Chapter 2

Combined regression preliminary test estimator in double sampling with partial information on the auxiliary variable.
2.1 Introduction

It is a well known fact that for estimating the population mean \( \mu_y \) of a random variable \( Y \), precision of the estimator can be increased when information on an auxiliary variable \( X \), highly correlated with \( Y \) is readily available on all the units of the population. When the relationship between \( Y \) and \( X \) is found to be approximately linear but does not pass through the origin, linear regression estimate may be used, and the estimator is given by (Cochran, 1977)

\[
t_i = \bar{y} + b(\mu_x - \bar{x})
\]

where \( b \) is an estimate of the change in \( y \) when \( x \) is increased by unity, \( \bar{y} \) and \( \bar{x} \) are the sample means of the variables \( Y \) and \( X \) respectively and \( \mu_x \) is the population mean of \( X \).

The regression estimator given in (2.1) requires advance knowledge about \( \mu_x \), the population mean of the auxiliary variable \( X \). When such information is lacking, double sampling technique can be utilized, wherein it is sometimes considered relatively cheaper to take a large preliminary sample in which \( x_i \) alone is measured and is used for estimating the population characteristic like mean and total. The purpose of this sample is to furnish a good estimate of \( \mu_x \). Another independent or sub-sample observes both \((x_i, y_i)\) meant to estimate \((\bar{x}, \bar{y})\) for using it in the regression estimator.

To use the linear regression estimator \( t_i \) it is usually assumed that population mean \( \mu_x \) is known. However in certain practical situation, \( \mu_x \) is not
known a prior, in which case the technique of double sampling is applied. In the first preliminary sample of size \( n' \), we measure only \( x_i \) and use it for the estimation of \( \mu_x \); in the second sample, a random sub sample of size \( n (<n') \), from the preliminary sample, we observed both \( x_i \) and \( y_i \). Under double sampling the regression estimate (2.1) becomes

\[
t_2 = \bar{y}_n + b(\bar{x}_n - \bar{x}_n') \quad \text{...........}(2.2)
\]

where \( \bar{x}_n' \) is the mean of \( x_i \) in the first sample and \( (\bar{x}_n, \bar{y}_n) \) are the means of \( x_i \) and \( y_i \) in the second sample and \( b \) is the least square regression coefficient of \( Y \) on \( X \) which can be computed from the second sample.

Han (1973) described that the precision of an estimator can be improved if an auxiliary variable is used in a regression estimator based on double sampling with partial information on auxiliary variable. Sometimes there are situations where we have partial information about the mean \( \mu_x \) of the auxiliary variable \( X \). In order to utilize the partial information, Han (1973) suggested the use of a preliminary test and constructed a preliminary test estimator using double sampling with partial information on the auxiliary variable as follows;

\[
t_3 = \begin{cases} 
(y_n - \rho \bar{x}_n) & \text{if } |\bar{x}_n' - x| \leq Z_\alpha / \sqrt{n'} \\
(y_n + \rho(\bar{x}_n - x)) & \text{if } |\bar{x}_n' - x| > Z_\alpha / \sqrt{n'} 
\end{cases} \quad \text{...........}(2.3)
\]

where \( Z_\alpha \) is the \( 100(1-\alpha/2)\% \) point of \( N(0,1) \) and \( \alpha \) is the level of significance of the preliminary test.
In estimating the population mean \( \mu_y \) of the random variable \( Y \), suppose that in addition to information on an auxiliary variable \( X \), information on yet another auxiliary variable \( Z \) is available. When \( \mu_x \) and \( \mu_z \) are not available, one can take a preliminary sample to estimate these by the use of double sampling. In such a situation an estimator using \( X \) and \( Z \) is being suggested by Mukerjee et al. (1987).

Das (1992), Das and Bez (1995) and Das (2003), suggested some preliminary test estimators for the population mean in double sampling with two auxiliary variables, alternative to the usual regression estimator, when the experimenter has partial information on one and/or both auxiliary variables.

The present work is aimed to proceed in accordance to further enhance the work done by Han (1973) and Das (1995) and several other authors to find an appropriate estimator through the use of preliminary test estimation and double sampling procedures.

### 2.2 The combined regression preliminary test estimator (CRPTE) in double sampling

It is known that in many of the large scale surveys, it is inevitable to adopt stratification for the purpose of preparing a frame from which the sample can be extracted. Stratification produces gain in precision in the estimate of the characteristics of the whole population. It consists of classifying the population units in a certain number of groups called strata and selecting samples independently from each group. The division of population into strata can be done in such a way that the values of the study variable are homogeneous
within each stratum. Stratification can also be operationally convenient and economical if the sampling frame is available in the form of sub-frames.

An appropriate estimator for the population as a whole can be obtained by suitably combining the stratum-wise estimators of the characteristics under consideration. Stratification enables that the demarcation of the strata boundaries and the allocation of the total sample size to the strata may be done so as to make the estimator most efficient from the point of view of sampling variability and cost. Though the main advantage of using stratified sampling is the possible increase in efficiency per unit of cost in estimating the population characteristics, the method is also useful in situation when estimators are required with specific margins of errors not only for population as a whole but for certain groups of units. When appropriately used, the variance of the estimated mean of the study variable Y under stratification is usually less than that of the variance under simple random sampling (Cochran, 1977).

The present study attempts to proceed in accordance to further enhance the work done by Han (1973) and Das (1995), by utilizing double sampling and use preliminary test on the partial information on auxiliary variable in stratified sampling. Cochran (1977) mentioned that as with ratio estimates, there are two types of regression estimates that can be constructed in stratified random sampling. In the first estimate, $\bar{y}_{ih}$, a separate regression estimate is computed for each stratum mean, that is,

$$\bar{y}_{ih} = \bar{y}_h + b_h (\mu - \bar{x}_h),$$

for every $h$ and with $w_h = \frac{N_h}{N}$, $\bar{y}_h = \sum w_s \bar{y}_{ish}$
where $b_h$ is the within-stratum least square estimate of $B_h$ and $w_s$ is the stratum weight, $(\bar{y}_{hs}, \bar{x}_s)$ are the stratum means of $Y$ and $X$ respectively. Further, he discussed that this estimate is appropriate when it is thought that the true regression coefficients $B_h$ vary from stratum to stratum.

In the second regression estimate, $\bar{y}_{i.rc}$, the combined regression estimate is computed as

$$\bar{y}_{i.rc} = \bar{y}_u + b(\mu_s - \bar{x}_u)$$

where

$$\bar{y}_u = \sum_s w_s \bar{y}_s \quad \text{and} \quad \bar{x}_u = \sum_s w_s \bar{x}_s$$

where $b$ is the estimate of combined regression coefficient and $w_s$ is the stratum weight. Cochran states that $\bar{y}_{i.rc}$ is appropriate when $b$, an estimate of $B_h$ are presumed to be the same in all strata.

In many of the studies done in stratification, it is known that though the within stratum regression coefficients $B_h$ might differ slightly from stratum to stratum, but as such the relationship between the pair $(X,Y)$ is always maintained. Considering the combined regression estimate, the whole population is stratified into different classes and samples are selected from each stratum by simple random sampling. The mean from each stratum are calculated and utilizing the stratum weight which is estimated by proportional allocation, the strata means are combined to obtain the desired combined regression estimate.
The combined linear regression estimator $\bar{y}_{pc}$ can be utilized under three situations. Firstly when the population mean $\mu_x$ is known as a consequence of which, the study reduces to usual combined regression method of estimation. Secondly in certain practical situations $\mu_x$ is not known a priori, in which case the technique of double sampling can be applied wherein a preliminary sample is obtained to estimate $\mu_x$ and the estimator of $\mu_y$ is given by

$$t_d = \bar{y}_{st} + b(\bar{x}_n - \bar{x}_s), \quad \text{..........................}(2.4)$$

where

$$\bar{y}_n = \sum_h w_h \bar{y}_h \quad \text{and} \quad \bar{x}_n = \sum_h w_h \bar{x}_h$$

Here $\bar{x}_n$ is the value of the mean of X obtained from the preliminary sample and is utilized to estimate $\mu_x$. Thirdly, in certain situations, the experimenter may have partial information about $\mu_x$. Under such circumstances a preliminary test estimator using double sampling procedure can be used.

In the present study, the third case will be considered where partial information about the mean of the auxiliary variable will be used. The first sample is a stratified simple random sample of size $n$ in which the pair $(x_{sn}, y_{sn})$ values are measured from $n_h$ units drawn from each stratum and consequently estimating the pair $(\bar{x}_n, \bar{y}_n)$, with $n = \sum_h n_h$. The second sample is a larger simple random sample of size $n' = n + m$ which is obtained by supplementing $m$. 
more independent units on X where only $x_i$ is measured and evaluates $\bar{x}_n$ which is utilized to estimate $\mu_z$.

In order to utilize the partial information a preliminary test is done about the hypothesis

$$H_0 : \mu_x = \mu_o \quad \text{against} \quad H_1 : \mu_x \neq \mu_o$$

where $\mu_o$ is the value obtained from the partial information. If $H_0$ is accepted then $\mu_o$ is used to replace $\mu_x$ in the regression estimator $\bar{y}_{\text{nc}}$ and if $H_0$ is rejected then the sample mean $\bar{x}_{n'}$ based on the preliminary sample is used in $\bar{y}_{\text{nc}}$.

We assume that the auxiliary variable X and the study variable Y are jointly normally distributed with parameters given by $(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$. The marginal distributions which is the distribution of the study variable Y and the auxiliary variable X follow normal distribution given as $Y \sim N(\mu_y, \sigma_y^2)$ and $X \sim N(\mu_x, \sigma_x^2)$. The regression estimator depends on weather the covariance matrix is known or not. If known, one may let $\sigma_x^2 = \sigma_y^2 = 1$ without loss of generality (WLOG). The strata population $(X_h, Y_h)$ being carved out from the parent population, are also jointly assumed to follow bivariate normal distribution with parameters written as $(\mu_{x_h}, \mu_{y_h}, \sigma_{x_h}^2, \sigma_{y_h}^2, \rho)$. The correlation coefficient within each stratum between the pair of variables $(X_h, Y_h)$ might differ slightly from strata to strata, but such a relationship of the pair $(X, Y)$ is maintained in $(X_h, Y_h)$ for every stratum h. Hence the strata
correlations are assumed to be equal to the population correlation coefficient $\rho$.

Since the population is assumed to follow normal distribution, the preliminary sample utilized to collect information on the auxiliary variable for the estimation of $\bar{x}_{n'}$ is also assumed to follow normal distribution and therefore $\bar{x}_{n'} \sim N(\mu_x, \sigma_x^2/n')$ and under the assumption $\sigma_x^2 = \sigma_y^2 = 1$, $\bar{x}_{n'} \sim N(\mu_x, 1/n')$.

Further marginal distributions of $X_h$ and $Y_h$ are also normal given as $X_h \sim N(\mu_{x_h}, \sigma_{x_h}^2)$ and $Y_h \sim N(\mu_{y_h}, \sigma_{y_h}^2)$.

For each stratum, the pair of variables $(X_h, Y_h)$ for every $h$, follows a bivariate normal distribution with mean $(\mu_{x_h}, \mu_{y_h})$ and covariance matrix given by

$$\Sigma_h = \begin{pmatrix} \sigma_{x_h}^2 & \rho \sigma_{x_h} \sigma_{y_h} \\ \rho \sigma_{x_h} \sigma_{y_h} & \sigma_{y_h}^2 \end{pmatrix}$$

The regression estimator depends on whether $\Sigma_h$ is known or not. If $\Sigma_h$ is known, one may let $\sigma_{x_h}^2 = \sigma_{y_h}^2 = 1$ (WLOG).

The stratum means are given by

$$\bar{x}_h = \frac{\sum_{i=1}^{n_h} x_{ih}}{n_h} \quad \text{and} \quad \bar{y}_h = \frac{\sum_{i=1}^{n_h} y_{ih}}{n_h}$$

are linear combination of normally distributed random variables $(X_h, Y_h)$.

Hence it can be easily observed that $\bar{x}_h$ and $\bar{y}_h$ also follow normal distribution given by

$$\bar{x}_h \sim N(\mu_{x_h}, \sigma_{x_h}^2/n_h) \quad \text{and} \quad \bar{y}_h \sim N(\mu_{y_h}, \sigma_{y_h}^2/n_h)$$
\[ x_h \sim N(\mu_x, 1/n_h) \quad \text{and} \quad y_h \sim N(\mu_y, 1/n_h) \]

under the assumption of \( \sigma_{x_h}^2 = \sigma_{y_h}^2 = 1 \)

Since it is assumed that the joint distribution of the pair (X,Y) is normal, so it follows that the joint distribution of \((x_h, y_h)\) is bivariate normal with mean as \((\mu_x, \mu_y)\) and covariance matrix given by

\[
\begin{pmatrix}
\sigma_{x_h}^2 / n_h & \rho \sigma_{x_h} \sigma_{y_h} / n_h \\
\rho \sigma_{x_h} \sigma_{y_h} / n_h & \sigma_{y_h}^2 / n_h
\end{pmatrix} = \begin{pmatrix}
1/n_h & \rho / n_h \\
\rho / n_h & 1/n_h
\end{pmatrix}
\]

In certain situation, the experimenter may have partial information about \(\mu_x\). In order to utilize the partial information, one can perform a preliminary test about the hypothesis

\[ H_0 : \mu_x = \mu_0, \quad \text{against} \quad H_1 : \mu_x \neq \mu_0 \]

where \(\mu_0\) is the value obtained from the partial information and \(\bar{x}_{n'}\) which is the value of the mean of X obtained from the preliminary sample through the use of double sampling is being used to test the hypothesis.

Now, when \(\mu_x\) is partially known, one can let \(\mu_0 = 0\) (WLOG), so that the hypothesis can be accepted when,

\[
\left| \frac{(x_{n'} - \mu_x)}{\text{SE}(x_{n'})} \right| \leq Z_{\alpha}
\]

\[
\Rightarrow \left| \frac{x_{n'}}{(1/\sqrt{n'})} \right| \leq Z_{\alpha}
\]

\[
\Rightarrow \left| x_{n'} \right| \leq Z_{\alpha} / \sqrt{n'}
\]
where \( Z_0 \) is the 100(1-\( \alpha \)/2)% point of \( N(0,1) \) and \( \alpha \) is the level of significance of the preliminary test.

Under the above assumption the Combined regression preliminary test estimator (CRPTE) in double sampling having partial information on the auxiliary variable \( X \) can be written as

\[
t_s = \begin{cases} 
(y_{st} - \rho \bar{x}_{st}) & \text{if } \left| \bar{x}_{st} \right| \leq Z_0 / \sqrt{n'} \\
(y_{st} + \rho (x_{st} - \bar{x}_{st})) & \text{if } \left| \bar{x}_{st} \right| > Z_0 / \sqrt{n'}
\end{cases}
\]  

(2.5)

where \( y_{st} = \sum_h W_h y_h \) and \( \bar{x}_{st} = \sum_h W_h x_h \)

and the regression coefficient \( b \) from \( y_{st} \) reduces to \( b = \rho (\sigma_x / \sigma_y) = \rho \) under the above assumptions.

2.3 Bias of the CRPTE

To evaluate the bias of \( t_s \), we consider that the joint distribution of \( (x_{st}, y_{st}, y_{st}) \) is a multivariate normal with mean \( (\mu_x, \mu_x, \mu_y) \) and covariance matrix given by

\[
\Sigma = \begin{pmatrix}
Var(x_{st}) & Cov(x_{st}, y_{st}) & Cov(x_{st}, y_{st}) \\
Cov(x_{st}, y_{st}) & Var(y_{st}) & Cov(y_{st}, y_{st}) \\
Cov(y_{st}, y_{st}) & Cov(y_{st}, y_{st}) & Var(y_{st})
\end{pmatrix}
\]  

(2.6)

Now

The variance of \( \bar{x}_{st} \) and \( \bar{y}_{st} \) being given by

\[
Var(\bar{x}_{st}) = \sum W_h^2 \sigma_{x, st}^2 / n_h \\
Var(\bar{y}_{st}) = \sum W_h^2 \sigma_{y, st}^2 / n_h
\]

(Cochran 1977)

which under the assumption \( \sigma_{x, st}^2 = \sigma_{y, st}^2 = 1 \), becomes
\[ \text{Var}(\bar{x}_n) = \sum W_h^2 / n_h \quad \text{Var}(\bar{y}_n) = \sum W_h^2 / n_h \]

When the samples are selected with proportional allocation then the stratum weight is given by \( W_h = (N_h / N) = (n_h / n) \)

Thus \( \sum W_h^2 / n_h = \sum W_h^2 / nW_h = (1/n) \sum W_h = (1/n) \quad (\text{as} \quad \sum W_h = 1) \)

Hence the covariance matrix (2.6) reduces to

\[
\sum = \begin{pmatrix}
1/n' & 1/n' & \rho/n' \\
1/n' & 1/n & \rho/n \\
\rho/n' & \rho/n & 1/n
\end{pmatrix}
\]

The derivation of bias of \( t_5 \) involves conditional expectations, the condition being the acceptance or rejection of the hypothesis considered in the preliminary test. Further the expectations can be obtained from the integrals involving probability density functions which are assumed to be normal. The bias of the estimator \( t_5 \) is derived as follows:

The Bias of an estimator is defined as

\[ \text{Bias}(t_5) = E(t_5) - \mu_Y \]  

where \( E(.) \) is the mathematical expectation

\[
\text{Bias}(t_5) = E \left\{ \bar{y}_{st} - \rho \bar{x}_{st} \right\} \quad \text{if} \quad |\bar{x}_{n'}| \leq Z_{\alpha} / \sqrt{n'}
\]
\[
+ E \left\{ \bar{y}_{st} + \rho (\bar{x}_{n'} - \bar{x}_{st}) \right\} \quad \text{if} \quad |\bar{x}_{n'}| > Z_{\alpha} / \sqrt{n'}
\]

\[
= E \left\{ \bar{y}_{st} - \rho \bar{x}_{st} \right\} + E \left\{ \rho \bar{x}_{n'} \right\} \quad \text{if} \quad |\bar{x}_{n'}| > Z_{\alpha} / \sqrt{n'}
\]

\[
\Rightarrow \text{Bias}(t_5) = \sum W_h \mu_{y_h} - \rho \sum W_h \mu_{x_h} + E \left\{ \rho \bar{x}_{n'} / |\bar{x}_{n'}| > Z_{\alpha} / \sqrt{n'} \right\} - \mu_Y
\]
It is given that
\[ \sum W_h \mu_{x_h} = \mu_x \quad \text{and} \quad \sum W_h \mu_{y_h} = \mu_y \] (Cochran, 1977)

Thus,
\[ \text{Bias } (t_5) = -\rho \mu_x + E\{\rho \overline{x}_n' \quad \text{if} \quad |\overline{x}_n'| > Z_\alpha / \sqrt{n'} \}
\]
\[ = -\rho \mu_x + E\{\rho \overline{x}_n' \quad \text{if} \quad \overline{x}_n' > Z_\alpha / \sqrt{n'} \} + E\{\rho \overline{x}_n' \quad \text{if} \quad \overline{x}_n' < -Z_\alpha / \sqrt{n'} \}
\]

\[ \Rightarrow \text{Bias } (t_5) = -\rho \mu_x + \rho \left\{ \int_{Z_\alpha / \sqrt{n'}}^{\infty} \overline{x}_n' f(\overline{x}_n') d\overline{x}_n' + \int_{-\infty}^{-Z_\alpha / \sqrt{n'}} \overline{x}_n' f(\overline{x}_n') d\overline{x}_n' \right\}
\]

\[ \text{..........................(2.9)} \]

where \( f(\overline{x}_n') \) is the probability density function of \( \overline{x}_n' \) which follows
\( N(\mu_x, 1/n') \) under the assumption \( \sigma^2_x = 1 \).

Let
\[ I = \int_{Z_\alpha / \sqrt{n'}}^{\infty} \overline{x}_n' f(\overline{x}_n') d\overline{x}_n' + \int_{-\infty}^{-Z_\alpha / \sqrt{n'}} \overline{x}_n' f(\overline{x}_n') d\overline{x}_n'
\]

\[ = \left( \sqrt{n'}/\sqrt{2\pi} \right) \left[ \int_{Z_\alpha / \sqrt{n'}}^{\infty} \overline{x}_n' \text{Exp} \left\{ \frac{-1/2}{(\overline{x}_n' - \mu_x)/(1/\sqrt{n'})^2} \right\} d\overline{x}_n' \right.
\]
\[ \left. + \int_{-\infty}^{-Z_\alpha / \sqrt{n'}} \overline{x}_n' \text{Exp} \left\{ \frac{-1/2}{(\overline{x}_n' - \mu_x)/(1/\sqrt{n'})^2} \right\} d\overline{x}_n' \right]
\]
Putting \( w = \frac{(x \nu - \mu_s)(1 / \sqrt{n'})}{\sqrt{n'}} \Rightarrow dw = \sqrt{n'} dx \), we have

When \( \tilde{x}_n = Z_a / \sqrt{n'} \) then \( w = \sqrt{n'} (Z_a / \sqrt{n'} - \mu_s) = Z_a - \sqrt{n'} \mu_s = A \) and

When \( \tilde{x}_n = -Z_a / \sqrt{n'} \) then \( w = \sqrt{n'} (-Z_a / \sqrt{n'} - \mu_s) = -Z_a - \sqrt{n'} \mu_s = B \)

Therefore,

\[
I = \frac{1}{\sqrt{2\pi}} \int_A^B (\mu_x + (w / \sqrt{n'})) \exp \left( \frac{-1}{2} w^2 \right) dw
\]

\[
+ \int_{-\infty}^B (\mu_x + (w / \sqrt{n'})) \exp \left( \frac{-1}{2} w^2 \right) dw
\]

\[
= \frac{1}{\sqrt{2\pi}} \mu_x \left\{ \int_A^B \exp \left( \frac{-1}{2} w^2 \right) dw + \int_{-\infty}^B \exp \left( \frac{-1}{2} w^2 \right) dw \right\}
\]

\[
+ \left( \frac{1}{\sqrt{2\pi}} \right) \left( \frac{1}{\sqrt{n'}} \right) \left\{ \int_A^B w \exp \left( \frac{-1}{2} w^2 \right) dw + \int_{-\infty}^B w \exp \left( \frac{-1}{2} w^2 \right) dw \right\}
\]

\[
= \mu_x \left\{ 1 - \Phi(A) + \Phi(B) \right\}
\]

\[
+ \left( \frac{1}{\sqrt{2\pi n'}} \right) \left\{ \int_A^B w \exp \left( \frac{-1}{2} w^2 \right) dw + \int_{-\infty}^B w \exp \left( \frac{-1}{2} w^2 \right) dw \right\}
\]

where \( \Phi(.) \) is the cumulative distribution function of \( N(0,1) \).
Again

putting $\left( w^2 / 2 \right) = t \Rightarrow wdw = dt$

\[
I = \mu_x \left\{ \Phi \left( B \right) - \Phi \left( A \right) \right\} + \frac{1}{\sqrt{2\pi}} \left\{ \int_{A^2/2}^{\infty} \exp \left( -t \right) dt + \int_{-\infty}^{B^2/2} \exp \left( -t \right) dt \right\}
\]

\[
= \mu_x \left\{ \Phi \left( B \right) - \Phi \left( A \right) \right\} + \frac{1}{\sqrt{2\pi n'}} \left\{ \left( \frac{1}{\sqrt{2\pi}} \right)^2 \exp \left( -A^2/2 \right) - \exp \left( -B^2/2 \right) \right\}
\]

\[
= \mu_x \left\{ \Phi \left( B \right) - \Phi \left( A \right) \right\} + \frac{1}{\sqrt{n'}} \left\{ \phi \left( A \right) - \phi \left( B \right) \right\}
\]

where $\phi(.)$ is the density function of $\text{N}(0,1)$.

Therefore from (2.9),

\[
\text{Bias} \left( t_s \right) = -\rho \mu_x + \rho \left\{ \mu_x \left\{ \Phi \left( A \right) - \Phi \left( B \right) \right\} + \frac{1}{\sqrt{n'}} \left\{ \phi \left( A \right) - \phi \left( B \right) \right\} \right\}
\]

\[
= -\rho \mu_x \left\{ \Phi \left( A \right) - \Phi \left( B \right) \right\} + \rho \left( \frac{1}{\sqrt{n'}} \right) \left\{ \phi \left( A \right) - \phi \left( B \right) \right\}
\]

\[
\text{...................(2.10)}
\]

2.4 Discussion

From Equation (2.10), we know that

\[
\text{Bias} \left( t_s \right) = -\rho \mu_x \left\{ \Phi \left( A \right) - \Phi \left( B \right) \right\} + \rho \left( 1/\sqrt{n'} \right) \left\{ \phi \left( A \right) - \phi \left( B \right) \right\}
\]

As partial checks we have, the following:

When $\alpha = 0$, i.e when we always accept $H_0$, then $Z_s = \infty$.
Thus \( \text{Bias}(t_z) = -\rho \mu_x \) 

When \( \alpha = 1 \), then \( Z_z = 0 \). Thus \( \text{Bias}(t_z) = 0 \) 

The values of \( \text{Bias}(t_z) \) can be easily computed for different values of \( \mu_x \). In order to get an idea about the behavior of the bias function with respect to \( \mu_x \), \( \text{Bias}(t_z) \) is computed for a set of values of \( n' \), \( \alpha \) and \( \rho \) which are represented in Table 2.1 - 2.4 and Figure 2.1 - 2.4. When the bias of the proposed estimator is computed for different values of the mean of the auxiliary variables \( \mu_x \), it can be observed (Table 2.1 and Fig 2.1) that the behavior of the bias is symmetrical about \( \mu_x = 0 \). Thus it suffices to analyze the behavior of the bias for \( \mu_x \geq 0 \). It is found in general that \( \text{Bias}(t_z) \) has minimum value zero at \( \mu_x = 0 \). As \( \mu_x \) increases, the \( \text{Bias}(t_z) \) increases to a maximum and then gradually decreases to zero. The Figures (2.1 - 2.4) clearly show that when the mean of the auxiliary variable is close to the hypothetical value, then bias is very close to 0. Also as \( \mu_x \) moves away from the hypothetical value the bias increases, but after attaining maximum again gradually reduces to zero. This establishes the utility of the present study that the use of partial information and preliminary test reduces the bias of the proposed estimator.

Further, when the parameters \( \alpha \) and \( \rho \) are fixed (Table 2.2 and Fig 2.2), then the bias is inversely proportional to the square root of the size of the preliminary sample \( n' \). Therefore it can be concluded that with the increase in the preliminary sample size, the bias decreases. However, the bias is not affected by \( n \), the size of the stratified random sample.
2.5 Bias of the CRPTE computed numerically

The above analytical method used for computing the bias of the proposed estimator involves the evaluation of mathematical expectation of the random variables and consequently results in the computation of integrals within certain limits. This may sometimes become very cumbersome, hence an alternative method for the evaluation of the bias is sought with the help of numerical techniques.

During the pre-computer era, significant amount of energy were expended on the technique to find solution, rather than on definition and interpretation of the problem. Mathematical solutions are usually derived for some problems using analytical or exact methods. These solutions were often useful and provide insight into the behavior of some systems. However, analytical solutions can be derived only for a limited class of problems. These problems include those that can be approximated with linear models and those with simple geometry and low dimensionality. Consequently, analytical solutions are often of limited practical value because most real problems are non-linear and involve complex shapes and processes. Numerical methods are extremely powerful problem solving tools. They are capable of handling large systems of equations and complicated geometries that are common in many mathematical and physical phenomena and that are often impossible to solve analytically. Through the use of numerical methods one can successively approximate both simple and complex solutions to evaluate the roots of equations, solving system of equations, evaluate differentiation and integration numerically, finding solutions to ordinary and partial differential equations and many other
mathematical approximations with great precision and accuracy (Chapra and Canale, 1989).

Today's high speed computers provide an alternative for such complicated calculations. Using numerical techniques and computers to obtained solutions directly, one can approach these calculations without recourse to simplifying assumptions or time-intensive techniques. Although analytical solutions are extremely valuable both for problem solving and for providing insight, numerical methods represent alternatives that greatly enlarge one's own capabilities to confront and solve problems. As a result more time is available to the use of creative skills. Thus, more emphasis can be placed on problem formulation and interpretation of the solution.

The function to be differentiated or integrated usually will be typically a continuous function such as a polynomial, an exponential, or a trigonometric function or some other complicated function that is difficult to differentiate or integrate analytically. Sometimes one can come across a tabulated function where values of \( x \) and \( f(x) \) are given at a number of discrete points as is often the case with experimental or field data. In such instances, analytical solutions are difficult to obtain and therefore numerical analysis can be employed.

As we may see in chapter 2, the evaluation of bias of CRPTE involves computations of definite integrals. For this we may use numerical techniques. The most common approach for numerical integrations is the Newton-Cotes formulae which are based on replacing a complicated function with a simple polynomial that is easy to integrate. Three of the most widely used
Newton-Cotes formulae are the trapezoidal rule, Simpson's 1/3 rule, and Simpson's 3/8 rule. The error in approximating an integral by Simpson's 1/3 rule is

$$\left|\frac{(b-a)^3}{2880}f^4(\xi)\right|$$

where \(f(.)\) is the function to be integrated, \(a\) and \(b\) are the limits of integration and \(\xi\) is some number between \(a\) and \(b\).

The error is (asymptotically) proportional to \((b-a)^5\). Simpson's rule gains an extra order because the points at which the integrand is evaluated are distributed symmetrically in the interval \([a, b]\). It may be noted that Simpson's rule provides exact results for any polynomial of degree three or less, since the error term involves the fourth derivative of \(f\). The approximated formulae for numerical integration by Simpson's 1/3 rule is given as follows.

$$\int_{x_0}^{x_k} f(x)dx = \frac{h}{3} \left[ f(x_0) + 4 \left\{ \sum_{i=1,3,5,...}^{k-1} f(x_i) \right\} + 2 \left\{ \sum_{i=2,4,6,...}^{k-2} f(x_i) \right\} + f(x_k) \right]$$

where \((x_i-x_{i-1})=h\) \hspace{1cm} \text{(Jain, Iyengar and Jain, 2007)}

The evaluation of integrals by numerical techniques involves the computation of the numerical values of the function \(f(x)\) at different points \(a=x_0, x_1, x_2, ..., x_k=b\). These functional values are then substituted in Simpson's rule to get an approximate value of the integral of \(f(x)\). When the number of divisions or partitions of the range \((b-a)\) increases, say when \(k > 35\)
then the behavior of the bias function (Fig 2.5) becomes smooth and converges. Thus for all numerical integration in the present work, the number of divisions $k$ is fixed at 50.

As mentioned above, the increase in the number of divisions is a necessity, and as a result computation of the integral by Simpson’s rule becomes tedious and manual exercise is practically impossible. However the availability of high speed computation facilities makes it possible to evaluate the integral in a much faster and easier way. Thus by using numerical techniques and computers to obtain the bias one can approach these calculations without recourse to simplifying assumptions or time-intensive techniques.

In the present study alternative to analytical methods, attempt is also made to evaluate the bias of the suggested estimator $t_5$ numerically as follows;

The Bias of the proposed estimator is defined as

$$
\text{Bias}(t_5) = E(t_5) - \mu_y
$$

where $E(.)$ is the mathematical expectation

$$
\text{Bias}(t_5) = E\left\{y_{st} - \rho \bar{x}_{st}\right\} \text{if } |\bar{x}_{st}| \leq \frac{Z_a}{\sqrt{n'}}
$$

$$
+ \left\{E\left\{y_{st} + \rho (\bar{x}_{st} - \bar{x}_{st})\right\} \text{if } |\bar{x}_{st}| > \frac{Z_a}{\sqrt{n'}}\right\} - \mu_y
$$

$$
= E(y_{st} - \rho \bar{x}_{st}) + \left\{E(\rho \bar{x}_{st}) \text{if } |\bar{x}_{st}| > \frac{Z_a}{\sqrt{n'}}\right\} - \mu_y
$$

It is given that

$$
\sum W_n \mu_{x_n} = \mu_x \text{ and } \sum W_n \mu_{y_n} = \mu_y \text{ (Cochran, 1977)}
$$

Thus

$$
\text{Bias} (t_5) = \mu_y - \rho \mu_x + E\{\rho \bar{x}_{st}/|\bar{x}_{st}| > \frac{Z_a}{\sqrt{n'}}\} - \mu_y
$$
\[ I = \rho \mu_x + E\{\rho \frac{x_n}{x_{n'}} > Z_a / \sqrt{n'}\} \]
\[ + E\{\rho \frac{x_n}{x_{n'}} < -Z_a / \sqrt{n'}\} \] 
\( \ldots \ldots (2.11) \)

Let,
\[ I = E\{\frac{x_n}{x_{n'}} > Z_a / \sqrt{n'}\} + E\{\frac{x_n}{x_{n'}} < -Z_a / \sqrt{n'}\} \]
\[ = \int_{Z_a / \sqrt{n'}}^{\infty} x_n' f(x_{n'}) d\bar{x}_{n'} + \int_{-\infty}^{-Z_a / \sqrt{n'}} x_n' f(x_{n'}) d\bar{x}_{n'} \]

where \( f(\bar{x}_n) \) is the probability density function of \( \bar{x}_{n'} \sim N(\mu_x, 1/n') \) under the assumption \( \sigma_x^2 = 1 \)

\[ I = (\sqrt{n'} / \sqrt{2\pi}) \int_{Z_a / \sqrt{n'}}^{\infty} \bar{x}_n' \text{Exp}\{(-1/2)((\bar{x}_n' - \mu_x)/((1/\sqrt{n'}))^2)\} d\bar{x}_{n'} \]
\[ + (\sqrt{n'} / \sqrt{2\pi}) \int_{-\infty}^{-Z_a / \sqrt{n'}} \bar{x}_n' \text{Exp}\{(-1/2)((\bar{x}_n' - \mu_x)/((1/\sqrt{n'}))^2)\} d\bar{x}_{n'} \]

Putting \( w = (\bar{x}_n - \mu_x)/(1/\sqrt{n'}) \) \( \Rightarrow dw = \sqrt{n'} d\bar{x}_n \) we have

When \( \bar{x}_n = Z_a / (1/\sqrt{n'}) \) then \( w = \sqrt{n'} (Z_a / \sqrt{n'} - \mu_x) = Z_a - \sqrt{n'} \mu_x = A \) and

When \( \bar{x}_n = -Z_a / (1/\sqrt{n'}) \) then \( w = \sqrt{n'} (-Z_a / \sqrt{n'} - \mu_x) = -Z_a - \sqrt{n'} \mu_x = B \)

\[ I = (1/\sqrt{2\pi}) \int_{A}^{B} (\mu_x + (w/\sqrt{n'})) \text{Exp}((-1/2)w^2) dw \]
\[ + (1/\sqrt{2\pi}) \int_{-\infty}^{B} (\mu_x + (w/\sqrt{n'})) \text{Exp}((-1/2)w^2) dw \]
Substituting the values of $I$ in (2.11), it follows that

$$
\text{Bias } (t_5) = -\rho \mu_x + \rho I
$$

$$
\text{Bias } (t_5) = \rho (I - \mu_x) \quad \cdots \cdots \text{(2.12)}
$$

where

$$
I = \left(1/\sqrt{2\pi}\right) \left\{ \int_{A}^{B} \{ \mu_x + (w/\sqrt{n'}) \} \exp(-0.5w^2) \, dw \right\}
$$

$$
+ \left(1/\sqrt{2\pi}\right) \left\{ \int_{-\infty}^{B} \{ \mu_x + (w/\sqrt{n'}) \} \exp(-0.5w^2) \, dw \right\}
$$

$$
I = I_1 + I_2 \quad \cdots \cdots \text{(2.13)}
$$

The programs written on Fortran 77 (Rajaraman, 1997) were used in the numerical evaluation of the above integrals $I_1$ and $I_2$. (Appendix 1 and 2)

The integrals $I_1$ and $I_2$ involve the integrand function $g(w) = \{ \mu_x + (w/\sqrt{n'}) \} \exp(-0.5w^2)$ and in both, the limits of integration has infinity at one end. Such integration is tedious when evaluated by numerical techniques. The function $g(w)$ is plotted graphically for various values of $w$ (Fig 2.6). The graph also show that the function $g(w)$ tapers to zero for $w > 3$ and $w < -3$. As we know that integration is the process of finding the area under the curve, bounded by the rectangular axis and the two ordinates corresponding to the limits of the integral, so the area under the curve $g(w)$ for the entire integrating limits defined in both $I_1$ and $I_2$ is approximately equal to that when the limits of
integration is confined to between A to 3 for $I_1$ and between -3 to B for $I_2$. The values of $I_1$ and $I_2$ are given in Table 2.5.

2.6 Discussion

The values of $Bias(t_1)$ with respect to different values of $\mu_x$ are computed by the use of numerical techniques by substituting the output values of $I_1$, depicted in table 2.5 (a), (b), (c) in equation (2.12). In order to get an idea about the behavior of the bias function with respect to $\mu_x$, $Bias(t_1)$ is computed for a set of values of $\alpha$ and $\rho$ which are depicted in Table 2.6 - 2.7 and Figure 2.7 - 2.8. It is found that $Bias(t_1)$ is zero or minimum at $\mu_x = 0$. As $\mu_x$ increases, the $Bias(t_1)$ increases to a maximum and then gradually decreases to zero. The figure shows that when the mean of the auxiliary variable obtained by partial information is close to the hypothetical value, then the bias is very close to 0.

Han(1973) and Das and Bez(1995) in their paper constructed their estimators using analytical techniques exclusively. With the advent of modern and high speed digital computers, handling of complex statistical calculations can be done with ease in a short period of time. Also the advancement in field of Numerical techniques has simplified the method of solving complex mathematical analysis like the determination of roots of equations, solving systems of linear algebraic equations, differentiations and integrations, ordinary differential equations and partial differential equations.
In the present work, an attempt is being made to construct a preliminary test estimator in double sampling through stratification of the population. The combined linear regression estimator as suggested by Cochran (1997) is being used to construct a preliminary test estimator in double sampling for the present study. Bias is calculated analytically and the results are also plotted graphically. An attempt is also made in this chapter to evaluate the bias of the above estimators using numerical techniques and compared with that obtained analytically. Fig 2.9 show that the bias obtained by numerical methods depict a pattern similar to that obtained by analytical methods for increasing values of $\mu_1$. The differences in the values of bias between analytical and numerical methods of computation are minimal.
Table 2.1 Behaviour of Bias(t₅) computed analytically with respect to μₓ for n' = 200, ρ = 0.8, α = 0.01.

<table>
<thead>
<tr>
<th>μₓ</th>
<th>0</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias</td>
<td>0.0</td>
<td>0.059</td>
<td>0.042</td>
<td>0.006</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>μₓ</td>
<td>0</td>
<td>-0.1</td>
<td>-0.2</td>
<td>-0.3</td>
<td>-0.4</td>
<td>-0.5</td>
<td>-0.6</td>
<td>-0.7</td>
<td>-0.8</td>
<td>-0.9</td>
<td>-1.0</td>
</tr>
<tr>
<td>Bias</td>
<td>0.0</td>
<td>0.059</td>
<td>0.042</td>
<td>0.006</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Table 2.2 Behaviour of Bias(t₅) computed analytically with respect to μₓ for different values of n' and for ρ = 0.8, α = 0.01.

<table>
<thead>
<tr>
<th>n'</th>
<th>μₓ</th>
<th>0</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>120</td>
<td>0</td>
<td>0.069</td>
<td>0.077</td>
<td>0.039</td>
<td>0.006</td>
<td>0.0003</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>200</td>
<td>0</td>
<td>0.055</td>
<td>0.043</td>
<td>0.006</td>
<td>0.0001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>400</td>
<td>0</td>
<td>0.044</td>
<td>0.007</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 2.3  Behaviour of Bias($t_\theta$) computed analytically with respect to $\mu_x$ for different values of $\alpha$ and for $n' = 200$, $\rho = 0.8$

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$\mu_x$</th>
<th>0</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>0</td>
<td>0</td>
<td>0.059</td>
<td>0.042</td>
<td>0.006</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.05</td>
<td>0</td>
<td>0</td>
<td>0.037</td>
<td>0.015</td>
<td>0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.25</td>
<td>0</td>
<td>0</td>
<td>0.011</td>
<td>0.002</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2.4  Behaviour of Bias($t_\theta$) computed analytically with respect to $\mu_x$ for different values of $\rho$ and for $n' = 200$, $\alpha = 0.01$

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$\mu_x$</th>
<th>0</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.7</td>
<td>0</td>
<td>0</td>
<td>0.033</td>
<td>0.013</td>
<td>0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.8</td>
<td>0</td>
<td>0</td>
<td>0.037</td>
<td>0.015</td>
<td>0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.9</td>
<td>0</td>
<td>0</td>
<td>0.042</td>
<td>0.017</td>
<td>0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 2.5 Numerically computed values of $I$ with $n' = 200$
and $\rho = 0.8$ for (a) $\alpha = 0.01$ (b) $\alpha = 0.05$ (c) $\alpha = 0.25$

(a)

<table>
<thead>
<tr>
<th>$\mu_x$</th>
<th>$I_1$</th>
<th>$I_2$</th>
<th>$I$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0010</td>
<td>-0.0010</td>
<td>0</td>
</tr>
<tr>
<td>0.1</td>
<td>0.02648</td>
<td>-6.44E-06</td>
<td>0.018939</td>
</tr>
<tr>
<td>0.2</td>
<td>0.14694</td>
<td>-6.21E-09</td>
<td>0.10346</td>
</tr>
<tr>
<td>0.3</td>
<td>0.29256</td>
<td>0</td>
<td>0.25633</td>
</tr>
<tr>
<td>0.4</td>
<td>0.3997</td>
<td>0</td>
<td>0.39278</td>
</tr>
<tr>
<td>0.5</td>
<td>0.499505</td>
<td>0</td>
<td>0.499505</td>
</tr>
<tr>
<td>0.6</td>
<td>0.59987</td>
<td>0</td>
<td>0.59987</td>
</tr>
<tr>
<td>0.7</td>
<td>0.69985</td>
<td>0</td>
<td>0.69985</td>
</tr>
<tr>
<td>0.8</td>
<td>0.79983</td>
<td>0</td>
<td>0.79983</td>
</tr>
<tr>
<td>0.9</td>
<td>0.89981</td>
<td>0</td>
<td>0.89981</td>
</tr>
<tr>
<td>1</td>
<td>0.99979</td>
<td>0</td>
<td>0.99979</td>
</tr>
</tbody>
</table>

(b)

<table>
<thead>
<tr>
<th>$\mu_x$</th>
<th>$I_1$</th>
<th>$I_2$</th>
<th>$I$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0041</td>
<td>-0.0041</td>
<td>0</td>
</tr>
<tr>
<td>0.1</td>
<td>0.0536</td>
<td>-5.81E-05</td>
<td>0.04417</td>
</tr>
<tr>
<td>0.2</td>
<td>0.1808</td>
<td>-1.26E-07</td>
<td>0.15399</td>
</tr>
<tr>
<td>0.3</td>
<td>0.2987</td>
<td>0</td>
<td>0.28734</td>
</tr>
<tr>
<td>0.4</td>
<td>0.39876</td>
<td>0</td>
<td>0.39876</td>
</tr>
<tr>
<td>0.5</td>
<td>0.49986</td>
<td>0</td>
<td>0.49986</td>
</tr>
<tr>
<td>0.6</td>
<td>0.59987</td>
<td>0</td>
<td>0.59987</td>
</tr>
<tr>
<td>0.7</td>
<td>0.69985</td>
<td>0</td>
<td>0.69985</td>
</tr>
<tr>
<td>0.8</td>
<td>0.79983</td>
<td>0</td>
<td>0.79983</td>
</tr>
<tr>
<td>0.9</td>
<td>0.89981</td>
<td>0</td>
<td>0.89981</td>
</tr>
<tr>
<td>1</td>
<td>0.99979</td>
<td>0</td>
<td>0.99979</td>
</tr>
</tbody>
</table>

(c)

<table>
<thead>
<tr>
<th>$\mu_x$</th>
<th>$I_1$</th>
<th>$I_2$</th>
<th>$I$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0144</td>
<td>-0.0144</td>
<td>0</td>
</tr>
<tr>
<td>0.1</td>
<td>0.0873</td>
<td>-0.0005</td>
<td>0.0821</td>
</tr>
<tr>
<td>0.2</td>
<td>0.19745</td>
<td>-3.26E-06</td>
<td>0.191048</td>
</tr>
<tr>
<td>0.3</td>
<td>0.299</td>
<td>0</td>
<td>0.2987</td>
</tr>
<tr>
<td>0.4</td>
<td>0.39986</td>
<td>0</td>
<td>0.39986</td>
</tr>
<tr>
<td>0.5</td>
<td>0.49989</td>
<td>0</td>
<td>0.49989</td>
</tr>
<tr>
<td>0.6</td>
<td>0.59987</td>
<td>0</td>
<td>0.59987</td>
</tr>
<tr>
<td>0.7</td>
<td>0.69985</td>
<td>0</td>
<td>0.69985</td>
</tr>
<tr>
<td>0.8</td>
<td>0.79983</td>
<td>0</td>
<td>0.79983</td>
</tr>
<tr>
<td>0.9</td>
<td>0.89981</td>
<td>0</td>
<td>0.89981</td>
</tr>
<tr>
<td>1</td>
<td>0.99979</td>
<td>0</td>
<td>0.99979</td>
</tr>
</tbody>
</table>
Table 2.6  Behaviour of Bias($t_5$) computed numerically with respect to $\mu_x$ for different values of $\alpha$ and for $n' = 200, \rho = 0.8$

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$\mu_x$</th>
<th>0</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td></td>
<td>0</td>
<td>0.059</td>
<td>0.042</td>
<td>0.01</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td></td>
<td>0</td>
<td>0.037</td>
<td>0.015</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0.25</td>
<td></td>
<td>0</td>
<td>0.011</td>
<td>0.002</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.7  Behaviour of Bias($t_5$) computed numerically with respect to $\mu_x$ for different values of $\rho$ and for $n' = 200, \alpha = 0.01$

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$\mu_x$</th>
<th>0</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.7</td>
<td></td>
<td>0</td>
<td>0.033</td>
<td>0.013</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td></td>
<td>0</td>
<td>0.037</td>
<td>0.015</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0.9</td>
<td></td>
<td>0</td>
<td>0.05</td>
<td>0.042</td>
<td>0.01</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
Figure 2.1 Behaviour of Bias($t_\delta$) computed analytically with respect to $\mu_X$ for $n' = 200$, $p = 0.8$, $\alpha = 0.01$.

Figure 2.2 Behaviour of Bias($t_\delta$) computed analytically with respect to $\mu_X$ for different values of $n'$ and for $p = 0.8$, $\alpha = 0.01$. 
Figure 2.3 Behaviour of Bias(\(t_5\)) computed analytically with respect to \(\mu_x\) for different values of \(\alpha\) and for \(n' = 200\), \(\rho = 0.8\)

Figure 2.4 Behaviour of Bias(\(t_5\)) computed analytically with respect to \(\mu_x\) for different values of \(\rho\) and for \(n' = 200\), \(\alpha = 0.01\)
Figure 2.5  Behaviour of Bias(tₖ) with respect to \( \mu_x \) for different refinements \( k \) of the interval of integration and for \( \alpha = 0.01, \rho = 0.8, n' = 200 \).

Figure 2.6  Behaviour of the function \( g(w) \) with respect to \( w \) for different values of \( \mu_x \) and for \( n' = 200 \).
Figure 2.7  Behaviour of Bias(t₅) computed numerically with respect to μₓ for different values of α and for n' = 200, ρ = 0.8

Figure 2.8  Behaviour of Bias(t₅) computed numerically with respect to μₓ for different values of ρ and for n' = 200, α = 0.01
Figure 2.9 Comparative behaviour of Bias(t₅) with respect to μₓ for different values of α and for ρ = 0.8, n' = 200

[Graph showing the comparative behaviour of Bias(t₅) with respect to μₓ for different values of α and for ρ = 0.8, n' = 200]