}}\xi_x\right)}\right). \tag{4.3.3}$$

or,

$$t_{eji} = \left(S_y^2 + S_z^2\xi_z\right)exp\left(F\frac{\left(-\alpha b^{-1}S_x^{2b^{-1}}\xi_x\right)}{f\left(\alpha S_x^{2b^{-1}}+\beta\right)\left\{1+\frac{\left(f-1\right)}{f}\left(\frac{\alpha b^{-1}S_x^{2b^{-1}}\xi_x}{\left(\alpha S_x^{2b^{-1}}+\beta\right)}\right)\right\}}\right). \tag{4.3.4}$$

Simplifying and applying Taylor series on Eq. (4.3.4) up to first order, we have

$$t_{eji} \approx \left(S_y^2 + S_z^2\xi_z\right)exp\left(-\frac{F}{bf}\psi_j\xi_x\left\{1-\left(1-\frac{1}{f}\right)\frac{1}{b}\psi_j\xi_x\right\}\right), \tag{4.3.5}$$

where $\psi_j = \frac{\alpha S_x^{2b^{-1}}}{\left(\alpha S_x^{2b^{-1}}+\beta\right)}$.

or,

Further simplifying and expanding exponential series in Eq. (4.3.5), we have

$$t_{eji} \approx \left(S_y^2 + S_z^2\xi_z\right)\begin{pmatrix} 1 - w\frac{1}{b}\psi_j\xi_x \\ \qquad + w\left(1-\frac{1}{f}\right)\frac{1}{b^2}\psi_j^2\xi_x^2 + \frac{1}{2b^2}w^2\psi_j^2\xi_x^2 \end{pmatrix}, \qquad (4.3.6)$$

or,

$$t_{eji} \approx \left(S_y^2 + S_z^2\xi_z\right)\left(1 - w\frac{1}{b}\psi_j\xi_x + \left\{\left(1-\frac{1}{f}\right) + \frac{w}{2}\right\}\frac{1}{b^2}w\psi_j^2\xi_x^2\right), \qquad (4.3.7)$$

On simplification of Eq. (4.3.7), we have

$$\begin{aligned} t_{eji} \approx\ & S_y^2 + S_z^2\xi_z - wS_y^2\frac{1}{b}\psi_j\xi_x \\ & - wS_z^2\frac{1}{b}\psi_j\xi_z\xi_x + \left\{\left(1-\frac{1}{f}\right) + \frac{w}{2}\right\}S_y^2 w\frac{1}{b^2}\psi_j^2\xi_x^2, \end{aligned} \qquad (4.3.8)$$

or,

$$\begin{aligned} t_{eji} - S_y^2 \approx\ & S_z^2\xi_z - wS_y^2\frac{1}{b}\psi_j\xi_x \\ & - wS_z^2\frac{1}{b}\psi_j\xi_z\xi_x + \left\{\left(1-\frac{1}{f}\right) + \frac{w}{2}\right\}S_y^2 w\frac{1}{b^2}\psi_j^2\xi_x^2, \end{aligned} \qquad (4.3.9)$$

Simplifying and applying expectation on Eq. (4.3.9), we have

$$E\left(t_{eji} - S_y^2\right) \approx E\begin{bmatrix} S_z^2\xi_z - wS_y^2\frac{1}{b}\psi_j\xi_x \\ \qquad - wS_z^2\frac{1}{b}\psi_j\xi_z\xi_x + \left\{\left(1-\frac{1}{f}\right) + \frac{w}{2}\right\}wS_y^2\frac{1}{b^2}\psi_j^2\xi_x^2, \end{bmatrix}. \qquad (4.3.10)$$

Final bias of the proposed generalized exponential estimator given in Eq. (4.3.10) is

$$Bias\left(t_{eji}\right) \approx w\frac{1}{b^2}\psi_j S_y^2 \lambda_{40}^* \left[\left\{\left(1-\frac{1}{f}\right)+\frac{w}{2}\right\}\psi_j - b\Re_{zy}\Theta\right], \qquad (4.3.11)$$

where $\Theta = \dfrac{\lambda_{22}^*}{\lambda_{04}^*}$ and $\Re_{zy} = \dfrac{S_z^2}{S_y^2}$.

In order to derive the *MSE* of the proposed generalized exponential estimator, we simplifying up to first-order Eq. (4.3.5) and get

$$t_{eji} \approx \left(S_y^2 + S_z^2\xi_z\right) exp\left(-w\frac{1}{b}\psi_j\xi_x\right), \qquad (4.3.12)$$

On expanding exponential series up to first order of Eq. (4.3.12), we have

$$t_{eji} \approx \left(S_y^2 + S_z^2\xi_z\right)\left(1 - w\frac{1}{b}\psi_j\xi_x\right), \qquad (4.3.13)$$

On further simplification of Eq. (4.3.13), we have

$$t_{eji} - S_y^2 \approx \left(S_z^2\xi_z - w\frac{1}{b}S_y^2\psi_j\xi_x\right), \qquad (4.3.14)$$

Simplifying and fascinating expectation on both sides of the Eq. (4.3.14), we have

$$E\left(t_{eji} - S_y^2\right)^2 \approx E\left(S_z^4\xi_z^2 + w^2\frac{1}{b^2}S_y^4\psi_j^2\xi_x^2 - 2w\frac{1}{b}S_z^2S_y^2\psi_j\xi_z\xi_x\right). \qquad (4.3.15)$$

The MSE of the proposed generalized exponential estimator given in Eq. (4.3.15) is

$$MSE\left(t_{eji}\right) \approx \theta S_y^4\left(\frac{1}{b^2}w\ \psi_j\lambda_{04}^*\left(w\psi_j - 2b\Re_{zy}\Theta\right) + \Re_{zy}^2\lambda_{40}^*\right). \qquad (4.3.16)$$

In order to get the least MSE, partially differentiating Eq. (4.3.16) with respect to f and equating to zero, we have

$$\frac{\partial MSE\left(t_{eji}\right)}{\partial w}=0.$$

or,

$$w\frac{1}{b}S_y^2\psi_j\lambda_{04}^* = S_z^2\lambda_{22}^*$$

or,

$$w = \frac{1}{\psi_j}b\frac{S_z^2}{S_y^2}\frac{\lambda_{22}^*}{\lambda_{04}^*}.$$

or,

$$\hat{f} = F\psi_j\left(b\Re_{zy}\Theta\right)^{-1}.$$

Substituting "*f*" in Eq. (4.3.16), we have the lowest*MSE* of the proposed generalized exponential estimator is.

$$MSE_{min}\left(t_{eji}\right) \approx \theta S_z^4\lambda_{40}^*\left(1-\frac{\lambda_{22}^{*2}}{\lambda_{40}^*\lambda_{04}^*}\right). \tag{4.3.17}$$

or,

$$MSE_{min}\left(t_{eji}\right) \approx \theta S_z^4\lambda_{40}^*\left(1-\chi^2\right), \tag{4.3.18}$$

where $\chi = \dfrac{\lambda_{22}^*}{\left(\lambda_{40}^*\lambda_{04}^*\right)^{1/2}}$.

The special cases of the proposed generalized exponential estimator are obtained at different values and summarized in Appendix A.

4.4 MATHEMATICAL COMPARISONS

In this section, we mathematically compared the proposed exponential estimator with the existing estimators.

- The proposed exponential estimator will perform improved than the sample variance estimator if

$$MSE_{min}\left(t_{eji}\right) < var\left(t_y\right).$$

which implies that,

$$\theta S_z^4 \lambda_{40}^* \left(1-\chi^2\right) < \theta S_z^4 \lambda_{40}^*.$$

or,

$$\chi^2 > 1$$

- The proposed exponential estimator will perform improved than the classical ratio estimator if

$$MSE_{min}\left(t_{eji}\right) < MSE\left(t_r\right).$$

which implies that,

$$\theta S_z^4 \lambda_{40}^* \left(1-\chi^2\right) < \theta S_z^4 \left(\lambda_{40}^* + \lambda_{04}^* - 2\lambda_{22}^*\right).$$

or,

$$\chi^2 > 2\Theta - \left(\lambda_{04}^* / \lambda_{40}^*\right).$$

- The proposed exponential estimator will perform healthier than the basic exponential ratio estimator if

$$MSE_{min}\left(t_{eji}\right) < MSE\left(t_{e1i}\right).$$

which implies that,

$$\theta S_z^4 \lambda_{40}^* \left(1-\chi^2\right) < \theta S_z^4 \left(\lambda_{40}^* + (1/4)\lambda_{04}^* - \lambda_{22}^*\right).$$

or,

$$\chi^2 > \Theta - \left(\lambda_{04}^* / 4\lambda_{40}^*\right).$$

4.5 NUMERICAL STUDY

In this section, "we have conducted a numerical study to check the performance of the proposed exponential estimator over the existing estimators considered in this research. We have conducted the simulation study for the population generated by normal distribution with different mean and covariance matrix. Further, we also used the real data-set of information and communication technologies population, which was initially used by Sousa *et al.* (2010) to defend the mathematical properties of the proposed exponential estimator. The efficiency of the proposed estimators is evaluated on the basis of absolute bias (AB), MSE and percentage relative efficiency (PRE). The formulas of *AB*,*MSE*and *PRE*are defined as”

$$AB(\Omega)=\left|\frac{1}{q}\sum_{i=1}^{q}\left(\Omega-S_y^2\right)\right|.$$

$$MSE(\Omega)=\frac{1}{q}\sum_{i=1}^{q}\left(\Omega-S_y^2\right)^2.$$

and

$$PRE(\Omega)=\frac{Var\left(s_y^2\right)}{MSE(\Omega)}\text{x}100,$$

where $\Omega = t_y,\ t_r,\ t_{eji},\ t_{e1i},\ t_{e2i},\ t_{e3i}\ ,...,\ t_{e12i}.$

4.5.1 Simulation Study

In this sub-section, we have conducted a simulation study to evaluate the performance of the proposed generalized exponential estimators using a bivariate normal population with means and variance covariance matrix to represent the distribution of Y and X. The characteristics of populations are:

Population

$$\mu = \begin{bmatrix} 2 & 2 \end{bmatrix} \quad \Sigma = \begin{bmatrix} 06 & 03 \\ & 02 \end{bmatrix} \text{ and } \rho_{yx} = 0.8684$$

The “following steps have been coded in *R*-language to compute the *AB* and *PRE* of all the estimators considered in this study.

- **Step 1:** From the described population, we computed the population variance of the auxiliary variable *x*.
- **Step 2:** We consider different sample sizes for each population to generate the samples using *SRSWOR*.

- **Step 3:** For each sample size, the values of the proposed generalized estimator and existing estimators considered in this study are computed.
- **Step 4:** The process in Step-2 and Step-3 is repeated 20,000 times and the results of *AB*, *MSE* and *PRE* are reported in Table-4.1.

Table 4.1
Theoretical and Empirical *AB*, *MSE* and *PRE* of all the Estimators for Artificial Population

			Empirical		*Theoretical*		
N	n	***Estimators***	***AB***	***MSE***	***AB***	***MSE***	***PRE***
5000	500	t_y	0.00985	0.12753	0.00875	0.12299	100.0000
		t_r	0.00322	0.06863	0.00212	0.06408	191.9257
		t_{eji}	0.01102	0.06182	0.01002	0.06046	198.4746
		t_{e1i}	0.00905	0.06852	0.00795	0.06398	192.2334
		t_{e2i}	0.01007	0.09064	0.00897	0.08610	142.8484
		t_{e3i}	0.00331	0.06842	0.00221	0.06388	192.5448
		t_{e4i}	0.01012	0.10724	0.00902	0.10270	119.7616
		t_{e5i}	0.00888	0.09403	0.00777	0.08949	137.4362
		t_{e6i}	0.00980	0.10932	0.00869	0.10478	117.3839
		t_{e7i}	0.00444	0.07226	0.00334	0.06771	181.6355
		t_{e8i}	0.00993	0.11806	0.00883	0.11352	108.3458
		t_{e9i}	0.01325	0.06827	0.01215	0.06372	193.0089
		t_{e10i}	0.01155	0.06847	0.01045	0.06392	192.4026
		t_{e11i}	0.01661	0.24124	0.01550	0.23670	51.96098
		t_{e12i}	0.01070	0.09062	0.00960	0.08608	142.8790
5000	1000	t_y	0.00773	0.06665	0.00662	0.06211	100
		t_r	0.00859	0.03704	0.00749	0.03249	191.1432
		t_{eji}	0.00523	0.03132	0.00443	0.03054	199.9556
		t_{e1i}	0.00689	0.03630	0.00579	0.03176	195.5868
		t_{e2i}	0.00700	0.04759	0.00589	0.04304	144.2919
		t_{e3i}	0.00857	0.03698	0.00747	0.03244	191.4652
		t_{e4i}	0.00728	0.05615	0.00618	0.05160	120.3577
		t_{e5i}	0.00767	0.04933	0.00657	0.04479	138.6761
		t_{e6i}	0.00748	0.05722	0.00638	0.05268	117.9033
		t_{e7i}	0.00936	0.03817	0.00825	0.03362	184.7116

N	n	Estimators	Empirical AB	MSE	Theoretical AB	MSE	PRE
		t_{e8i}	0.00755	0.06174	0.00645	0.05720	108.581
		t_{e9i}	0.00354	0.03695	0.00244	0.03241	191.6226
		t_{e10i}	0.00563	0.03628	0.00453	0.03174	195.6956
		t_{e11i}	0.00283	0.13058	0.00172	0.12604	49.27626
		t_{e12i}	0.00668	0.04758	0.00558	0.04304	144.3134
5000	2000	t_y	0.00187	0.02663	0.00077	0.02209	100
		t_r	0.00331	0.01500	0.00221	0.01046	211.2243
		t_{eji}	0.00035	0.009821	0.00025	0.009344	215.6432
		t_{e1i}	0.00213	0.01522	0.00102	0.01067	206.9603
		t_{e2i}	0.00188	0.01953	0.00078	0.01498	147.4338
		t_{e3i}	0.00331	0.01499	0.00221	0.01045	211.3767
		t_{e4i}	0.00185	0.02273	0.00074	0.01819	121.4611
		t_{e5i}	0.00210	0.02018	0.00100	0.01563	141.2976
		t_{e6i}	0.00190	0.02313	0.00080	0.01859	118.8541
		t_{e7i}	0.00297	0.01593	0.00187	0.01139	193.9629
		t_{e8i}	0.00187	0.02481	0.00076	0.02027	108.9846
		t_{e9i}	0.00145	0.01498	0.00035	0.01044	211.6174
		t_{e10i}	0.00166	0.01522	0.00056	0.01067	206.9686
		t_{e11i}	0.00118	0.04795	0.00008	0.04341	50.89057
		t_{e12i}	0.00177	0.01953	0.00066	0.01498	147.4325

From the results summarized in Table-4.1, we observed that the proposed class of exponential estimator performed better than the non-exponential estimators for the estimation of population variance of sensitive study variable in the presence of non-sensitive auxiliary variable except t_{e11i}. Moreover, the precision of the proposed class of exponential estimator increases by the increase of the sample size.The estimators, t_{e3i}, t_{e9i}, and t_{e10i}performed well than the existing estimators and the other members of the proposed class of exponential estimator.Consequently, we may say that the proposed class of exponential estimator is superior then the other existing estimators for different sample sizes.

4.5.2 Real Data Application

In this section, we have considered a real population based on the survey of information and communication technologies (*ICT*) usage in initiatives in 2009 seat in Portugal (Smilhily and Storm 2010). We took three different sample sizes (n = 1000, 2000 and 3000) to assess the performance of the estimators considered in this research. The theoretical and empirical*AB*,and *MSE* of all the estimators are given in Table-4.2."

The main characteristics of the populations are

$$N = 5336, \quad \bar{X} = 22.99, \quad \bar{Y} = 30.19,$$
$$S_x = 172.09, \quad S_y = 138.65, \quad \rho_{yx} = 0.9632.$$

Table 4.2
Theoretical and Empirical *AB*, *MSE* and *PRE*
of all Estimators for Real Population

			Empirical		*Theoretical*		
N	*n*	***Estimators***	***AB***	***MSE***	***AB***	***MSE***	***PRE***
		t_y	4.81	216264.25	4.78	216263.50	100.00
		t_r	5.40	33453.75	5.36	33453.00	646.47
		t_{eji}	5.23	33412.32	5.28	33413.43	648.98
		t_{e1i}	6.51	72468.55	6.47	72467.80	298.43
		t_{e2i}	6.01	131258.35	5.98	131257.60	164.76
		t_{e3i}	5.40	33437.97	5.37	33437.22	646.77
5336	500	t_{e4i}	5.50	170483.05	5.46	170482.30	126.85
		t_{e5i}	5.22	139413.45	5.18	139412.70	155.12
		t_{e6i}	5.26	175248.85	5.22	175248.10	123.40
		t_{e7i}	3.68	83266.13	3.65	83265.38	259.73
		t_{e8i}	5.09	195108.45	5.06	195107.70	110.84
		t_{e9i}	11.02	33491.43	10.98	33490.68	645.74
		t_{e10i}	7.91	72459.55	7.88	72458.80	298.46

N	n	Estimators	Empirical AB	Empirical MSE	Theoretical AB	Theoretical MSE	PRE
		t_{e11i}	17.21	269106.85	17.18	269106.10	80.36
		t_{e12i}	6.36	131252.55	6.33	131251.80	164.77
5336	1000	t_y	23.09	1322309.75	23.06	1322309.00	100.00
		t_r	19.46	208133.35	19.43	208132.60	635.32
		t_{eji}	18.98	207042.15	19.01	207043.40	641.43
		t_{e1i}	12.41	430084.15	12.38	430083.40	307.45
		t_{e2i}	15.55	791948.35	15.51	791947.60	166.97
		t_{e3i}	19.23	207283.15	19.19	207282.40	637.93
		t_{e4i}	18.77	1036007.75	18.73	1036007.00	127.64
		t_{e5i}	20.53	842902.85	20.50	842902.10	156.88
		t_{e6i}	20.29	1065886.75	20.25	1065886.00	124.06
		t_{e7i}	30.17	496436.95	30.13	496436.20	266.36
		t_{e8i}	21.31	1189891.75	21.27	1189891.00	111.13
		t_{e9i}	15.95	209336.85	15.92	209336.10	631.67
		t_{e10i}	3.58	429521.75	3.55	429521.00	307.86
		t_{e11i}	54.40	1748382.75	54.37	1748382.00	75.63
		t_{e12i}	13.34	791643.45	13.30	791642.70	167.03
1000	2000	t_y	82.39	1416382.75	82.35	1416382.00	100.00
		t_r	61.58	209175.25	61.55	209174.50	671.45
		t_{eji}	60.95	207931.05	60.98	208001.40	682.58
		t_{e1i}	81.52	450344.85	81.49	450344.10	314.51
		t_{e2i}	84.52	842355.95	84.49	842355.20	168.15
		t_{e3i}	61.91	208099.05	61.88	208098.30	680.63
		t_{e4i}	83.95	1106564.75	83.91	1106564.00	128.00
		t_{e5i}	79.65	897368.65	79.62	897367.90	157.84
		t_{e6i}	82.66	1138858.75	82.63	1138858.00	124.37
		t_{e7i}	63.81	521842.05	63.78	521841.30	271.42

			Empirical		Theoretical		
N	n	**Estimators**	**AB**	**MSE**	**AB**	**MSE**	**PRE**
		t_{e8i}	82.94	1273090.75	82.90	1273090.00	111.26
		t_{e9i}	99.59	210015.35	99.55	210014.60	674.42
		t_{e10i}	91.00	449901.85	90.96	449901.10	314.82
		t_{e11i}	116.41	1867118.75	116.38	1867118.00	75.86
		t_{e12i}	86.69	842097.45	86.66	842096.70	168.20

As expected, the numerical findings gain from the simulation study support the outcome obtained in the real data illustration. The results presented in Table 4.2 showed that the proposed class of estimator is more efficient than the existing estimators at different sample sizes. The estimator, t_{e3i}, executed well among all the existing and the members of the proposed class of exponential estimators. Thus, from the results of real data study, we conclude that the proposed class of estimator is efficient for estimating the population variance of a sensitive study variable usingbenchmark variable in *SRS* design except t_{e11i}. On the basis of numerical results, the generalized exponential estimator found to be more superior than all the estimators used in this research at different sample sizes.

4.6 CONCLUSION

In this chapter, we have considered the generalized exponential estimator of a sensitive study variable with single auxiliary variable for estimating the finite population variance based on additive scrambled response model. We examined the proposed and existing estimators using theoretical comparisons and numerical study. From the findings, we concluded that the proposed generalized exponential estimator performed efficiently than the existing estimators using both real and artificial populations. Overall, we may conclude that the proposed estimator is efficient and useful for real-life surveys.

CHAPTER 5:

GENERALIZED EXPONENTIAL ESTIMATORS FOR VARIANCE ESTIMATION OF SENSITIVE STUDY VARIABLE USING TWO AUXILIARY VARIABLES

5.1 INTRODUCTION

In this chapter, a generalized exponential estimator is proposed for confidential study variable in the presence of two benchmark variables using *RRM*. The mathematical equations of bias and *MSE* of the proposed exponential estimator are originatedby means of Taylor and exponential series. The expression of smallest*MSE* of the proposed exponential estimator is also obtained using the optimization. Many special cases are also produced as the family of the proposed estimator using different values of scalar constants.

5.2 SAMPLING METHODOLOGY, NOTATIONS AND FORMULAS

"Let us consider a finite population, say $P = \{P_1, P_2, P_3, \ldots, P_N\}$ of N units labeled from $i = 1, 2, 3, \ldots N$. Let Y be a sensitive survey variable and X and U are non-sensitive supporting variables which are correlated with the main sensitive study variable Y and uncorrelated with each other. Let S be a scrambling variable and assumed to be independent from the sensitive study variable Y and the non-sensitive auxiliary variables X and U. Further assumed that the scrambling variable follows normal distribution with zero mean and constant variance. The reported response, say Z, is obtained using the additive model $Z = Y + S$, given by Pollock and Beck (1976). Let y_i, z_i, x_i and u_i be the observed values of Y, Z, $X$and U obtained during the surveys. Suppose a random sample of size n units is drawn from the populations of Y, X and U using *SRSWOR*".

Let the means and variances of actual response Y, reported response Z and auxiliary variable X are denoted as $\mu_y\left(=N^{-1}\sum_{i=1}^{N} y_i\right)$, $\mu_z\left(=N^{-1}\sum_{i=1}^{N} z_i\right)$, $\mu_x\left(=N^{-1}\sum_{i=1}^{N} x_i\right)$, $\mu_u\left(=N^{-1}\sum_{i=1}^{N} u_i\right)$, and $S_y^2\left(=(N-1)^{-1}\sum_{i=1}^{N}\left(y_i-\bar{Y}\right)^2\right)$,

$S_z^2\left(=(N-1)^{-1}\sum_{i=1}^{N}\left(z_i-\bar{Z}\right)^2\right)$, $S_x^2\left(=(N-1)^{-1}\sum_{i=1}^{N}\left(x_i-\bar{X}\right)^2\right)$ and $S_u^2\left(=(N-1)^{-1}\sum_{i=1}^{N}\left(u_i-\bar{U}\right)^2\right)$ respectively.

Let ξ_y, ξ_z, ξ_x and ξ_u sampling error terms of actual response *Y*, reported response *Z* and the auxiliary variable *X* for population variances are defined as

$$\left.\begin{aligned}
& s_y^2 = S_y^2\left(1+\xi_y\right), s_z^2 = S_z^2\left(1+\xi_z\right), s_x^2 = S_x^2\left(1+\xi_x\right), s_u^2 = S_u^2\left(1+\xi_u\right) \\
& \quad E\left(\xi_y\right) = E\left(\xi_z\right) = E\left(\xi_x\right) = E\left(\xi_u\right) = 0 \\
& \quad E\left(\xi_z^2\right) = \left(\lambda_{400}-1\right) = \lambda_{400}^*, E\left(\xi_x^2\right) = \left(\lambda_{040}-1\right) \\
& \quad = \lambda_{040}^*, E\left(\xi_u^2\right) = \left(\lambda_{004}-1\right) = \lambda_{004}^*, \\
& \quad E\left(\xi_z\xi_x\right) = \left(\lambda_{220}-1\right) = \lambda_{220}^*, E\left(\xi_z\xi_u\right) = \left(\lambda_{202}-1\right) \\
& \quad = \lambda_{202}^* \text{ and } E\left(\xi_x\xi_u\right) = \left(\lambda_{022}-1\right) = \lambda_{022}^*
\end{aligned}\right| \qquad (5.2.1)$$

where

$$\lambda_{pqr} = \frac{\mu_{pqr}}{\mu_{200}^{p/2}\mu_{020}^{q/2}\mu_{002}^{r/2}}$$

and

$$\mu_{pqr} = \frac{1}{N-1}\sum_{i=1}^{N}\left(z_i-\bar{Z}\right)^p\left(x_i-\bar{X}\right)^q\left(u_i-\bar{U}\right)^r.$$

5.3 PROPOSED GENERALIZED EXPONENTIAL ESTIMATOR WITH TWO AUXILIARY VARIABLES

In this section, we have proposed a generalized exponential estimator for the estimation of population variance of confidential study variable using two benchmark variables in SRS design. The proposed exponential estimator is defined as

$$t_{eji}^{*}=\left(s_z^2-S_s^2\right)exp\left(\left(\begin{array}{c} M_1\dfrac{\left(\alpha S_x^{2b^{-1}}+\beta\right)-\left(\alpha s_x^{2b^{-1}}+\beta\right)}{\left(\alpha S_x^{2b^{-1}}+\beta\right)+\left(m_1-1\right)\left(\alpha s_x^{2b^{-1}}+\beta\right)} \\ +M_2\dfrac{\left(\alpha S_u^{2b^{-1}}+\beta\right)-\left(\alpha s_u^{2b^{-1}}+\beta\right)}{\left(\alpha S_u^{2b^{-1}}+\beta\right)+\left(m_2-1\right)\left(\alpha s_u^{2b^{-1}}+\beta\right)} \end{array}\right)\right). \tag{5.3.1}$$

where b ($b>0$) and M_1andM_2($+\infty<M_1,M_2<-\infty$) are two real constants and supposed to be known. Further, α $\left(\alpha\neq 0\right)$ and β be the scalar quantities which can assume different combinations of real number and known parameters of auxiliary variable to modify the form of proposed exponential estimators.The optimized constants m_1 and m_2 are used to find the least MSE. The proposed exponential estimator given in Eq. (5.3.1) can be re-write in the form of errors given in Eq. (5.2.1) as

$$t_{eji}^{*}=\left(S_z^2\left(1+\xi_z\right)-S_s^2\right)$$
$$exp\left(\begin{array}{c} M_1\dfrac{\left(\alpha S_x^{2b^{-1}}+\beta\right)-\left(\alpha S_x^{2b^{-1}}\left(1+\xi_x\right)^{b^{-1}}+\beta\right)}{\left(\alpha S_x^{2b^{-1}}+\beta\right)+\left(f_1-1\right)\left(\alpha S_x^{2b^{-1}}\left(1+\xi_x\right)^{b^{-1}}+\beta\right)} \\ +M_2\dfrac{\left(\alpha S_u^{2b^{-1}}+\beta\right)-\left(\alpha S_u^{2b^{-1}}\left(1+\xi_u\right)^{b^{-1}}+\beta\right)}{\left(\alpha S_u^{2b^{-1}}+\beta\right)+\left(m_2-1\right)\left(\alpha S_u^{2b^{-1}}\left(1+\xi_u\right)^{b^{-1}}+\beta\right)} \end{array}\right). \tag{5.3.2}$$

On simplification of Eq. (5.3.2), we have

$$t_{eji}^{*}=\left(S_z^2+S_z^2\xi_z-S_s^2\right)$$

$$exp\left(\begin{array}{l} M_1\dfrac{\left(\alpha S_x^{2b^{-1}}+\beta\right)-\left(\left(\alpha S_x^{2b^{-1}}+\beta\right)-\alpha b^{-1}S_x^{2b^{-1}}\xi_x\right)}{\left(\alpha S_x^{2b^{-1}}+\beta\right)+\left(m_1-1\right)\left(\left(\alpha S_x^{2b^{-1}}+\beta\right)+\alpha b^{-1}S_x^{2b^{-1}}\xi_x\right)} \\ +M_2\dfrac{\left(\alpha S_u^{2b^{-1}}+\beta\right)-\left(\left(\alpha S_u^{2b^{-1}}+\beta\right)-\alpha b^{-1}S_u^{2b^{-1}}\xi_u\right)}{\left(\alpha S_u^{2b^{-1}}+\beta\right)+\left(m_2-1\right)\left(\left(\alpha S_u^{2b^{-1}}+\beta\right)+\alpha b^{-1}S_u^{2b^{-1}}\xi_u\right)} \end{array}\right). \quad (5.3.3)$$

or,

$$t_{eji}^{*}=\left(S_y^2+S_z^2\xi_z\right)$$

$$exp\left(\begin{array}{l} M_1\dfrac{\left(-\alpha b^{-1}S_x^{2b^{-1}}\xi_x\right)}{m_1\left(\alpha S_x^{2b^{-1}}+\beta\right)\left\{1+\dfrac{\left(m_1-1\right)}{m_1}\left(\dfrac{\alpha b^{-1}S_x^{2b^{-1}}\xi_x}{\left(\alpha S_x^{2b^{-1}}+\beta\right)}\right)\right\}} \\ +M_2\dfrac{\left(-\alpha b^{-1}S_u^{2b^{-1}}\xi_u\right)}{m_2\left(\alpha S_u^{2b^{-1}}+\beta\right)\left\{1+\dfrac{\left(m_2-1\right)}{m_2}\left(\dfrac{\alpha b^{-1}S_u^{2b^{-1}}\xi_u}{\left(\alpha S_u^{2b^{-1}}+\beta\right)}\right)\right\}} \end{array}\right). \quad (5.3.4)$$

Simplifying and applying Taylor series on Eq. (5.3.4) up to first order, we have

$$t_{eji}^{*}\approx\left(S_y^2+S_z^2\xi_z\right)exp\left(-\left(\begin{array}{l} \dfrac{M_1}{bm_1}\psi_i\xi_x\left\{1-\left(1-\dfrac{1}{m_1}\right)\dfrac{1}{b}\psi_i\xi_x\right\} \\ +\dfrac{M_2}{bm_2}\psi_j\xi_u\left\{1-\left(1-\dfrac{1}{m_2}\right)\dfrac{1}{b}\psi_j\xi_u\right\} \end{array}\right)\right), \quad (5.3.5)$$

where $\psi_i=\dfrac{\alpha S_x^{2b^{-1}}}{\left(\alpha S_x^{2b^{-1}}+\beta\right)}$ and $\psi_j=\dfrac{\alpha S_u^{2b^{-1}}}{\left(\alpha S_u^{2b^{-1}}+\beta\right)}$.

or,

$$t_{eji}^{*} \approx \left(S_y^2 + S_z^2\xi_z\right) exp\left(-\frac{1}{b}\left(\begin{array}{r}\left\{w_1\psi_i\xi_x - w_1\left(1-\frac{1}{m_1}\right)\frac{1}{b}\psi_i^2\xi_x^2\right\} \\ +\left\{w_2\psi_j\xi_u - w_2\left(1-\frac{1}{m_2}\right)\frac{1}{b}\psi_j^2\xi_u^2\right\}\end{array}\right)\right), \tag{5.3.6}$$

Further simplifying and expanding exponential series in Eq. (5.3.6), we have

$$t_{eji}^{*} \approx \left(S_y^2 + S_z^2\xi_z\right)\left(\begin{array}{l}1-\frac{1}{b}\left(\begin{array}{l}\left\{w_1\psi_i\xi_x - w_1\left(1-\frac{1}{m_1}\right)\frac{1}{b}\psi_i^2\xi_x^2\right\} \\ +\left\{w_2\psi_j\xi_u - w_2\left(1-\frac{1}{m_2}\right)\frac{1}{b}\psi_j^2\xi_u^2\right\}\end{array}\right) \\ +\frac{1}{2b^2}\left(w_1^2\psi_i^2\xi_x^2 + w_2^2\psi_j^2\xi_u^2 + 2w_1 w_2\psi_i\psi_j\xi_x\xi_u\right)\end{array}\right), \tag{5.3.7}$$

On simplification of Eq. (5.3.7), we have

$$t_{eji}^{*} \approx \left(\begin{array}{l}\left(S_y^2 + S_z^2\xi_z\right) + \frac{1}{2b^2}S_y^2\left(w_1^2\psi_i^2\xi_x^2 + w_2^2\psi_j^2\xi_u^2 + 2w_1 w_2\psi_i\psi_j\xi_x\xi_u\right) \\ -\frac{1}{b}\left(\begin{array}{l}\left\{w_1\psi_i S_y^2\xi_x + w_1\psi_i S_z^2\xi_x\xi_z - w_1\left(1-\frac{1}{m_1}\right)\frac{1}{b}S_y^2\psi_i^2\xi_x^2\right\} \\ +\left\{w_2\psi_j S_y^2\xi_u + w_2\psi_j S_z^2\xi_u\xi_z - w_2\left(1-\frac{1}{m_2}\right)\frac{1}{b}S_y^2\psi_j^2\xi_u^2\right\}\end{array}\right)\end{array}\right), \tag{5.3.8}$$

or,

$$t_{eji}^{*}-S_y^2 \approx S_z^2\xi_z+\frac{1}{2b^2}S_y^2\left(w_1^2\psi_i^2\xi_x^2+w_2^2\psi_j^2\xi_u^2+2w_1w_2\psi_i\psi_j\xi_x\xi_u\right)$$
$$-\frac{1}{b}\left(\begin{array}{l}\left\{w_1\psi_iS_y^2\xi_x+w_1\psi_iS_z^2\xi_x\xi_z-w_1\left(1-\frac{1}{m_1}\right)\frac{1}{b}S_y^2\psi_i^2\xi_x^2\right\}\\+\left\{w_2\psi_jS_y^2\xi_u+w_2\psi_jS_z^2\xi_u\xi_z-w_2\left(1-\frac{1}{m_2}\right)\frac{1}{b}S_y^2\psi_j^2\xi_u^2\right\}\end{array}\right), \tag{5.3.9}$$

Simplifying and applying expectation on Eq. (5.3.9), we have

$$E\left(t_{eji}^{*}-S_y^2\right)\approx E\left(\begin{array}{l}S_z^2\xi_z+\frac{1}{2b^2}S_y^2\left(w_1^2\psi_i^2\xi_x^2+w_2^2\psi_j^2\xi_u^2+2w_1w_2\psi_i\psi_j\xi_x\xi_u\right)\\-\frac{1}{b}\left(\begin{array}{l}\left\{w_1\psi_iS_y^2\xi_x+w_1\psi_iS_z^2\xi_x\xi_z-w_1\left(1-\frac{1}{m_1}\right)\frac{1}{b}S_y^2\psi_i^2\xi_x^2\right\}\\+\left\{w_2\psi_jS_y^2\xi_u+w_2\psi_jS_z^2\xi_u\xi_z-w_2\left(1-\frac{1}{m_2}\right)\frac{1}{b}S_y^2\psi_j^2\xi_u^2\right\}\end{array}\right)\end{array}\right), \tag{5.3.10}$$

Final expression of bias of the proposed generalized exponential estimator given in Eq. (5.3.10) is

$$Bias\left(t_{eji}^{*}\right)\approx\theta\left(\begin{array}{l}\frac{1}{2b^2}S_y^2\left(w_1^2\psi_i^2\lambda_{040}^{*}+w_2^2\psi_j^2\lambda_{004}^{*}+2w_1w_2\psi_i\psi_j\lambda_{022}^{*}\right)\\-\frac{1}{b}\left(\begin{array}{l}\left\{w_1\psi_iS_z^2\lambda_{220}^{*}\xi_z\xi_x-w_1\left(1-\frac{1}{m_1}\right)\frac{1}{b}S_y^2\psi_i^2\lambda_{040}^{*}\right\}\\+\left\{w_2\psi_jS_z^2\lambda_{202}^{*}-w_2\left(1-\frac{1}{m_2}\right)\frac{1}{b}S_y^2\psi_j^2\lambda_{004}^{*}\right\}\end{array}\right)\end{array}\right), \tag{5.3.11}$$

or,

$$Bias\left(t_{eji}^{*}\right) \approx \theta \left(\begin{array}{l} \left(\left(1-\frac{1}{m_1}\right)+\frac{w_1}{2}\right) w_1 \frac{1}{b^2} S_y^2 \psi_i^2 \lambda_{040}^{*} \\ \qquad +\left(\left(1-\frac{1}{m_2}\right)+\frac{w_2}{2}\right) w_2 \frac{1}{b^2} S_y^2 \psi_j^2 \lambda_{004}^{*} \\ \frac{1}{2b^2} \left(\begin{array}{l} S_y^2 2 w_1 w_2 \psi_i \psi_j \lambda_{022}^{*} \\ \qquad -2bS_z^2 \left(w_1 \psi_i \lambda_{220}^{*} - w_2 \psi_j \lambda_{202}^{*}\right) \end{array} \right) \end{array} \right), \tag{5.3.12}$$

In order to derive the expression of *MSE* of the proposed generalized exponential estimator, we simplifying Eq. (5.3.6) up to first-order and get

$$t_{eji}^{*} \approx \left(S_y^2 + S_z^2 \xi_z\right) exp\left(-\frac{1}{b}\left(w_1 \psi_i \xi_x + w_2 \psi_j \xi_u\right)\right), \tag{5.3.13}$$

On expanding exponential series up to first order of Eq. (5.3.13), we have

$$t_{eji}^{*} \approx \left(S_y^2 + S_z^2 \xi_z\right)\left(1-\frac{1}{b}\left(w_1 \psi_i \xi_x + w_2 \psi_j \xi_u\right)\right), \tag{5.3.14}$$

On further simplification of Eq. (5.3.14), we have

$$t_{eji}^{*} \approx \left(S_y^2 + S_z^2 \xi_z - \frac{1}{b} S_y^2 \left(w_1 \psi_i \xi_x + w_2 \psi_j \xi_u\right)\right), \tag{5.3.15}$$

Simplifying and taking expectation on both sides of the Eq. (5.3.15), we have

$$E\left(t_{eji}^{*} - S_y^2\right)^2 \approx E\left(S_z^2 \xi_z - \frac{1}{b} S_y^2 \left(w_1 \psi_i \xi_x + w_2 \psi_j \xi_u\right)\right)^2. \tag{5.3.16}$$

or,

$$E\left(t_{eji}^{*} - S_y^2\right)^2 \approx E\left(\begin{array}{l} S_z^4 \xi_z^2 - 2\frac{1}{b} S_z^2 S_y^2 \left(w_1 \psi_i \xi_z \xi_x + w_2 \psi_j \xi_z \xi_u\right) \\ +\frac{1}{b^2} S_y^4 \left(w_1^2 \psi_i^2 \xi_x^2 + w_2^2 \psi_j^2 \xi_u^2 + 2 w_1 w_2 \psi_i \psi_j \xi_x \xi_u\right) \end{array} \right). \tag{5.3.17}$$

Final expression of *MSE* of the proposed generalized exponential estimator given in Eq. (5.3.18) is

$$MSE\left(t_{eji}^{*}\right) \approx \theta \begin{pmatrix} S_z^4\lambda_{400}^{*} - 2\dfrac{1}{b}S_z^2 S_y^2\left(w_1\psi_i\lambda_{220}^{*} + w_2\psi_j\lambda_{202}^{*}\right) \\ +\dfrac{1}{b^2}S_y^4\left(w_1^2\psi_i^2\lambda_{040}^{*} + w_2^2\psi_j^2\lambda_{004}^{*} + 2w_1w_2\psi_i\psi_j\lambda_{022}^{*}\right) \end{pmatrix}. \tag{5.3.19}$$

In order to get the minimum MSE, partially differentiating Eq. (5.3.19) with respect to f_1 and f_2 and equating to zero as

$$\frac{\partial MSE\left(t_{errh}\right)}{\partial w_1} = 0. \qquad \frac{\partial MSE\left(t_{errh}\right)}{\partial w_2} = 0.$$

Now we partially differentiating *MSE* (t_{errh}) with f_1 and equating to zero as

$$\frac{1}{b} S_y^2 \left(w_1 \psi_i \lambda_{040}^* + w_2 \psi_j \lambda_{022}^* \right) = S_z^2 \lambda_{220}^* .$$

or,

$$\left(w_1 \psi_i \lambda_{040}^* + w_2 \psi_j \lambda_{022}^* \right) = b \frac{S_z^2}{S_y^2} \lambda_{220}^* .$$

or,

$$w_1 \psi_i \lambda_{040}^* = b \Re_{zy} \lambda_{220}^* - w_2 \psi_j \lambda_{022}^* .$$

or,

$$w_1 = \frac{\left(b \Re_{zy} \lambda_{220}^* - w_2 \psi_j \lambda_{022}^* \right)}{\psi_i \lambda_{040}^*} .$$

or,

$$\hat{m}_1 = M_1 \frac{\psi_i \lambda_{040}^*}{\left(b \Re_{zy} \lambda_{220}^* - w_2 \psi_j \lambda_{022}^* \right)} .$$

Similarly,we partially differentiating MSE (terrh) with f2 and equating to zero as

$$\frac{1}{b} S_y^2 \left(w_2 \psi_j \lambda_{004}^* + w_1 \psi_i \lambda_{022}^* \right) = S_z^2 \lambda_{202}^* .$$

or,

$$\left(w_2 \psi_j \lambda_{004}^* + w_1 \psi_i \lambda_{022}^* \right) = b \frac{S_z^2}{S_y^2} \lambda_{202}^* .$$

or,

$$w_2 \psi_j \lambda_{004}^* = b \Re_{zy} \lambda_{202}^* - w_1 \psi_i \lambda_{022}^* .$$

or,

$$w_2 = \frac{\left(b \Re_{zy} \lambda_{202}^* - w_1 \psi_i \lambda_{022}^* \right)}{\psi_j \lambda_{004}^*} .$$

or,

$$\hat{m}_2 = M_2 \frac{\psi_j \lambda_{004}^*}{\left(b \Re_{zy} \lambda_{202}^* - w_1 \psi_i \lambda_{022}^* \right)} .$$

In order to obtain the optimum values, solving w_1 and w_2 simultaneously as:

$$w_1 = \frac{\left(b\Re_{zy}\lambda^*_{220} - \psi_j\lambda^*_{022}\frac{\left(b\Re_{zy}\lambda^*_{202} - w_1\psi_i\lambda^*_{022}\right)}{\psi_j\lambda^*_{004}} \right)}{\psi_i\lambda^*_{040}}.$$

or,

$$w_1 = \frac{\left(b\Re_{zy}\psi_j\lambda^*_{220}\lambda^*_{004} - b\Re_{zy}\psi_j\lambda^*_{022}\lambda^*_{202} + w_1\psi_i\psi_j\lambda^{*2}_{022}\right)}{\psi_i\psi_j\lambda^*_{040}\lambda^*_{004}}.$$

or,

$$w_1\left(\lambda^*_{040}\lambda^*_{004} - \lambda^{*2}_{022}\right)\psi_i\psi_j = b\Re_{zy}\psi_j\left(\lambda^*_{220}\lambda^*_{004} - \lambda^*_{022}\lambda^*_{202}\right).$$

or,

$$w_1 = \frac{b\Re_{zy}}{\psi_i}\frac{\left(\lambda^*_{220}\lambda^*_{004} - \lambda^*_{022}\lambda^*_{202}\right)}{\left(\lambda^*_{040}\lambda^*_{004} - \lambda^{*2}_{022}\right)}.$$

or,

$$\hat{m}_1 = \frac{M_1\psi_i}{b\Re_{zy}}\frac{\left(\lambda^*_{040}\lambda^*_{004} - \lambda^{*2}_{022}\right)}{\left(\lambda^*_{220}\lambda^*_{004} - \lambda^*_{022}\lambda^*_{202}\right)} = \hat{A}_1.$$

Similarly,

$$w_2 = \frac{\left(b\Re_{zy}\psi_i\lambda^*_{040}\lambda^*_{202} - b\Re_{zy}\psi_i\lambda^*_{022}\lambda^*_{220} + w_2\psi_i\psi_j\lambda^{*2}_{022}\right)}{\psi_i\psi_j\lambda^*_{040}\lambda^*_{004}}.$$

or

$$w_2 = \frac{\left(b\Re_{zy}\psi_i\lambda^*_{040}\lambda^*_{202} - b\Re_{zy}\psi_i\lambda^*_{022}\lambda^*_{220} + w_2\psi_i\psi_j\lambda^{*2}_{022}\right)}{\psi_i\psi_j\lambda^*_{040}\lambda^*_{004}}.$$

or,

$$w_2\left(\lambda^*_{040}\lambda^*_{004} - \lambda^{*2}_{022}\right)\psi_i\psi_j = b\Re_{zy}\psi_i\left(\lambda^*_{202}\lambda^*_{040} - \lambda^*_{022}\lambda^*_{220}\right).$$

or,

$$w_2 = \frac{b\Re_{zy}}{\psi_j}\frac{\left(\lambda^*_{202}\lambda^*_{040} - \lambda^*_{022}\lambda^*_{220}\right)}{\left(\lambda^*_{040}\lambda^*_{004} - \lambda^{*2}_{022}\right)}.$$

or,

$$\hat{m}_2 = \frac{M_2\psi_j}{b\Re_{zy}} \frac{\left(\lambda^*_{040}\lambda^*_{004} - \lambda^{*2}_{022}\right)}{\left(\lambda^*_{202}\lambda^*_{040} - \lambda^*_{022}\lambda^*_{220}\right)} = \hat{A}_2.$$

Substituting "w_1"and "w_2" in Eq. (5.3.19), we have the minimum *MSE* of the proposed generalized exponential estimator having two auxiliary variables is.

$$MSE\left(t^*_{eji}\right) \approx \theta \begin{pmatrix} S_z^4\lambda^*_{400} - 2\frac{1}{b}S_z^2S_y^2\left(\hat{A}_1\psi_i\lambda^*_{220} + \hat{A}_2\psi_j\lambda^*_{202}\right) \\ +\frac{1}{b^2}S_y^4\left(\hat{A}_1^2\psi_i^2\lambda^*_{040} + \hat{A}_2^2\psi_j^2\lambda^*_{004} + 2\hat{A}_1\hat{A}_2\psi_i\psi_j\lambda^*_{022}\right) \end{pmatrix}. \tag{5.3.20}$$

or,

$$MSE\left(t^*_{eji}\right) \approx \theta\left(S_z^4\lambda^*_{400} - 2\frac{1}{b}S_z^2S_y^2B_1 + \frac{1}{b^2}S_y^4B_2\right),$$

where

$$B_1 = \left(\hat{A}_1\psi_i\lambda^*_{220} + \hat{A}_2\psi_j\lambda^*_{202}\right).$$

$$B_2 = \left(\hat{A}_1^2\psi_i^2\lambda^*_{040} + \hat{A}_2^2\psi_j^2\lambda^*_{004} + 2\hat{A}_1\hat{A}_2\psi_i\psi_j\lambda^*_{022}\right).$$

The special cases of the proposed generalized exponential estimators are summarized in Appendix B.

5.4 MATHEMATICAL COMPARISONS

In this section, we mathematically compared the proposed exponential estimator having two auxiliary variables with the usual sample variance, classical ratio and exponentialratio estimatorsbased on sensitive study variable in the presence of two auxiliary variables under *SRS* design.

- The proposed exponential estimator with two auxiliary variables will perform better than the sample variance estimator if

$$MSE_{min}\left(t^*_{eji}\right) < var\left(t^*_y\right).$$

which implies that,

$$\theta\left(S_z^4\lambda^*_{400} - 2\frac{1}{b}S_z^2S_y^2B_1 + \frac{1}{b^2}S_y^4B_2\right) < \theta S_z^4\lambda^*_{400}.$$

or,

$$\Re_{zy} > \frac{B_2}{2B_1}.$$

- The proposed exponential estimator with two auxiliary variableswill perform better than the classical ratio estimator if

$$MSE_{min}\left(t_{eji}^*\right) < MSE\left(t_r^*\right).$$

which implies that,

$$\theta\left(S_z^4\lambda_{400}^* - 2\frac{1}{b}S_z^2S_y^2B_1 + \frac{1}{b^2}S_y^4B_2\right) < \theta S_z^4\lambda_{400}^*$$
$$+\theta S_y^4\left[\lambda_{040}^* + \lambda_{004}^* + 2\lambda_{022}^* - 2\Re_{zy}\left(\lambda_{220}^* + \lambda_{202}^*\right)\right].$$

or,

$$B_2 < b^2\left[\lambda_{040}^* + \lambda_{004}^* + 2\lambda_{022}^* - 2\Re_{zy}\left(\lambda_{220}^* + \lambda_{202}^*\right)\right] - 2\Re_{zy}B_1.$$

- The proposed exponential estimator with two auxiliary variableswill perform better than the exponential ratio estimator if

$$MSE_{min}\left(t_{eji}^*\right) < MSE\left(t_{e1i}^*\right).$$

which implies that,

$$\theta\left(S_z^4\lambda_{400}^* - 2\frac{1}{b}S_z^2S_y^2B_1 + \frac{1}{b^2}S_y^4B_2\right) < \theta S_z^4\lambda_{400}^*$$
$$+\theta\frac{1}{4}S_y^4\left[\lambda_{040}^* + \lambda_{004}^* + 2\lambda_{022}^* - 4\Re_{zy}\left(\lambda_{220}^* + \lambda_{202}^*\right)\right].$$

or,

$$B_2 < \frac{1}{4}\left[b^2\left(\lambda_{040}^* + \lambda_{004}^* + 2\lambda_{022}^* - 4\Re_{zy}\left(\lambda_{220}^* + \lambda_{202}^*\right)\right) - 8\Re_{zy}B_1\right].$$

5.5 SIMULATION STUDY

In this section, we have conducted asimulation study to defend the mathematical properties and to check the performance of the proposed generalized exponential estimator numerically over the existing estimators considered in this research. We have generated artificial populations by multivariate normal distribution with different mean and variance-covariance matrices. We compared the proposed exponential estimator and its special cases with the usual sensitive variance estimator and ratio estimator. We considered a bi-variate population of size $N = 1000$ with the same theoretical means of Y, X and U as $\mu = [5\ 5\ 5]$ but different covariance matrices.The scrambling variable follows normal distribution with zero mean and standard deviation is 0.1 of the standard deviation of X. The covariance matrix with correlation coefficientsis defined as

$$\sum = \begin{bmatrix} 6 & 3 & 2.9 \\ & 2 & 1.1 \\ & & 2 \end{bmatrix}.$$

$$\rho_{yx} = 0.8661, \rho_{yu} = 0.8447$$

The efficiency of the proposed estimators is evaluated on the basis of absolute bias (AB), MSE and Percentage Relative Efficiency (PRE). The formulas of AB,$MSE$and $PRE$are defined as

$$AB(\Omega) = \left| \frac{1}{q} \sum_{i=1}^{q} \left(\Omega - S_y^2 \right) \right|.$$

$$MSE(\Omega) = \frac{1}{q} \sum_{i=1}^{q} \left(\Omega - S_y^2 \right)^2.$$

and

$$PRE(\Omega) = \frac{var\left(S_y^2\right)}{MSE(\Omega)}.$$

where

$$\Omega = t_y^*,\ t_r^*,\ t_{eji}^*,\ t_{e1i}^*,\ t_{e2i}^*,\ t_{e3i}^*\ ,...,\ t_{e12i}^*.$$

The following steps have been coded in R-language to compute the AB and MSE of all the estimators considered in this study.

- **Step 1:** From the described populations, we computed the population variances of the auxiliary variable $x$and u.
- **Step 2:** We consider different sample sizes to choose the samples using $SRSWOR$.

- **Step 3:** For each sample size, the values of the proposed generalized exponential estimator, its special cases and the existing estimators considered in this study are computed.

- **Step 4:** The process in Step-2 and Step-3 is repeated 20,000 times and the results of *AB*,*MSE*, and *PRE*are reported in Table-5.1.

Table 5.1: Theoretical and Empirical *AB*, *MSE* and *PRE* of all the Estimators for Populations-I

			Empirical		*Theoretical*		
N	n	***Estimators***	***AB***	***MSE***	***AB***	***MSE***	***PRE***
1000	50	t_y^*	0.024725	1.274042	0.022383	1.228770	100.0000
		t_r^*	0.178764	1.158901	0.176422	1.113629	110.3393
		t_{eji}^*	0.045232	0.315331	0.045032	0.315419	391.3421
		t_{e1i}^*	0.008642	0.36701	0.006300	0.321738	381.9162
		t_{e2i}^*	0.026369	0.619493	0.024027	0.574221	213.9888
		t_{e3i}^*	0.155734	1.018609	0.153392	0.973337	126.243
		t_{e4i}^*	0.029631	0.894995	0.027289	0.849723	144.6082
		t_{e5i}^*	0.006921	0.563025	0.004579	0.517753	237.3272
		t_{e6i}^*	0.023113	0.850953	0.020771	0.805681	152.5132
		t_{e7i}^*	0.105623	0.411635	0.103281	0.366363	335.3963
		t_{e8i}^*	0.026898	1.045083	0.024556	0.999811	122.9002
		t_{e9i}^*	0.048174	0.964296	0.045832	0.919024	133.7037
		t_{e10i}^*	0.048433	0.361331	0.046091	0.316059	388.778
		t_{e11i}^*	0.033112	0.887381	0.030770	0.842109	145.9157
		t_{e12i}^*	0.039647	0.616677	0.037305	0.571405	215.0436
1000	100	t_y^*	0.025153	0.699971	0.022811	0.654699	100.0000
		t_r^*	0.126453	0.510358	0.124111	0.465086	140.7692
		t_{eji}^*	0.010223	0.143241	0.010132	0.143421	459.3435
		t_{e1i}^*	0.161011	3.560493	0.158669	3.515221	18.6247
		t_{e2i}^*	0.039984	0.189976	0.037642	0.144704	452.4377
		t_{e3i}^*	0.024103	0.344456	0.021761	0.299184	218.8282
		t_{e4i}^*	0.120273	0.481479	0.117931	0.436207	150.0889
		t_{e5i}^*	0.022493	0.496808	0.020151	0.451536	144.9937
		t_{e6i}^*	0.038043	0.310368	0.035701	0.265096	246.9667
		t_{e7i}^*	0.025631	0.472149	0.023289	0.426877	153.3694

N	n	Estimators	Empirical AB	Empirical MSE	Theoretical AB	Theoretical MSE	PRE
		t^*_{e8i}	0.086642	0.209725	0.084300	0.164453	398.1057
		t^*_{e9i}	0.0239	0.577696	0.021558	0.532424	122.9656
		t^*_{e10i}	0.019863	0.462892	0.017521	0.417620	156.769
		t^*_{e11i}	0.014327	0.189361	0.011985	0.144089	454.3707
		t^*_{e12i}	0.020804	0.493089	0.018462	0.447817	146.1978
1000	200	t^*_{y}	0.007829	0.329583	0.005487	0.284311	100.0000
		t^*_{r}	0.036855	0.254852	0.034513	0.209580	135.6574
		t^*_{eji}	0.012345	0.053432	0.012003	0.050328	558.4342
		t^*_{e1i}	0.147066	2.321397	0.144724	2.276125	12.49103
		t^*_{e2i}	0.004276	0.096716	0.001934	0.051444	552.6608
		t^*_{e3i}	0.010065	0.164951	0.007723	0.119679	237.56
		t^*_{e4i}	0.035553	0.248798	0.033211	0.203526	139.6924
		t^*_{e5i}	0.009955	0.235152	0.007613	0.189880	149.7316
		t^*_{e6i}	0.003787	0.149315	0.001445	0.104043	273.2628
		t^*_{e7i}	0.00861	0.223673	0.006268	0.178401	159.3661
		t^*_{e8i}	0.022477	0.104003	0.020135	0.058731	484.0891
		t^*_{e9i}	0.00892	0.27266	0.006578	0.227388	125.0334
		t^*_{e10i}	0.016896	0.244704	0.014554	0.199432	142.5606
		t^*_{e11i}	0.016345	0.09669	0.014003	0.051418	552.9352
		t^*_{e12i}	0.010764	0.233481	0.008422	0.188209	151.0612

From the resultssummarized in Table-5.1, we observed that the proposed class of exponential estimator performed better than the competing estimators. Moreover, the precision of the proposed class of exponential estimator increases by the increase of the sample sizes.“The proposed exponential estimator t^*_{e11i} performed well than the existing estimators and the other members of the proposed class of exponential estimators but the generalized estimator found to be superior among all the estimators considered in this study. Thus, from the results, we conclude that the proposed class of exponential estimator is more efficient over the non-exponential estimators for estimating the population variance of a sensitive study variable in the presence of two auxiliary variables under*SRS* design”.

5.6 CONCLUSION

In this chapter, we have considered the generalized exponential estimator of a sensitive study variable with two auxiliary variables for estimating the finite population variance based on additive scrambled response model. We examine the performance of proposed and existing estimators using theoretical comparisons and simulation study. From the findings, we concluded that the proposed generalized exponential estimator performed efficiently than the existing estimators having two auxiliary variables. Overall, we may conclude that the proposed estimator is efficient and useful for real-life surveys.

CHAPTER 6:

CONCLUSION AND RECOMMENDATION

6.1 CONCLUSION

"There are many situations in real-life surveys when the study variable is sensitive or confidential in nature. The questions based on such sensitive variables can raise the refusal to the answer or fabricated answers given intentionally. The estimates obtained from a direct survey on sensitive questions would be subject to high response bias.In this research, we have proposed an exponential estimator for sensitive study variable in the presence of auxiliary variable. The exponential estimators can produce families of estimators using different choices (real numbers or the parameters of the auxiliary variable) of scalar constants". We have also proposed exponential estimator using two auxiliary variables using randomized response technique. The properties of the proposedestimators are derived up to the second-order approximation using well-known Taylor and exponential series. The equations of minimum *MSE* of the proposed exponential estimators are obtained by differentiating the equations of *MSE* with respect to the optimized parameters. Various families and sub-families of the proposed estimators are obtained using different choices of generalized constants. The optimum conditions are also obtained in which the proposed estimators are compared over the classical variance and ratio estimators.

A broad numerical study is conducted using real and artificial populations generated through bivariate normal distribution. From the numerical findings, it is concluded that the exponential estimators perform outstanding for both real and artificial populations on different sample sizes for the estimation of scrambled responses under certain conditions. The exponential ratio type estimators perform better for the population having positive correlation with the supporting variable whereas the exponential product estimators perform well for negatively correlated population. Accordingly, on the basis of numerical study, we may conclude that the proposed exponential estimators are superior to the classical variance estimator at different levels of correlation for the estimation of finite population variance under *SRS* design.

In this study, we have considered the exponential-type estimators for estimating the population variance of a sensitive study variable using randomized responses having single scrambling variable. In future, we will considergeneralized estimator based on exponential and non-exponential functions and different randomized response models with two scrambling variables in the presence of multi-auxiliary variables.

REFERENCES

1. Arnab, R. and Dorffner, G. (2006). Randomized response technique for complex survey designs, *Statistical Papers*, 48: 131-141.

2. Das, A.K. and Tripathi, T.P. (1978). Use of auxiliary information in estimating the finite population variance. *Sankhya C*, 40: 139-148.

3. Diana, G., and Perri, P. F. (2011). A class of estimators for quantitative sensitive data. *Statistical Papers*, *52*(3): 633-650.

4. Eichhron, B.H. and Hayre, L.S. (1983). Scrambled randomized response methods for obtaining sensitive quantitative data. *Journal of Statistical Planning and Inference,* 7: 307-316.

5. Greenberg, B.G., Abul-Ela A.L.A., Simmons W.R. and Horvitz D.G. (1969). The unrelated question randomized response model: theoretical framework. *Journal of the American Statistical Association*, 64(1): 520-539.

6. Greenberg, B.G., Kuebler, R.R. Jr., Abernathy, J.R. and Hovertiz, D.G. (1971). Application of the randomized response techniques in obtaining quantitative data. *Journal of the American Statistical Association*, 66: 243-250.

7. Grover, L.K. (2007). A wide and efficient class of estimators of population variance under sub-sampling scheme. *Model Assisted Statistics and Applications IOS Press*, 2: 41-51.

8. Gupta R.K., Singh, S. and Mangat, N. S. (1992). Some chain ratio-type estimators for estimating finite population variance. *Aligarh Journal of Statistics*, 12: 65-69.

9. Gupta, S. N., Gupta, B.C. and Singh, S. (2002). Estimation of sensitivity level of personal interview survey questions. *Journal of Statistical Planning and Inference*, 100 (2): 239–47.

10. Gupta, S.N., Thornton, B., Shabbir, J. and Singhal, S. (2006). A comparison of multiplicative and additive optional RRT models, *Journal of Statistical Theory and Applications,* 5(3): 226-239.

11. Gupta, S. and Shabbir, J. (2004). Sensitivity estimation for personal interview survey questions. *Statistica*, 64: 643-653.

12. Gupta, S., Shabbir, J. and Sehra, S. (2010). Mean and sensitivity estimation in optional randomized response models. *Journal of Statistical Planning and Inference*, 140(10): 2870-2874.

13. Gupta, S., Shabbir, J., Sousa, R., and Real, P. C., (2012). Estimation of the mean of a sensitive variable in the presence of auxiliary information. *Communications in Statistics-Theory and Methods,* 41: 1-12.

14. Huang, K.C. (2010). Unbiased estimators of mean, variance and sensitivity level for quantitative characteristics in finite population sampling, *Metrika*, 71: 341-352.

15. Isaki, C.T. (1983). Variance estimation using auxiliary information. *Journal of American Statistical Association*. 381: 117-123.

16. Kalucha, G., Gupta, S. and Shabbir, J. (2016): A Two-Step Approach to Ratio and Regression Estimation of Finite Population Mean Using Optional Randomized Response Models. *Hacettepe Journal of Mathematics and Statistics*, 45(6): 1819 – 1830.

17. Khalil, S., Noor-ul-Amin, M. and Hanif, M. (2018). Estimation of population mean for a sensitive variable in the presence of measurement error, *Journal of Statistics and Management Systems,* 21(1): 81-91.

18. Khalil, S., Zhang, Q., and Gupta, S. (2019). Mean estimation of sensitive variables under measurement errors using optional RRT models. *Communications in Statistics-Simulation and Computation*, 1-10.

19. Koyuncu, N., Gupta, S. and Sousa, R. (2014). Exponential type estimators of the mean of a sensitive variable in the presence of non-sensitive auxiliary information. *Communications in Statistics–Simulation and Computation,* 43 (7): 1583–94.

20. Perri, P. F. (2008). Modified randomized devices for Simmons' model. *Model Assisted Statistics and Applications,* 3: 233–9.

21. Pollock, K. H., and Y. Beck. (1976). A comparison of three randomized response models for quantitative data. *Journal of the American Statistical Association,* 71(356): 884–6.

22. Saha, A. (2008). A randomized response technique for quantitative data under unequal probability sampling. *Journal of Statistical Theory and Practice*, 2(4): 589–96. doi:10.1080/15598608.2008.10411897.

23. Saleem, I., Sanaullah, A. and Hanif, M. (2019): Double-sampling regression-cum-exponential estimator of the mean of a sensitive variable, *Mathematical Population Studies*, DOI: 10.1080/08898480.2019.1565273.

24. Samiuddin, A. and Hanif, M. (2007). Estimation of population mean in single phase and two-phase sampling with or without additional information. *Pakistan Journal of Statistics*, 23(2): 99-118.

25. Singh, M.P. (1967). Ratio-cum-product method of estimation. *Metrika*, 12: 34-42.

26. Singh, V. K. and Singh, G. N. (1991). Chain-type regression estimators with two auxiliary variables under double sampling scheme. *Metron-International Journal of Statistics*, 49: 279-289.

27. Sousa, R., Shabbir ,J., Real, P.C., and Gupta, S. (2010). Ratio Estimation of the Mean of a Sensitive Variable in the Presence of Auxiliary Information. *Journal of Statistical Theory and Practice*, 4(3): 495-507.

28. Warner, S.L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60(309): 63-69.

APPENDIX-A:

Some ratio and product type exponential estimators are mentioned below using different values of constants. Each estimator given in Table A1 can produce a family of estimators using different choices of α and β.

A1. Special Cases of Proposed Exponential Estimators

Estimators	***F***	***B***	***α***	***β***	***f***
$t_y = \left(s_z^2 - S_s^2\right).$	0	1	1	0	2
Exponential Ratio Type Estimators					
$t_{e1i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right).$	1	1	1	0	2
$t_{e2i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x - s_x}{S_x + s_x}\right).$	1	2	1	0	2
$t_{e3i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x^4 - s_x^4}{S_x^4 + s_x^4}\right).$	1	½	1	0	2
$t_{e4i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{\sqrt{S_x} - \sqrt{s_x}}{\sqrt{S_x} + \sqrt{s_x}}\right).$	1	4	1	0	2
$t_{e5i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x^2 - s_x^2}{S_x^2 + (5/2) s_x^2}\right).$	1	1	1	0	7/2
$t_{e6i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x - s_x}{S_x + (5/2) s_x}\right).$	1	2	1	0	7/2
$t_{e7i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x^4 - s_x^4}{S_x^4 + (5/2) s_x^4}\right).$	1	1/2	1	0	7/2
$t_{e8i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{\sqrt{S_x} - \sqrt{s_x}}{\sqrt{S_x} + (5/2)\sqrt{s_x}}\right).$	1	4	1	0	7/2
$t_{e9i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x^2 - s_x^2}{S_x^2}\right).$	1	1	1	0	1
$t_{e10i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x - s_x}{S_x}\right).$	1	2	1	0	1
$t_{e11i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x^4 - s_x^4}{S_x^4}\right).$	1	1/2	1	0	1
$t_{e12i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{\sqrt{S_x} - \sqrt{s_x}}{\sqrt{S_x}}\right).$	1	4	1	0	1

Estimators	***F***	***B***	***α***	***β***	***f***
Exponential Product Type Estimators					
$t_{e13i}=\left(s_z^2-S_s^2\right)exp\left(\frac{s_x^2-S_x^2}{S_x^2+s_x^2}\right).$	-1	1	1	0	2
$t_{e14i}=\left(s_z^2-S_s^2\right)exp\left(\frac{s_x-S_x}{S_x+s_x}\right).$	-1	2	1	0	2
$t_{e15i}=\left(s_z^2-S_s^2\right)exp\left(\frac{s_x^4-S_x^4}{S_x^4+s_x^4}\right).$	-1	1/2	1	0	2
$t_{e16i}=\left(s_z^2-S_s^2\right)exp\left(\frac{\sqrt{s_x}-\sqrt{S_x}}{\sqrt{S_x}+\sqrt{s_x}}\right).$	-1	4	1	0	2
$t_{e17i}=\left(s_z^2-S_s^2\right)exp\left(\frac{s_x^2-S_x^2}{S_x^2+(1/2)s_x^2}\right).$	-1	1	1	0	3/2
$t_{e18i}=\left(s_z^2-S_s^2\right)exp\left(\frac{s_x-S_x}{S_x+(1/2)s_x}\right).$	-1	2	1	0	3/2
$t_{e19i}=\left(s_z^2-S_s^2\right)exp\left(\frac{s_x^4-S_x^4}{S_x^4+(1/2)s_x^4}\right).$	-1	1/2	1	0	3/2
$t_{e20i}=\left(s_z^2-S_s^2\right)exp\left(\frac{\sqrt{s_x}-\sqrt{S_x}}{\sqrt{S_x}+(1/2)\sqrt{s_x}}\right).$	-1	4	1	0	3/2
$t_{e21i}=\left(s_z^2-S_s^2\right)exp\left(\frac{s_x^2-S_x^2}{S_x^2}\right).$	-1	1	1	0	1
$t_{e22i}=\left(s_z^2-S_s^2\right)exp\left(\frac{s_x-S_x}{S_x}\right).$	-1	2	1	0	1
$t_{e23i}=\left(s_z^2-S_s^2\right)exp\left(\frac{s_x^4-S_x^4}{S_x^4}\right).$	-1	1/2	1	0	1
$t_{e24i}=\left(s_z^2-S_s^2\right)exp\left(\frac{\sqrt{s_x}-\sqrt{S_x}}{\sqrt{S_x}}\right).$	-1	4	1	0	1

APPENDIX-B:

Some ratio and product type exponential estimators having two auxiliary variables are mentioned below using different values of constants. Many estimators can also be produce using different combinations or ratio-product of the auxiliary variables. Each estimator given in Table B1 can produce a family of estimators using different choices of α and β.

B1. Special Cases of Proposed Exponential Estimators

Estimators	***M_1***	***b***	***α***	***β***	***m_1***	***M_2***	***m_2***
$t_y^* = \left(s_z^2 - S_s^2\right).$	0	1	1	0	2	0	2
Exponential Ratio Type Estimators							
$t_{e1i}^* = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2} + \frac{S_u^2 - s_u^2}{S_u^2 + s_u^2}\right).$	1	1	1	0	2	1	2
$t_{e2i}^* = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x - s_x}{S_x + s_x} + \frac{S_u - s_u}{S_u + s_u}\right).$	1	2	1	0	2	1	2
$t_{e3i}^* = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x^4 - s_x^4}{S_x^4 + s_x^4} + \frac{S_u^4 - s_u^4}{S_u^4 + s_u^4}\right).$	1	1/2	1	0	2	1	2
$t_{e4i}^* = \left(s_z^2 - S_s^2\right) exp\left(\frac{\sqrt{S_x} - \sqrt{s_x}}{\sqrt{S_x} + \sqrt{s_x}} + \frac{\sqrt{S_u} - \sqrt{s_u}}{\sqrt{S_u} + \sqrt{s_u}}\right).$	1	4	1	0	2	1	2
$t_{e5i}^* = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x^2 - s_x^2}{S_x^2 + (5/2) s_x^2} + \frac{S_u^2 - s_u^2}{S_u^2 + (5/2) s_u^2}\right).$	1	1	1	0	7/2	1	7/2
$t_{e6i}^* = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x - s_x}{S_x + (5/2) s_x} + \frac{S_u - s_u}{S_u + (5/2) s_u}\right).$	1	2	1	0	7/2	1	7/2
$t_{e7i}^* = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x^4 - s_x^4}{S_x^4 + (5/2) s_x^4} + \frac{S_u^4 - s_u^4}{S_u^4 + (5/2) s_u^4}\right).$	1	1/2	1	0	7/2	1	7/2
$t_{e8i}^* = \left(s_z^2 - S_s^2\right)$ $exp\left(\frac{\sqrt{S_x} - \sqrt{s_x}}{\sqrt{S_x} + (5/2)\sqrt{s_x}} + \frac{\sqrt{S_u} - \sqrt{s_u}}{\sqrt{S_u} + (5/2)\sqrt{s_u}}\right).$	1	4	1	0	7/2	1	7/2
$t_{e9i}^* = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x^2 - s_x^2}{S_x^2} + \frac{S_u^2 - s_u^2}{S_u^2}\right).$	1	1	1	0	1	1	1
$t_{e10i}^* = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x - s_x}{S_x} + \frac{S_u - s_u}{S_u}\right).$	1	2	1	0	1	1	1
$t_{e11i}^* = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x^4 - s_x^4}{S_x^4} + \frac{S_u^4 - s_u^4}{S_u^4}\right).$	1	1/2	1	0	1	1	1
$t_{e12i}^* = \left(s_z^2 - S_s^2\right) exp\left(\frac{\sqrt{S_x} - \sqrt{s_x}}{\sqrt{S_x}} + \frac{\sqrt{S_u} - \sqrt{s_u}}{\sqrt{S_u}}\right).$	1	4	1	0	1	1	1

Estimators	***M_1***	***b***	***α***	***β***	***m_1***	***M_2***	***m_2***
Exponential Product Type Estimators							
$t^*_{e13i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{s_x^2 - S_x^2}{S_x^2 + s_x^2} + \frac{s_u^2 - S_u^2}{S_u^2 + s_u^2}\right).$	-1	1	1	0	2	-1	2
$t^*_{e14i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{s_x - S_x}{S_x + s_x} + \frac{s_u - S_u}{S_u + s_u}\right).$	-1	2	1	0	2	-1	2
$t^*_{e15i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{s_x^4 - S_x^4}{S_x^4 + s_x^4} + \frac{s_u^4 - S_u^4}{S_u^4 + s_u^4}\right).$	-1	1/2	1	0	2	-1	2
$t^*_{e16i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{\sqrt{s_x} - \sqrt{S_x}}{\sqrt{S_x} + \sqrt{s_x}} + \frac{\sqrt{s_u} - \sqrt{S_u}}{\sqrt{S_u} + \sqrt{s_u}}\right).$	-1	4	1	0	2	-1	2
$t^*_{e17i} = \left(s_z^2 - S_s^2\right)$ $exp\left(\frac{s_x^2 - S_x^2}{S_x^2 + (5/2)s_x^2} + \frac{s_u^2 - S_u^2}{S_u^2 + (5/2)s_u^2}\right).$	-1	1	1	0	7/2	-1	7/2
$t^*_{e18i} = \left(s_z^2 - S_s^2\right)$ $exp\left(\frac{s_x - S_x}{S_x + (5/2)s_x} + \frac{s_u - S_u}{S_u + (5/2)s_u}\right).$	-1	2	1	0	7/2	-1	7/2
$t^*_{e19i} = \left(s_z^2 - S_s^2\right)$ $exp\left(\frac{s_x^4 - S_x^4}{S_x^4 + (5/2)s_x^4} + \frac{s_u^4 - S_u^4}{S_u^4 + (5/2)s_u^4}\right).$	-1	1/2	1	0	7/2	-1	7/2
$t^*_{e20i} = \left(s_z^2 - S_s^2\right)$ $exp\left(\frac{\sqrt{s_x} - \sqrt{S_x}}{\sqrt{S_x} + (5/2)\sqrt{s_x}} + \frac{\sqrt{s_x} - \sqrt{S_x}}{\sqrt{S_x} + (5/2)\sqrt{s_x}}\right).$	-1	4	1	0	7/2	-1	7/2
$t^*_{e21i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{s_x^2 - S_x^2}{S_x^2} + \frac{s_u^2 - S_u^2}{S_u^2}\right).$	-1	1	1	0	1	-1	1
$t^*_{e22i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{s_x - S_x}{S_x} + \frac{s_u - S_u}{S_u}\right).$	-1	2	1	0	1	-1	1
$t^*_{e23i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{s_x^4 - S_x^4}{S_x^4} + \frac{s_u^4 - S_u^4}{S_u^4}\right).$	-1	1/2	1	0	1	-1	1
$t^*_{e24i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{\sqrt{s_x} - \sqrt{S_x}}{\sqrt{S_x}} + \frac{\sqrt{s_u} - \sqrt{S_u}}{\sqrt{S_u}}\right).$	-1	4	1	0	1	-1	1

Estimators	***M_1***	***b***	***α***	***β***	***m_1***	***M_2***	***m_2***
Exponential Ratio-Product Estimators							
$t^*_{e25i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2} + \frac{s_u^2 - S_u^2}{S_u^2 + s_u^2}\right).$	1	1	1	0	2	-1	2
$t^*_{e26i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x - s_x}{S_x + s_x} + \frac{s_u - S_u}{S_u + s_u}\right).$	1	2	1	0	2	-1	2
$t^*_{e27i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x^4 - s_x^4}{S_x^4 + s_x^4} + \frac{s_u^4 - S_u^4}{S_u^4 + s_u^4}\right).$	1	1/2	1	0	2	-1	2
$t^*_{e28i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{\sqrt{S_x} - \sqrt{s_x}}{\sqrt{S_x} + \sqrt{s_x}} + \frac{\sqrt{s_u} - \sqrt{S_u}}{\sqrt{S_u} + \sqrt{s_u}}\right).$	1	4	1	0	2	-1	2
$t^*_{e29i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x^2 - s_x^2}{S_x^2 + (5/2)s_x^2} + \frac{s_u^2 - S_u^2}{S_u^2 + (5/2)s_u^2}\right).$	1	1	1	0	7/2	-1	7/2
$t^*_{e30i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x - s_x}{S_x + (5/2)s_x} + \frac{s_u - S_u}{S_u + (5/2)s_u}\right).$	1	2	1	0	7/2	-1	7/2
$t^*_{e31i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x^4 - s_x^4}{S_x^4 + (5/2)s_x^4} + \frac{s_u^4 - S_u^4}{S_u^4 + (5/2)s_u^4}\right).$	1	1/2	1	0	7/2	-1	7/2
$t^*_{e32i} = \left(s_z^2 - S_s^2\right)$ $exp\left(\frac{\sqrt{S_x} - \sqrt{s_x}}{\sqrt{S_x} + (5/2)\sqrt{s_x}} + \frac{\sqrt{s_u} - \sqrt{S_u}}{\sqrt{S_u} + (5/2)\sqrt{s_u}}\right).$	1	4	1	0	7/2	-1	7/2
$t^*_{e33i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x^2 - s_x^2}{S_x^2} + \frac{s_u^2 - S_u^2}{S_u^2}\right).$	1	1	1	0	1	-1	1
$t^*_{e34i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x - s_x}{S_x} + \frac{s_u - S_u}{S_u}\right).$	1	2	1	0	1	-1	1
$t^*_{e35i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{S_x^4 - s_x^4}{S_x^4} + \frac{s_u^4 - S_u^4}{S_u^4}\right).$	1	1/2	1	0	1	-1	1
$t^*_{e36i} = \left(s_z^2 - S_s^2\right) exp\left(\frac{\sqrt{S_x} - \sqrt{s_x}}{\sqrt{S_x}} + \frac{\sqrt{s_u} - \sqrt{S_u}}{\sqrt{S_u}}\right).$	1	4	1	0	1	-1	1

Estimators	***M₁***	***b***	***α***	***β***	***m₁***	***M₂***	***m₂***
Exponential Product-Ratio Estimators							
$t^{*}_{e37i}=\left(s_z^2-S_s^2\right)exp\left(\frac{s_x^2-S_x^2}{S_x^2+s_x^2}+\frac{S_u^2-s_u^2}{S_u^2+s_u^2}\right).$	-1	1	1	0	2	1	2
$t^{*}_{e38i}=\left(s_z^2-S_s^2\right)exp\left(\frac{s_x-S_x}{S_x+s_x}+\frac{S_u-s_u}{S_u+s_u}\right).$	-1	2	1	0	2	1	2
$t^{*}_{e39i}=\left(s_z^2-S_s^2\right)exp\left(\frac{s_x^4-S_x^4}{S_x^4+s_x^4}+\frac{S_u^4-s_u^4}{S_u^4+s_u^4}\right).$	-1	½	1	0	2	1	2
$t^{*}_{e40i}=\left(s_z^2-S_s^2\right)exp\left(\frac{\sqrt{s_x}-\sqrt{S_x}}{\sqrt{S_x}+\sqrt{s_x}}+\frac{\sqrt{S_u}-\sqrt{s_u}}{\sqrt{S_u}+\sqrt{s_u}}\right).$	-1	4	1	0	2	1	2
$t^{*}_{e41i}=\left(s_z^2-S_s^2\right)exp\left(\frac{s_x^2-S_x^2}{S_x^2+\left(5/2\right)s_x^2}+\frac{S_u^2-s_u^2}{S_u^2+\left(5/2\right)s_u^2}\right).$	-1	1	1	0	7/2	1	7/2
$t^{*}_{e42i}=\left(s_z^2-S_s^2\right)exp\left(\frac{s_x-S_x}{S_x+\left(5/2\right)s_x}+\frac{S_u-s_u}{S_u+\left(5/2\right)s_u}\right).$	-1	2	1	0	7/2	1	7/2
$t^{*}_{e43i}=\left(s_z^2-S_s^2\right)exp\left(\frac{s_x^4-S_x^4}{S_x^4+\left(5/2\right)s_x^4}+\frac{S_u^4-s_u^4}{S_u^4+\left(5/2\right)s_u^4}\right).$	-1	½	1	0	7/2	1	7/2
$t^{*}_{e44i}=\left(s_z^2-S_s^2\right)$ $exp\left(\frac{\sqrt{s_x}-\sqrt{S_x}}{\sqrt{S_x}+\left(5/2\right)\sqrt{s_x}}+\frac{\sqrt{S_u}-\sqrt{s_u}}{\sqrt{S_u}+\left(5/2\right)\sqrt{s_u}}\right).$	-1	4	1	0	7/2	1	7/2
$t^{*}_{e45i}=\left(s_z^2-S_s^2\right)exp\left(\frac{s_x^2-S_x^2}{S_x^2}+\frac{S_u^2-s_u^2}{S_u^2}\right).$	-1	1	1	0	1	1	1
$t^{*}_{e46i}=\left(s_z^2-S_s^2\right)exp\left(\frac{s_x-S_x}{S_x}+\frac{S_u-s_u}{S_u}\right).$	-1	2	1	0	1	1	1
$t^{*}_{e47i}=\left(s_z^2-S_s^2\right)exp\left(\frac{s_x^4-S_x^4}{S_x^4}+\frac{S_u^4-s_u^4}{S_u^4}\right).$	-1	1/2	1	0	1	1	1
$t^{*}_{e48i}=\left(s_z^2-S_s^2\right)exp\left(\frac{\sqrt{s_x}-\sqrt{S_x}}{\sqrt{S_x}}+\frac{\sqrt{S_u}-\sqrt{s_u}}{\sqrt{S_u}}\right).$	-1	4	1	0	1	1	1