TITLE PAGE

USMANU DANFODIYO UNIVERSITY, SOKOTO

(POSTGRADUATE SCHOOL)

RATIO-TYPE ESTIMATORS IN STRATIFIED RANDOM SAMPLING USING AUXILIARY ATTRIBUTE

A Dissertation

Submitted to the

Postgraduate School

USMANU DANFODIYO UNIVERSITY, SOKOTO

In Partial Fulfillment of the Requirements

For the Award of the Degree of

MASTER OF SCIENCE (STATISTICS)

BY

AUDU AHMED

(10211311006)

DEPARTMENT OF MATHEMATICS

OCTOBER, 2012

CERTIFICATION

This dissertation by Audu, Ahmed (10211311006) has met the requirements for the award of the Degree of Master of Science (Statistics) of the Usmanu Danfodiyo University, Sokoto, and is approved for its contribution to knowledge.

________________ ____________
Prof. R. A. Ipinyomi Date
External Examiner

________________ ____________
Dr. R. V. K Singh Date
Major Supervisor

________________ ____________
Dr. A. Danbaba Date
Co – Supervisor I

________________ ____________
Dr. B. Z. Abubakar Date
Co –Supervisor II

________________ ____________
Dr. I. J. Uwanta Date
Head of Department

DEDICATION

I dedicated this research work to ALMIGHTY ALLAH (SubuhanahuWata'ala) and entire Jama'a Al-SunnahWaljamaha World-Wide.

ACKNOWLEDGEMENTS

In the name of Allah, Most Gracious, Most Merciful. I wish to express my sincere gratitude to Allah (SWT) for granting me opportunity to carry out this research work. My special thanks go to my major supervisor, Dr. R. V. K. Singh for supervision, assistance, advice and for the many hours he spent guiding me in my dissertation and in my career.

I sincerely, thank and appreciate the effort of my co-supervisors Dr. A. Danbaba and Dr. B. Z. Abubakar for their guidance and advices. I would also like to express my gratitude to all lecturers in the Department of Mathematics, especially Prof. S. U. Gulumbe for his guidance and advice, Head of Department in person of Dr. I. J. Uwanta, Dr. U. Ahmed, Mal. Umar, Mal. N. Dauran, Mal. B. Nakone, Mal. Yakubu Musa and the rest of them who assisted me in one way or another.

My profound gratitude goes to my Parents, friends and my entire family. In addition, my profound gratitude goes to my beloved wife Tawakaltu, Princess Raddiyah Ahmad and Prince Abdullahi Ahmad. I would not forget with my close friends like Shamsudden Suleiman, Abdumuhaimin A. Sanusi and the rest of them for their perseverance, encouragement and constant prayers for me while undergoing the program.

TABLE OF CONTENTS

LIST OF TABLES

ABBREVIATIONS/NOTATIONS

CV- Coefficient of Variation

MSE- Mean Square Error

P. R- Population Ratio

RSD- Relative Standard Deviation

R.B-Relative Bias

R. E.- Relative Efficiency

F- Female

M- Male

SPSS- Statistical Package for Social Science

ABSTRACT

A problem of the ratio-type estimators in Stratified Sampling is the use of non-attribute auxiliary information. In this study, some ratio-type estimators in stratified random sampling using attribute as auxiliary information are proposed. The sample mean of study variable and proportion of auxiliary attribute were transformed linearly and using auxiliary parameters respectively. Biases and mean square errors (MSE) for these estimators were derived. The MSE of these estimators were compared with the MSE of the traditional combined ratio estimator. The results show that the proposed estimators are more efficient and less bias than the combined ratio estimate in all conditions. An empirical study was also conducted using students height data from each faculty of the Usmanu Danfodiyo University, Sokoto. The results also show that the proposed estimators are more efficient and less bias than the combined ratio estimator. In addition, formulae for determination of sample sizes when the proposed estimators are adopted under various allocations (Optimum, Neyman and Proportional) for fixed cost and desired precision were obtained.

CHAPTER ONE

INTRODUCTION

1.1 INTRODUCTION

Prior knowledge about population mean along with coefficient of variation, kurtosis and correlation of the population of an auxiliary variable are known to be very useful particularly when the ratio, product and regression estimators are used for estimation of population mean of a variable of interest. The use of auxiliary information can increase the precision of an estimator when study variable is highly correlated with auxiliary variable. Srivastava and Jhajj (1981) suggested a class of estimators of the population mean, provided that the mean and variance of the auxiliary variable are known. Singh and Tailor (2003) considered a modified ratio estimator by exploiting the known value of correlation coefficient of the auxiliary variable. Singh and Upadhyaya (1999) suggested two ratio-type estimators when the coefficient of variation and kurtosis of the auxiliary variable are known.

However, the fact that the known population proportion of an attribute also provides similar type of information has not drawn as much attention. In several situations, instead of existence of auxiliary variables there exists some auxiliary attributes, which are highly correlated with study variable (Singh *et. al*.,2008). For example, sex and height of the persons, amount of milk produced by a particular breed of cow, amount of yield of wheat crop by a particular variety of wheat etc. (Jhajj *et. al*., 2006). In such situations, taking the advantage of point-biserial correlation between the study variable and the auxiliary attribute, the estimators of parameters of interest can be constructed by using prior knowledge of the parameters of auxiliary attribute.

It is often useful to incorporate auxiliary information of the population in a sampling procedure. In practice, auxiliary information can be obtained in different ways. For example, the sampling frames often used in official statistics production may include auxiliary information on the population elements or these data are extracted from administrative registers and are merged with the sampling frame elements. In other words, aggregate-level of auxiliary information can be obtained from different sources, such as published official statistics. Use of auxiliary information in sampling and estimation can be very useful in the construction of an efficient sampling design.

In the estimation of population parameters, auxiliary information is used to improve efficiency for the variable of interest. Whenever there is auxiliary information, the researcher wants to utilize it in the method of estimation to obtain the most efficient estimator.

In simple random sampling, the variance of the estimate (say, of population mean $\bar{Y}$) depends, apart from the sample size, on the variability of the character y in the population. If the population is very heterogeneous and considerations of cost limit the size of the sample, it may be found impossible to get a sufficiently precise estimate by taking a simple random sample from entire population. And populations encountered in practice are generally very heterogeneous (Raj and Chandhok, 1998). In surveys of manufacturing establishments, for example, it can be found that some establishment are very large, that is, they employ 1000 or more persons, but there are many others which have only two or three persons on their rolls. Any estimate made from a direct random sample taken from the totality of such establishments would be subject to exceedingly large sampling fluctuations. But suppose it is possible to divide this population into parts or strata on the basis of, say employment, thereby separating the very large ones, the medium-sized ones and the smaller ones. If a random of establishments is now taken from each stratum, it should be possible to make a better estimate of the strata average, which in turn should help in producing a better of the population average. Similarly, if a sample is selected with probability proportionate to x from the entire population, the variance of the population-total estimate may be very high because the ratio of y to x varies considerably over the population. If a way can be found of subdividing the population so that the variation of the ratio of y to x is considerably reduced within the subdivisions or strata, a better estimate of the population can be made. This is the basic consideration involved in the use of stratification for improving the precision of estimation (Raj and Chandhok, 1998).

1.2 CENSUS VERSUS SAMPLE SURVEY

Broadly speaking, information on population may be collected in two ways. Either every unit in the population is enumerated (called complete enumeration, or census) or enumeration is limited to only a part or sample selected from the population (called sample enumeration or sample survey). A sample survey will usually be less costly than a complete census because the expense of covering all units would be greater than that of covering only a sample fraction. Also, it will take less time to collect and process data from a sample than from a census. But economy is not the only consideration; the most important point is whether the accuracy of the results would be adequate for the end in view. It is a curious fact that the results from a carefully planned and well executed sample survey are expected to be more accurate (near to the aim of study) than those from a complete enumeration that can be taken. A complete census ordinarily requires a huge and unwieldy organization and therefore many types of errors creep in which cannot be controlled adequately. In a sample survey the volume of work is reduced considerably, and it becomes possible to employ persons of higher caliber, train them suitably, and supervise their work adequately. In a properly designed sample survey it is also possible to make a valid estimate of the margin of error and hence decide whether the results are sufficiently accurate. A complete census does not reveal by its self the margin of uncertainty to which it is subject. But there is not always a choice of one versus the other. For example, if the data are required for every small administrative area in a country, no sample survey of a reasonable size will be able to deliver the desired information; only a complete census can do this (Raj and Chandhok, 1998).

1.3 RANDOM SAMPLING

Simple random sampling is a method of selecting n units out of the N such that every one of the ${}_N C_n$ distinct samples has an equal chance of being drawn. In practice a simple random sample is drawn unit by unit. The units in the population are numbered from 1 to N . A series of random numbers between 1 and N is then drawn, either by means of a table of random numbers or by means of a computer program that produces such a table. At any draw the process used must give an equal chance of selection to any number in the population not already drawn. The units that bear these n numbers constitute the sample. It is easily verified that all ${}_N C_n$ distinct samples have an equal chance of being selected by this method. Consider one distinct sample, that is, one set of n specified units. At the first draw the probability that some one of the n specified units is selected is n/N . At the second draw the probability that some one of the remaining $(n-1)$ specified units is drawn is, and so on. Hence the probability that all n specified units are selected in n draws is

$$\frac{n}{N}\cdot\frac{(n-1)}{(N-1)}\cdot\frac{(n-2)}{(N-2)}\cdots\frac{1}{(N-n+1)}=\frac{n!(N-n)!}{N!}=\frac{1}{{}_N C_n}$$

Since a number that has been drawn is removed from the population for all subsequent draws, this method is also called random sampling without replacement. Random sampling with replacement is entirely feasible; at any draw, all N members of the population are given equal chance of being drawn, no matter how often they have been drawn. The formulas for the variances and estimated variances of estimates made from the sample are often simpler when sampling is with replacement than when it is without replacement. For this reason sampling with replacement is sometimes used in the more complex sampling plans (Cochran, 1977).

1.4 DEFINITION OF BASIC TERMS

Sample:- A sample is a group of units selected from larger group (population). By studying the sample, it is hoped to draw valid conclusions about the larger group. A sample is generally selected for study because the population is too large to study in its entirety. The sample should be representative of general population. This is often best achieved by random sampling. Also, before collecting the sample, it is important that the researcher carefully and completely defines the population, including a description of the members to be included (Cochran, 1977).

Parameter:- A parameter is a value usually unknown (and which therefore has to be estimated), used to represent a certain population characteristic. Within a population, a parameter is fixed value which does not vary. They are often denoted by Greek letters (Cochran, 1977).

Statistic:- A statistic is a quantity that is calculated from a sample data. It is used to give information about unknown values in the corresponding population. it is possible to draw more than one sample from the same population and the value of a statistic will in general vary from sample to sample. Therefore, statistic is a random variable (Cochran, 1977).

Estimator:- An estimator is a rule for calculating an estimate of a given quantity based on observed data. There are point and interval estimators. The point estimator yields single-valued results, although this includes the possibility of single vector-valued results and results that can be expressed as a single function. This is in contrast to an interval estimator, where the results would be a range of plausible values (or vectors or functions). An estimator is a statistic, (that is, a function of data) that is used to infer the value of an unknown parameter in statistical model. The parameter being estimated is sometimes called estimand. It can be either finite-dimensional (in parametric and semi-parametric) or finite-dimensional (in nonparametric and semi-nonparametric models). If the parameter is denoted by θ, then the estimator is typically written as $\hat{\theta}$. Being a function of data, the estimator is a random variable (Cochran, 1977).

Bias:-the bias of an estimator is the difference between this estimator's expected value and the true value of the parameter being estimated. An estimator with zero bias is called unbiased. Otherwise the estimator is said to be biased. Suppose we have a statistical model parameterized by θ giving rise to a probability distribution for observed data $p(x \setminus \theta)$ and a statistic $\hat{\theta}$ which serves as an estimator based on the any observed data

x. That is, we assume that our data follows some unknown distribution $p(x \setminus \theta)$ (where θ is a fixed constant that is part of this distribution, but is unknown), and then we construct some estimators $\hat{\theta}$ that maps observed data to values that we hope are close to. Then the bias of this estimator is defined to be;

$Bias(\hat{\theta}) = E(\hat{\theta}) - \theta$ (Cochran, 1977).

Mean Square Error:- The MSE of an estimator is one of many ways to quantify the difference between values implies by an estimator and the true values of the quantity being estimated. MSE is a risk function, corresponding to the expected value of the squared error loss or quadratic loss. MSE measures the average of the squares of the errors. The error is the amount by which the value implied by the estimator differs from the quantity to be estimated. The difference occurs because of randomness or because the estimator doesn't account for information that could produce a more accurate estimate. The MSE is the second moment (about the origin) of the error, and thus incorporates both the variance of the estimator and its bias. For an unbiased estimator, the MSE is the variance. The MSE of an estimator $\hat{\theta}$ with respect to the estimated θ is defined mathematically as;

$$MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2$$

$$= \text{var}(\hat{\theta}) + (Bias(\hat{\theta}))^2$$

The MSE thus assess the quality of an estimator in terms of its variation and unbiasedness (Cochran, 1977).

Kurtosis:- Kurtosis is any measure of the peakedness of the probability distribution of a real-valued random variable. It is descriptor of the shape of probability distributions. One common measure of kurtosis originated by Pearson, is based on a scaled version of the fourth moment of the data or population. For this measure, higher kurtosis means more of the variance is the result of infrequent extreme deviations as opposed to frequent modestly sized deviations. Distributions with negative or positive excess are called platykurtic or leptokurtic respectively. The fourth standardized moment is defined as;

$\beta_2 = \frac{\mu_4}{\sigma^4}$, Where μ_4 is the fourth moment about the mean and σ is the standard deviation (Cochran, 1977).

Point-biserial correlation coefficient:- Point-biserial correlation coefficient denoted by ρ_{pb} is a correlation used when one variable (e.g Y) is dichotomous ; Y can either be naturally dichotomous like gender or an artificial dichotomous variable. Point-biserial correlation is mathematically equivalent to the Pearson product moment correlation; that is, if we have one continuously measured variable X and a dichotomous variable Y. This can be shown by assigning two distinct numerical values (say, 1 and 2) to dichotomous variable. The Point-biserial correlation coefficient is given as;

$$\rho_{pb} = \frac{M_1 - M_2}{s_n}\sqrt{\frac{n_1 n_2}{n^2}}$$

Where $s_n = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}$ the standard deviation for X , M_1 is the mean value on the continuous variable for all data points in group 1, M_2 is the mean value on the continuous variable for all data points in group 2, n_1 is the number of data point in group1, n_2 is the number of data point in group 2 and n is the sample size (John, 2008).

Coefficient of Variation (CV):- CV is a normalized measure of dispersion of a probability distribution. It is known as unitized risk or the variation coefficient. The absolute value of CV is sometimes known as relative standard deviation (RSD), which is express as a percentage. The CV is defined as the ratio of the standard deviation to the mean;

$CV = \frac{\sigma}{\mu}$ which is the inverse of the signal-to-noise ratio. It shows the extent of variability in relation to mean of the population (Cochran, 1977).

1.5 AIM AND OBJECTIVES

The aim of this research work is to develop some ratio-type estimators under stratified random sampling scheme using auxiliary attributes that will produce more precise estimates than the conventional estimator.

The above aim is achieved through the following objectives;

1. To linearly transform the sample mean of the variable of interest.
2. To transform the proportion of auxiliary attributes using auxiliary parameters like kurtosis, coefficient of variation and coefficient of Point-biserial correlation.
3. To obtain the biases and mean square errors of the proposed estimators up to first order approximation using Taylors' expansion.

4. To obtain the conditions for efficiency of the proposed estimators over the conventional estimator.

1.6 SIGNIFICANCE OF THE STUDY

Ratio estimators of population parameters are more precise than their simple random sampling estimators' counterparts (Cochran 1942). The mean square error of ratio estimator can be reduced with the application of transformation on the study and auxiliary variables (Chaudhuri and Adrikari 1979). Situations arise when the available auxiliary information are inform of attributes instead of variables. Based on these situations, some ratio-type estimators had been proposed by several researchers in simple random sampling which regards the population units as homogeneous. There are possibilities in which population units are heterogeneous as a whole but homogeneous within sub-populations (strata). In such situations, there is need to develop estimators that capture the variability within and between the strata for population parameters of interest with emphasis to bias reduction and efficiency improvement.

1.7 SCOPE AND LIMITATION

This research work primarily considers some ratio-type estimators in stratified sampling using attribute as auxiliary information. The transformation of the study variable mean is linear and kurtosis, coefficient of variation and coefficient of point-biserial correlation are the parameters of auxiliary attribute used for the transformation of proportion of auxiliary attribute. The data used for the empirical study was taken from Students Pre-medical Registration, Usmanu Danfodiyo University, Sokoto (2011/2012 Session). The results of the analysis are limited to the data used, the set of the proposed estimators and the sample sizes taken from the data used. In future research, efforts will be made toward modification of the proposed estimators to obtain unbiased or almost unbiased estimators with higher precisions.

CHAPTER TWO

LITERATURE REVIEW

2.1 RATIO ESTIMATORS

The work incorporated in this dissertation is associated mainly with the study of efficiency of ratio-type estimators. Although, a significant contribution by the survey statisticians has been made towards the development of ratio and product estimators using the auxiliary information at the estimation stage; a review of only those items of dissertation work has been made in this section which have an immediate bearing on and relevance to the present work. However, the review presents only a proper orientation and perspective to the present work.

It was late 1930s and early 1940s when a few major centers emerged for research in and application of sampling methodology. One was the statistical laboratory at Ames, Iowa, another was the Bureau of Census in United States. At the same time major advances were occurred at a few other locations.

In the decade of 1940-1950, the contributions of Cochran (1942), Deming (1956), Hurvitz (1952), Hansen (1951), Jessen (1942), Madow (1953), Sukhatme (1947), Yates (1948) and others helped in lying the foundation of modern sampling theory. The work done during this period related mainly with the development of sampling and estimation procedures under different sets of assumptions about the population values.

In 1942, Cochran made a particular important contribution to the modern sampling theory by suggesting methods of utilizing the auxiliary information for the purpose of estimation of the population mean in order to increase the precision of the estimates (Cochran 1942).

Several well-known procedures use auxiliary information at the estimation stage. This is most commonly used way of utilizing auxiliary information which gives rise to some estimators that are known today, and in under certain conditions, these estimators are more efficient than the estimators based on simple random sampling. He used the ratio estimator of the form

$$\bar{y}_r = \frac{\bar{y}}{\bar{x}}\bar{X} \; ; \; \bar{x} \neq 0 \qquad (2.1)$$

Where $\bar{y}$ and $\bar{x}$ are the sample means of the characteristics under study and auxiliary characteristics respectively based on a sample drawn under simple random sampling design and $\bar{X}$ is the population mean of the auxiliary characteristic X .

The aim of this method was to use the ratio of sample means of two characteristics which would be almost stable in sampling fluctuation and thus, would provide a better estimate of the true value. It has been a well-known fact that $\bar{y}_r$ is more efficient than the sample mean estimator, $\bar{y}$, where no auxiliary information is used.

Contrary to the situation of ratio estimator, if variants Y and X are negatively correlated, then the product estimator $\bar{y}_p$ given by

$$\bar{y}_p = \frac{\bar{y}}{\bar{X}}\bar{x} \; ; \; \bar{X} \neq 0 \qquad (2.2)$$

which was proposed by Murthy (1964), has been observed to give higher precision than $\bar{y}$, the sample mean estimator.

The expression for the bias and mean square error (MSE) of $\bar{y}_r$ and $\bar{y}_p$ have been derived by Cochran (1942) and Murthy (1964) respectively which are also available in the books by Cochran (1977), Jessen (1978) and Murthy (1967).

Since the ratio estimator was observed to be more precise than the usual sample mean estimator under different conditions, several researchers diverted their attention in the direction of modifying estimation procedure so that unbiased or less bias estimators could be obtained. The work of Hartley and Ross (1954) deserves special attention in this direction. Several other authors also proposed unbiased or almost unbiased ratio-type estimators. Beale (1962) and Tin (1965) proposed bias-adjusted ratio estimators which are equally efficient in finite population. Rao (1981a, 1981b), Rao and Webster (1966) and Rao and Beagle (1967) also suggested some almost unbiased ratio-type estimators.

Several authors have used prior value of certain population parameter(s) to find more precise estimates. Searls (1964) used coefficient of variation of study character at estimation stage. In practice this coefficient of variation is seldom known. Motivated by Searls (1964) work, Sen (1978), Sisodiya and Dwivedi (1981), and Upadhyaya and Singh (1984) used the known coefficient of variation of the auxiliary character for estimating population mean of a study character in ratio method of estimation. The use of prior value of coefficient of kurtosis in estimating the population variance of study character was first made by Singh *et. al*. (1973). Later used by Searls and Intarapanich (1990), Singh and Upadhyaya (1999), Singh (2003) and Singh et. al. (2004) in the estimation population mean of study character. Recently Singh and Tailor (2003) proposed a modified ratio estimator by using estimator by using the known value of correlation coefficient.

2.2 RANKED SET SAMPLING

A ranked set sampling is a ratio estimator in which the observations for variable of interest and auxiliary variables are ranked instead of being quantified.

Mclntyre (1952) introduced ranked set sampling method for estimating the mean of pasture yield. In situations the sampling units in a study can be more easily ranked than quantified, Mclntyre suggested the mean of m sample units based on ranked set sampling as an estimator of the population mean. This estimator is unbiased estimator with a smaller variance compared to the usual sample mean based on simple random sample of the same size. Takahasi and Wakimoto (1968) provided the mathematical properties of ranked set sampling. Samawi and Muttlak (1996) used ranked set sampling method to estimate the population ratio and showed that it provided a more efficient estimator compared to the one obtained by simple random sample.

Al Odat (2009) suggested a modification of estimating a ratio in rank set sampling. The variance of the suggested estimator was obtained and compared with the variance of the traditional ratio rank set sampling. By this comparison, the condition which Make the proposed estimator more efficient than the traditional estimator was found.

2.3 STRATIFIED RATIO ESTIMATOR

Kadilar and Cingi (2003) examined some ratio-type estimators in stratified random sampling and obtained their mean square error equations. By these equations, mean square error of the estimators have been compared in theory and by this comparison, the conditions which the estimators have smaller mean square error with respect to each other was found. These theoretical conditions are also satisfied by the results of an application with original data. In this application, it was concluded that the traditional combined ratio estimator is more efficient than ratio estimators developed in recent years. This conclusion shows that, in the forthcoming studies, new ratio-type estimators should be proposed not only in simple random sampling but also in stratified random sampling.

Kadilar and Cingi (2005) suggested a new ratio estimator in stratified sampling based on the Prasad (1989) estimator. The estimator was compared with traditional estimator theoretically and from numerical demonstration, the proposed estimator was found to be more efficient than the combined ratio estimate in all conditions.

Naik and Gupta (1996) defined ratio estimator of population mean when the prior information of population proportion of units, possessing the same attribute is available taking into consideration the point-biserial correlation between a variable and an attribute as;

$$t_{NG} = \bar{y}\frac{P}{p},\ p \neq 0 \tag{2.3}$$

where $\bar{y}$ is the sample mean of study variable, p is the proportion of the sample unit possessing the auxiliary attribute and P is the proportion of the population units possessing the auxiliary attribute. The Naik and Gupta (1996) estimator in stratified sampling is defined as;

Consider a random sample of size $n = \sum_{h=1}^{k} n_h$ to be taken from population of size $N = \sum_{h=1}^{k} N_h$ stratified into k strata. Let a sample of size $n_{h,}\ \left(h = 1, 2, ..., k\right)$ be drawn by simple random sampling without replacement from a stratum h of size N_h. Let y_i and ϕ_i denote the observations on a random variable y and ϕ respectively for i^{th} unit $(i = 1, 2, ..., N)$. Suppose there is complete dichotomy in the population with respect to the presence or absence of an attribute, say ϕ, and it is assumed that attribute ϕ takes only two values 0 and 1 as

$$\phi_i = \begin{cases} 1, if\ i^{th}\ unit\ of\ the\ population\ possesses\ attribute\ \phi \\ 0, otherwise \end{cases}$$

Then we have the following definitions;

$A = \sum_{i=1}^{N} \phi_i -$ denotes the total number of units in the population possessing attribute ϕ.

$A_h = \sum_{i=1}^{N_h} \phi_{hi} -$ denotes the total number of units in the stratum h possessing attribute ϕ.

$a_h = \sum_{i=1}^{n_h} \phi_{hi} -$ denotes the total number of units in the sample drawn from stratum h possessing attribute ϕ.

$P = \frac{A}{N} -$ denotes the proportion of units in the population possessing attribute ϕ.

$P_h = \frac{A_h}{N_h} -$ denotes the proportion of units in the stratum h possessing attribute ϕ.

$p_h = \frac{a_h}{n_h}$ – denotes the proportion of units in the sample drawn from stratum h possessing attribute ϕ.

In stratified random sampling, we proposed the following combined ratio estimators for the population mean using auxiliary attribute

$$\hat{\bar{T}} = \frac{\bar{y}_{st}}{p_{st}} P = R_{st} P \tag{2.4}$$

Where $\bar{y}_{st} = \sum_{h=1}^{k} W_h \bar{y}_h$, $\bar{y}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} y_i$, $p_{st} = \sum_{h=1}^{k} W_h p_h$, $W_h = \frac{N_h}{N}$ and $R_{st} = \frac{\bar{y}_{st}}{p_{st}}$

Singh *et. al*. (2008) proposed some ratio estimators for estimating the population mean of the variable under study, which make use of information regarding population proportion possessing certain attribute. Under simple random sampling without replacement scheme, the expressions of bias and mean square error up to first order approximation are derived. The results obtained were illustrated numerically by taking some empirical population considered in the literature.

Dash *et. al*. (2011) suggested a class of estimators in stratified sampling using transformation on both study and auxiliary variables. The proposed class of estimators can be reduced to an infinite number of estimators namely HT-estimator, Usual ratio and product estimators, Kadilar and Cingi (2005) estimator, Shabbir and Gupta (2006) estimator using Bedi (1996) transformation. Its optimum properties were studied and also compared with several competitive estimators and found to be more efficient.

In literature, it has been shown by various authors such as Mohanty and Das (1971), Reddy (1974), Reddy and Rao (1977), Srivenkatarama(1978), Chaudhuri and Adrikari (1979) and Mahalanobis (1994) suggested that the bias and mean square error of ratio estimator $\bar{y}_R$ can be reduced with the application of transformation on the auxiliary variable X. Singh *et. al*. (1973) used coefficient of kurtosis to improve the efficiency of the ratio estimator in the estimation procedure of variance.

The facts that the ratio estimators have superiority over sample mean estimators when the correlation between the study and auxiliary variables in the population is positively high, led the statisticians to focus their attention on the modifications of such conventional estimators so that the modified estimators can work efficiently even if the correlation is low. Consequently, a number of modified ratio estimators came into existence in recent past. Such estimators are generally developed either using one or more unknown constants or introducing a convex linear combination of population and sample means of the auxiliary characteristics with unknown weights. In both the cases, optimum choices of unknown parameters are made by minimizing the mean square error of the modified estimators so that they become superior to the conventional one. Srivenkatarama (1980), Srivenkatarama and Tracy (1979, 1980 and 1981), Srivastava and Jhajj (1981) and others have suggested modified ratio type estimators and studied their properties theoretically and empirically.

CHAPTER THREE

MATERIALS AND METHODOLOGY

3.1 INTRODUCTION

Sampling technique adopted in this research work is stratified random sampling for estimating population mean for random variable y using parameters of auxiliary information (attribute). This enables the researchers to obtain ratio estimators, their mean square errors and biases, as well as identifying the estimators with least mean square error and bias. The estimator with least mean square error and bias is preferred and identified as the most efficient estimators among the ratio estimators defined in the work.

3.2 DATA USED FOR THE ANALYSIS

The data used for the analysis was taken from Students Pre-Medical Registration, Usmanu Danfodiyo University, Sokoto (2011/2012 Session). The height and sex of students were taken according to their respective faculties. The data is shown in the Appendix I.

3.3 SOFTWARE USED FOR THE ANALYSIS

SPSS 17 was used to obtain the strata parameters and JAVA program was written to return the values of biases and MSE of both proposed estimators and conventional estimator. The source code is given in Appendix II.

3.4 PROPOSED ESTIMATORS

The proposed estimators can be generally written in the form;

$$\hat{\bar{T}}_i^* = \frac{\sum_{h=1}^{k} W_h(\bar{y}_h - b_{\phi h}(p_h - P_h))}{\sum_{h=1}^{k} W_h(m_{1i}p_h + m_{2i})}(m_{1i}P + m_{2i}),\ (i = 1, 2, 3, ...) \qquad (3.1)$$

Where $m_{1i}(\neq 0)$ and m_{2i} are either real numbers or the functions of the known parameters of the attribute such as coefficient of variation C_p, coefficient of kurtosis $B_2(\phi)$ and point biserial correlation coefficient ρ_{pb}, $b_{\phi h} = \frac{s_{y\phi h}}{s_{\phi h}^2}$ is the slope of regression equation between the study variable and auxiliary attribute, $s_{\phi h}^2 = \frac{1}{n_h - 1}\sum_{i=1}^{n_h}(\phi_i - p_h)^2$ and $s_{y\phi h} = \frac{1}{n_h - 1}\sum_{i=1}^{n_h}(\phi_i - p_h)(y_i - \bar{y}_h)$.

Remarks 1: When $b_{\phi h} = 0$, $m_{1i} = 1$ and $m_{2i} = 0$, the proposed estimators reduce to traditional estimator.

The following scheme presents the possible estimators of the population mean, which can be obtained by suitable choice of constants m_{1i} and m_{2i}.

Table 3.1: Proposed Estimators based on the available Population Parameters.

Estimator	Values of m_{1i}	Values of m_{2i}
$\hat{\bar{T}}_1^* = \dfrac{\sum_{h=1}^{k} W_h\left(\bar{y}_h - b_{\phi h}\left(p_h - P_h\right)\right)}{\sum_{h=1}^{k} W_h p_h} P$	1	0
$\hat{\bar{T}}_2^* = \dfrac{\sum_{h=1}^{k} W_h\left(\bar{y}_h - b_{\phi h}\left(p_h - P_h\right)\right)}{\sum_{h=1}^{k} W_h\left(p_h + \beta_2(\phi)\right)}\left(P + \beta_2(\phi)\right)$	1	$\beta_2(\phi)$
$\hat{\bar{T}}_3^* = \dfrac{\sum_{h=1}^{k} W_h\left(\bar{y}_h - b_{\phi h}\left(p_h - P_h\right)\right)}{\sum_{h=1}^{k} W_h\left(p_h + C_p\right)}\left(P + C_p\right)$	1	C_p
$\hat{\bar{T}}_4^* = \dfrac{\sum_{h=1}^{k} W_h\left(\bar{y}_h - b_{\phi h}\left(p_h - P_h\right)\right)}{\sum_{h=1}^{k} W_h\left(p_h + \rho_{pb}\right)}\left(P + \rho_{pb}\right)$	1	ρ_{pb}
$\hat{\bar{T}}_5^* = \dfrac{\sum_{h=1}^{k} W_h\left(\bar{y}_h - b_{\phi h}\left(p_h - P_h\right)\right)}{\sum_{h=1}^{k} W_h\left(\beta(\phi) p_h + C_p\right)}\left(\beta(\phi) P + C_p\right)$	$\beta_2(\phi)$	C_p
$\hat{\bar{T}}_6^* = \dfrac{\sum_{h=1}^{k} W_h\left(\bar{y}_h - b_{\phi h}\left(p_h - P_h\right)\right)}{\sum_{h=1}^{k} W_h\left(C_p p_h + \beta_2(\phi)\right)}\left(C_p P + \beta_2(\phi)\right)$	C_p	$\beta_2(\phi)$

Table 3.1: Continued

$\hat{\hat{T}}_7^* = \dfrac{\sum_{h=1}^{k} W_h\left(\bar{y}_h - b_{\phi h}\left(p_h - P_h\right)\right)}{\sum_{h=1}^{k} W_h\left(C_p p_h + \rho_{pb}\right)}\left(C_p P + \rho_{pb}\right)$	C_p	ρ_{pb}
$\hat{\hat{T}}_8^* = \dfrac{\sum_{h=1}^{k} W_h\left(\bar{y}_h - b_{\phi h}\left(p_h - P_h\right)\right)}{\sum_{h=1}^{k} W_h\left(\rho_{pb} p_h + C_p\right)}\left(\rho_{pb} P + C_p\right)$	ρ_{pb}	C_p
$\hat{\hat{T}}_9^* = \dfrac{\sum_{h=1}^{k} W_h\left(\bar{y}_h - b_{\phi h}\left(p_h - P_h\right)\right)}{\sum_{h=1}^{k} W_h\left(\beta_2(\phi) p_h + \rho_{pb}\right)}\left(\beta_2(\phi) P + \rho_{pb}\right)$	$\beta_2(\phi)$	ρ_{pb}
$\hat{\hat{T}}_{10}^* = \dfrac{\sum_{h=1}^{k} W_h\left(\bar{y}_h - b_{\phi h}\left(p_h - P_h\right)\right)}{\sum_{h=1}^{k} W_h\left(\rho_{pb} p_h + \beta_2(\phi)\right)}\left(\rho_{pb} P + \beta_2(\phi)\right)$	ρ_{pb}	$\beta_2(\phi)$

Equation (3.1) can be written in the form

$$\hat{\hat{T}}_i^* = \frac{\bar{y}_{st}^{\ *}}{p_{st}^{\ *}} P^* = R_{st}^{\ *} P^*, \ (i = 1, 2, ..., 10) \tag{3.2}$$

Where $\bar{y}_{st}^{\ *} = \sum_{h=1}^{k} W_h(\bar{y}_h - b_{\phi h}(p_h - P_h))$, $p_{st}^{\ *} = \sum_{h=1}^{k} W_h(m_{1i} p_h + m_{2i})$, $P^* = m_{1i} P + m_{2i}$,

$R_{st}^* = \dfrac{\bar{y}_{st}^*}{p_{st}^*}$

3.5 BIAS AND MEAN SQUARE ERROR (MSE) OF ESTIMATOR $\hat{\hat{T}}$

The traditional estimator is defined as

$$\hat{\hat{T}} = \frac{\bar{y}_{st}}{p_{st}} P = R_{st} P$$

The population ratio $R = \frac{\bar{Y}}{P}$ which implies that $\bar{Y} = RP$.

The expectation of the estimator is given as

$$E\left(\hat{\hat{T}}\right) = PE\left(\frac{\bar{y}_{st}}{p_{st}}\right) = PE\left(\frac{\bar{Y} + \bar{\varepsilon}'}{P + \bar{\varepsilon}}\right) \tag{3.3}$$

Where $\bar{\varepsilon}' = \bar{y}_{st} - \bar{Y}$, $\bar{\varepsilon} = p_{st} - P$

$$E\left(\hat{\hat{T}}\right) = \bar{Y}E\left(\frac{1 + \frac{\bar{\varepsilon}'}{\bar{Y}}}{1 + \frac{\bar{\varepsilon}}{P}}\right) = \bar{Y}E\left(\left(1 + \frac{\bar{\varepsilon}'}{\bar{Y}}\right)\left(1 + \frac{\bar{\varepsilon}}{P}\right)^{-1}\right)$$

$$= \bar{Y}E\left(1 - \frac{\bar{\varepsilon}}{P} + \frac{\bar{\varepsilon}'}{\bar{Y}} - \frac{\bar{\varepsilon}'\bar{\varepsilon}}{P\bar{Y}} + \frac{\bar{\varepsilon}^2}{P^2} - \ldots\right)$$

Then, the expectation of traditional estimator to first order approximation is given as

$$E\left(\hat{\hat{T}}\right) \cong \bar{Y}\left(1 - \frac{1}{P\bar{Y}} \operatorname{cov}\left(\bar{y}_{st} p_{st}\right) + \frac{1}{P^2} \operatorname{var}\left(p_{st}\right)\right) \tag{3.4}$$

$$\cong \bar{Y}\left(1 - \frac{1}{P\bar{Y}} \sum_{h=1}^{k} W_h^2 \gamma_h S_{y\phi h} + \frac{1}{P^2} \sum_{h=1}^{k} W_h^2 \gamma_h S_{\phi h}^2\right)$$

The bias of the estimator is given by

$$Bias(\hat{\bar{T}}) = E(\hat{\bar{T}}) - \bar{Y}$$

$$= E(R_{st}P) - RP$$

$$= P(E(R_{st}) - R)$$

$$Bias\left(\hat{\bar{T}}\right) = PBias(R_{st}) \tag{3.5}$$

Now,

$$Bias(R_{st}) = E(R_{st}) - R = E(R_{st} - R)$$

$$= E\left(\frac{\bar{y}_{st}}{p_{st}} - R\right) = E\left(\frac{\bar{y}_{st} - Rp_{st}}{p_{st}}\right)$$

$$= E\left(\frac{\bar{y}_{st} - Rp_{st}}{P + \bar{\varepsilon}}\right), \text{ where } \bar{\varepsilon} = p_{st} - P$$

Hence the $Bias(R_{st})$ is the value at $\theta = 1$ of the function

$$f(\theta) = E\left(\frac{\bar{y}_{st} - Rp_{st}}{P + \theta\bar{\varepsilon}}\right)$$

We shall find Taylor's expansion of $f(\theta)$ around $\theta = 0$. This expression is

$$f(\theta) = f(0) + \theta f'(0) + \frac{\theta^2}{2!} f''(0) + \ldots$$

Hence, the bias of R_{st} is given by

$$Bias(R_n) = f(0) + f'(0) + \frac{1}{2!} f''(0) + ...$$

The results of the derivatives are given as

$$f(0) = E\left(\frac{\bar{y}_{st} - Rp_{st}}{P}\right),\ f'(\theta) = -E\left(\frac{(\bar{y}_{st} - Rp_{st})\bar{\varepsilon}}{(P + \theta\bar{\varepsilon})^2}\right),\ f'(0) = -E\left(\frac{(\bar{y}_{st} - Rp_{st})\bar{\varepsilon}}{P^2}\right),$$

$$f''(\theta) = E\left(\frac{(\bar{y}_{st} - Rp_{st})\bar{\varepsilon}^2}{(P + \theta\bar{\varepsilon})^3}\right),\ f''(0) = E\left(\frac{(\bar{y}_{st} - Rp_{st})\bar{\varepsilon}^2}{P^3}\right)$$

So,

$$Bias(R_{st}) = E\left(\frac{\bar{y}_{st} - Rp_{st}}{P}\right) - E\left(\frac{(\bar{y}_{st} - Rp_{st})\bar{\varepsilon}}{P^2}\right) + E\left(\frac{(\bar{y}_{st} - Rp_{st})\bar{\varepsilon}^2}{P^3}\right) + ...$$

Then, the $Bias(R_{st})$ to first order approximation is given by

$$Bias(R_{st}) \cong E\left(\frac{\bar{y}_{st} - Rp_{st}}{P}\right) - E\left(\frac{(\bar{y}_{st} - Rp_{st})\bar{\varepsilon}}{P^2}\right)$$

$$\cong -\frac{E\left((\bar{y}_{st} - Rp_{st})\bar{\varepsilon}\right) - PE(\bar{y}_{st} - Rp_{st})}{P^2}$$

$$\cong -\frac{E\left((\bar{y}_{st} - Rp_{st})(p_{st} - P)\right) - E(p_{st})E(\bar{y}_{st} - Rp_{st})}{P^2}$$

$$\cong -\frac{1}{P^2}\operatorname{cov}\left((\bar{y}_{st} - Rp_{st})\, p_{st}\right)$$

$$Bias\left(R_{st}\right) \cong -\frac{1}{P^2}\left(\operatorname{cov}\left(\bar{y}_{st}\, p_{st}\right) - R\operatorname{var}\left(p_{st}\right)\right) \tag{3.6}$$

$$\cong -\frac{1}{P^2}\left(\sum_{h=1}^{k} W_h^2 \gamma_h S_{y\phi h} - R\sum_{h=1}^{k} W_h^2 \gamma_h S_{\phi h}^2\right)$$

$$\cong \frac{1}{P^2}\sum_{h=1}^{k} W_h^2 \gamma_h \left(RS_{\phi h}^2 - S_{y\phi h}\right) \tag{3.7}$$

Substituting eqn.(3.7) into eqn.(3.5) produces the bias of the estimator as

$$Bias\left(\hat{\bar{T}}\right) \cong \frac{1}{P}\sum_{h=1}^{k} W_h^2 \gamma_h \left(RS_{\phi h}^2 - S_{y\phi h}\right) \tag{3.8}$$

Now to obtain the mean square error (MSE) of the estimator;

$$MSE\left(\hat{\bar{T}}\right) = E\left(\hat{\bar{T}} - \bar{Y}\right)^2$$

$$MSE\left(\hat{\bar{T}}\right) = P^2 E\left(R_n - R\right)^2 \tag{3.9}$$

$$E\left(R_{st} - R\right)^2 = E\left(\frac{\bar{y}_{st}}{p_{st}} - R\right)^2 = E\left(\frac{\bar{y}_{st} - Rp_{st}}{p_{st}}\right)^2$$

$$MSE\left(R_{st}\right) = E\left(\frac{\bar{y}_{st} - Rp_{st}}{P + \bar{\varepsilon}}\right)^2, \text{ where } \bar{\varepsilon} = p_{st} - P$$

Hence, the $MSE\left(R_{st}\right)$ is the value at $\theta = 1$ of the function

$$g(\theta) = E\left(\frac{\bar{y}_{st} - Rp_{st}}{P + \theta\bar{\varepsilon}}\right)^2, \text{ where } \bar{\varepsilon} = p_{st} - P$$

We shall find Taylor's expansion of $g(\theta)$ around $\theta = 0$. The expression is given as

$$g(\theta) = g(0) + \theta g'(0) + \frac{\theta^2}{2!} g'(0) + \ldots$$

Hence, the mean square error of R_{st} is given by

$$MSE(R_{st}) = g(0) + g'(0) + \frac{1}{2!} g''(0) + \ldots$$

The results of the derivatives are given below

$$g(0) = E\frac{(\bar{y}_{st} - Rp_{st})^2}{P^2},\ g'(\theta) = -2E\left(\frac{(\bar{y}_{st} - Rp_{st})^2 \bar{\varepsilon}}{(P + \theta\bar{\varepsilon})^3}\right),\ g'(0) = -2\frac{E\left((\bar{y}_{st} - Rp_{st})^2 \bar{\varepsilon}\right)}{P^3}$$

So,

$$MSE(R_{st}) = \frac{E(\bar{y}_{st} - Rp_{st})^2}{P^2} - 2\frac{E\left((\bar{y}_{st} - Rp_{st})^2 \bar{\varepsilon}\right)}{P^3} + \ldots$$

Then the $MSE(R_{st})$ to first order approximation is given by

$$MSE(R_{st}) \cong \frac{1}{P^2} E(\bar{y}_{st} - Rp_{st})^2$$

$$\cong \frac{1}{P^2} E\left(\bar{y}_{st} - \bar{Y} + \bar{Y} - Rp_{st}\right)^2$$

$$MSE(R_{st}) \cong \frac{1}{P^2} E\left(\left(\bar{y}_{st} - \bar{Y}\right) - R\left(p_{st} - P\right)\right)^2$$

$$\cong \frac{1}{P^2}\left(E\left(\bar{y}_{st} - \bar{Y}\right)^2 + R^2 E\left(p_{st} - P\right)^2 - 2RE\left(\bar{y}_{st} - \bar{Y}\right)\left(p_{st} - P\right)\right)$$

$$MSE(R_{st}) \cong \frac{1}{P^2}\left(\text{var}\left(\bar{y}_{st}\right) + R^2 \text{var}\left(p_{st}\right) - 2R\,\text{cov}\left(\bar{y}_{st} p_{st}\right)\right) \qquad (3.10)$$

$$\cong \frac{1}{P^2}\left(\sum_{h=1}^{k} W_h^2 \gamma_h S_{yh}^2 + R^2 \sum_{h=1}^{k} W_h^2 \gamma_h S_{\phi h}^2 - 2R \sum_{h=1}^{k} W_h^2 \gamma_h S_{y\phi h}\right)$$

$$\cong \frac{1}{P^2} \sum_{h=1}^{k} W_h^2 \gamma_h \left(S_{yh}^2 + R^2 S_{\phi h}^2 - 2RS_{y\phi h}\right) \qquad (3.11)$$

Substitute eqn. (3.11) into eqn. (3.9), we obtained

$$MSE\left(\hat{\bar{T}}\right) \cong \sum_{h=1}^{k} W_h^2 \gamma_h \left(S_{yh}^2 + R^2 S_{\phi h}^2 - 2RS_{y\phi h}\right) \qquad (3.12)$$

Where $\gamma_h = \left(\frac{1}{n_h} - \frac{1}{N_h}\right)$, $w_h = \frac{N_h}{N}$ is the weight of stratum, $R = \frac{\bar{Y}}{P}$ is the population ratio, n_h is the number of units in sample from stratum h, N_h is the population size of stratum h, $S_{yh}^2 = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} \left(y_i - \bar{Y}_h\right)^2$ is the population variance of study variable in stratum h, $S_{\phi h}^2 = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} \left(\phi_i - P_h\right)^2$ is the population variance of auxiliary attribute in stratum h and $S_{y\phi h} = \frac{1}{N - 1} \sum_{i=1}^{N_h} \left(y_i - \bar{Y}_h\right)\left(\phi_i - P_h\right)$ is the population covariance of study variable and auxiliary attribute in stratum.

3.6 BIAS AND MEAN SQUARE ERROR OF THE PROPOSED ESTIMATORS $\hat{\bar{T}}_i^*$

The estimators are defined as

$$\hat{\bar{T}}_i^* = \frac{\bar{y}_{st}^*}{p_{st_i}} P_i^* = R_{st_i}^* P_i^*, \ (i = 1, 2, ..., 10)$$

Let $R_i^* = \frac{\bar{Y}}{P_i^*}$ which implies that $\bar{Y} = R_i^* P_i^*$

From equation (3.3), we defined the expectation of the proposed estimators $\hat{\bar{T}}_i^*$, $(i = 1, 2, ..., 10)$ as

$$E\left(\hat{\bar{T}}_i^*\right) = P_i^* E\left(\frac{\bar{y}_{st}^*}{p_{st_i}^*}\right) \tag{3.13}$$

Then, the expectation of the estimators up to first order approximation using equation (3.4) is defined as

$$E\left(\hat{\bar{T}}_i^*\right) \cong \bar{Y}\left(1 - \frac{1}{P_i^* \bar{Y}} \operatorname{cov}\left(\bar{y}_{st}^* p_{st_i}^*\right) + \frac{1}{P_i^{*^2}} \operatorname{var}\left(p_{st_i}^*\right)\right) \tag{3.14}$$

Now,

$$\operatorname{cov}\left(\bar{y}_{st}^* p_{st_i}^*\right) = E\left(\bar{y}_{st}^* p_{st_i}^*\right) - E\left(\bar{y}_{st}^*\right) E\left(p_{st_i}^*\right)$$

$$= E\left(\left(\sum_{h=1}^{k} W_h \left(\bar{y}_h - b_{\phi h}\left(p_h - P_h\right)\right)\right)\left(\sum_{h=1}^{k} W_h \left(m_{1i} p_h + m_{2i}\right)\right)\right)$$

$$\operatorname{cov}\left(\bar{y}_{st}^{*} p_{st_i}^{*}\right) = E\left(\left(\bar{y}_{st} - \lambda\right)\left(m_{1i} p_{st} + m_{2i}\right)\right) - E\left(\bar{y}_{st} - \lambda\right) E\left(m_{1i} p_{st} + m_{2i}\right)$$

$$-E\left(\sum_{h=1}^{k} W_h \left(\bar{y}_h - b_{\phi h}\left(p_h - P_h\right)\right)\right) E\left(\sum_{h=1}^{k} W_h \left(m_{1i} p_h + m_{2i}\right)\right)$$

where $\lambda = \sum_{h=1}^{k} w_h b_{\phi h}\left(p_h - P_h\right)$

$$\operatorname{cov}\left(\bar{y}_{st}^{*} p_{st_i}^{*}\right) = m_{1i} E\left(\bar{y}_{st} p_{st}\right) + m_{2i} E\left(\bar{y}_{st}\right) - m_{1i} E\left(p_{st}\lambda\right) - m_{2i} E\left(\lambda\right) - m_{1i} E\left(\bar{y}_{st}\right) E\left(p_{st}\right)$$

$$-m_{2i} E\left(\bar{y}_{st}\right) + m_{1i} E\left(p_{st}\right) E\left(\lambda\right) + m_{2i} E\left(\lambda\right)$$

$$= m_{1i} \operatorname{cov}\left(\bar{y}_{st} p_{st}\right) - m_{1i} \operatorname{var}\left(\lambda p_{st}\right)$$

$$= m_{1i} \sum_{h=1}^{k} W_h^2 \gamma_h S_{y\phi h} - m_{1i} \sum_{h=1}^{k} W_h^2 \gamma_h b_{\phi h} S_{\phi h}^2$$

$$= m_{1i} \sum_{h=1}^{k} W_h^2 \gamma_h S_{y\phi h} - m_{1i} \sum_{h=1}^{k} W_h^2 \gamma_h S_{y\phi h}$$

$$\therefore \operatorname{cov}\left(\bar{y}_{st}^{*} p_{st_i}^{*}\right) = 0 \qquad (3.15)$$

$$\operatorname{var}\left(p_{st_i}^{*}\right) = \operatorname{var}\left(m_{1i} p_{st} + m_{2i}\right) = m_{1i}^2 \sum_{h=1}^{k} W_h^2 \gamma_h S_{\phi h}^2 \qquad (3.16)$$

Substitute equations (3.15) and (3.16) into eqn. (3.14), the expectation of the proposed estimators becomes

$$E\left(\hat{\bar{T}}_i^{*}\right) \cong \bar{Y}\left(1 + \frac{m_{1i}^2}{P_i^{*2}} \sum_{h=1}^{k} W_h^2 \gamma_h S_{\phi h}^2\right) \qquad (3.17)$$

Taking the limit of equation (3.17) as sample size increases indefinitely, $\lim_{n\to\infty} E\left(\hat{\bar{T}}_i^*\right)$, the sample size from different stratum tends to their respective stratum population, $n_h \to N_h$ and the quantity $\gamma_h = \left(\frac{1}{n_h} - \frac{1}{N_h}\right) \to 0$. Consequently, the term with γ_h in equation (3.17) approaches zero. Conclusively, we can say that $\lim_{n\to\infty} E\left(\hat{\bar{T}}_i^*\right) = \bar{Y}$ which implies that the proposed estimators are asymptotically unbiased.

Using equation (3.5), we defined the bias of the estimators $\hat{\bar{T}}_i^*$, $(i = 1, 2, ..., 10)$ as

$$Bias\left(\hat{\bar{T}}_i^*\right) = P_i^* Bias\left(R_{st_i}^*\right) \tag{3.18}$$

From equation (3.6) the bias of $R_{n_i}^*$ to first order approximation can be defined as

$$Bias\left(R_{st_i}^*\right) \cong -\frac{1}{P_i^{*^2}}\left(\operatorname{cov}\left(\bar{y}_{st}^* p_{st_i}^*\right) - R_i^* \operatorname{var}\left(p_{st_i}^*\right)\right) \tag{3.19}$$

Substitute equations (3.15) and (3.16) into eqn. (3.19), the bias of $R_{n_i}^*$ becomes

$$Bias\left(R_{st_i}^*\right) \cong \frac{1}{P_i^{*^2}} R_i^* m_{1i}^2 \sum_{h=1}^{k} W_h^2 \gamma_h S_{\phi h}^2 \tag{3.20}$$

Substitute eqn. (3.20) into eqn, (3.18), we obtain the bias of $\hat{\bar{T}}_i^*$ to first order approximation as

$$Bias\left(\hat{\bar{T}}_i^*\right) = \frac{m_{1i}^2 \bar{Y}}{P_i^{*^2}} \sum_{h=1}^{k} W_h^2 \gamma_h S_{\phi h}^2, \qquad (i = 1, 2, ..., 10) \tag{3.21}$$

Let $\tau_i = \frac{m_{1i}}{P_i^*}$, the bias of $\hat{\bar{T}}_i^*$ can be written in the form

$$Bias\left(\hat{\bar{T}}_i^*\right) \cong \tau_i^2 \bar{Y} \sum_{h=1}^{k} W_h^2 \gamma_h S_{\phi h}^2, \quad \left(i = 1, 2, ..., 10\right) \tag{3.22}$$

Where $\tau_1 = \frac{1}{P}$, $\tau_2 = \frac{1}{P + \beta_2(\phi)}$, $\tau_3 = \frac{1}{P + C_p}$, $\tau_4 = \frac{1}{P + \rho_{pb}}$, $\tau_5 = \frac{\beta_2(\phi)}{P\beta_2(\phi) + C_p}$

$\tau_6 = \frac{C_p}{PC_p + \beta_2(\phi)}$, $\tau_7 = \frac{C_p}{PC_p + \rho_{pb}}$, $\tau_8 = \frac{\rho_{pb}}{P\rho_{pb} + C_p}$, $\tau_9 = \frac{\beta_2(\phi)}{P\beta_2(\phi) + \rho_{pb}}$, $\tau_{10} = \frac{\rho_{pb}}{P\rho_{pb} + \beta_2(\phi)}$

Using equation (3.9), the mean square error of $\hat{\bar{T}}_i^*$ to first order approximation can be defined as

$$MSE\left(\hat{\bar{T}}_i^*\right) \cong P_i^{*^2} MSE\left(R_{st_i}^*\right) \tag{3.23}$$

From equation (3.10)

$$MSE\left(R_{st_i}^*\right) \cong \frac{1}{P_i^{*^2}} \left(\text{var}\left(\bar{y}_{st}^*\right) + R_i^{*^2} \text{ var}\left(p_{st}^*\right) - 2R_i^* \text{cov}\left(\bar{y}_{st}^* p_{st}^*\right) \right)$$

$$\cong \frac{1}{P_i^{*^2}} \left(\text{var}\left(\bar{y}_{st}^*\right) + R_i^{*^2} \text{ var}\left(p_{st}^*\right) \right)$$

$$\cong \frac{1}{P_i^{*^2}} \left(\text{var}\left(\sum_{h=1}^{k} W_h \left(\bar{y}_h - b_{\phi h}\left(p_h - P_h\right)\right) \right) + R_i^{*^2} \text{ var}\left(\sum_{h=1}^{k} W_h \left(m_{1i} p_h + m_{2i}\right) \right) \right)$$

$$\cong \frac{1}{P_i^{*^2}} \left(\sum_{h=1}^{k} W_h^2 \left(\text{var}\left(\bar{y}_h\right) + \text{var}\left(b_{\phi h}\left(p_h - P_h\right)\right) - 2\,\text{cov}\left(b_{\phi h} \bar{y}_h \left(p_h - P_h\right)\right) \right) + R_i^{*^2} \sum_{h=1}^{k} W_h^2 m_{1i}^2 \text{ var}\left(p_h\right) \right) \tag{3.24}$$

The value of $b_{\phi h}$ that minimizes equation (3.24) is $B_{\phi h}=\frac{S_{y\phi h}}{S_{\phi h}^{2}}=\frac{\rho_{pbh}S_{yh}}{S_{\phi h}}$. The $MSE\left(R_{n_i}^{*}\right)$ is now written as;

$$MSE\left(R_{st_i}^{*}\right)\cong\frac{1}{P_i^{*^2}}\left(\sum_{h=1}^{k}W_h^2\left(\text{var}\left(\bar{y}_h\right)+B_{\phi h}^2\,\text{var}\left(p_h-P_h\right)-2B_{\phi h}\,\text{cov}\left(\bar{y}_h\left(p_h-P_h\right)\right)\right)+R_i^{*^2}\sum_{h=1}^{k}W_h^2m_{1i}^2\,\text{var}\left(p_h\right)\right)$$

$$\cong\frac{1}{P_i^{*^2}}\left(\sum_{h=1}^{k}W_h^2\left(\gamma_hS_{yh}^2+B_{\phi h}^2\gamma_hS_{\phi h}^2-2B_{\phi h}\gamma_hS_{y\phi h}\right)+R_i^{*^2}m_{1i}^2\sum_{h=1}^{k}W_h^2\gamma_hS_{\phi h}^2\right) \tag{3.25}$$

Substitute $S_{y\phi h}=B_{\phi h}S_{\phi h}^2$ in equation (3.25), then we have

$$MSE\left(R_{st_i}^{*}\right)\cong\frac{1}{P_i^{*^2}}\left(\sum_{h=1}^{k}W_h^2\gamma_h\left(S_{yh}^2+B_{\phi h}^2S_{\phi h}^2-2B_{\phi h}S_{\phi h}^2\right)+R_i^{*^2}m_{1i}^2\sum_{h=1}^{k}W_h^2\gamma_hS_{\phi h}^2\right)$$

$$\cong\frac{1}{P_i^{*^2}}\left(\sum_{h=1}^{k}W_h^2\gamma_h\left(S_{yh}^2-B_{\phi h}^2S_{\phi h}^2\right)+R_i^{*^2}m_{1i}^2\sum_{h=1}^{k}W_h^2\gamma_hS_{\phi h}^2\right)$$

$$\cong\frac{1}{P_i^{*^2}}\sum_{h=1}^{k}W_h^2\gamma_h\left(S_{yh}^2-B_{\phi h}^2S_{\phi h}^2+R_i^{*^2}m_{1i}^2S_{\phi h}^2\right)$$

Substitute $B_{\phi h}=\frac{\rho_{pbh}S_{yh}}{S_{\phi h}}$

$$\therefore MSE\left(R_{st_i}^{*}\right)\cong\frac{1}{P_i^{*^2}}\sum_{h=1}^{k}W_h^2\gamma_h\left(S_{yh}^2\left(1-\rho_{pbh}^2\right)+R_i^{*^2}m_{1i}^2S_{\phi h}^2\right) \tag{3.26}$$

Substitute eqn.(3.26) into eqn.(3.23), we obtained

$$MSE\left(\hat{\bar{T}}_i^{*}\right)=\sum_{h=1}^{k}W_h^2\gamma_h\left(R_i^{*^2}m_{1i}^2S_{\phi h}^2+S_{yh}^2\left(1-\rho_{pbh}^2\right)\right), \tag{3.27}$$

Taking the limit of equation (3.27) as sample size increases indefinitely, $\lim_{n\to\infty} MSE\left(\hat{\bar{T}}_i^*\right)$, the sample size from different stratum tends to their respective stratum population, $n_h \to N_h$ and the quantity $\gamma_h = \left(\frac{1}{n_h} - \frac{1}{N_h}\right) \to 0$. Consequently, the term with γ_h in equation (3.27) approaches zero. Conclusively, we can say that $\lim_{n\to\infty} MSE\left(\hat{\bar{T}}_i^*\right) = 0$ which implies that the proposed estimators are asymptotically consistent.

Here we defined population ratio for the estimators $\hat{\bar{T}}_i^*$, $(i = 1, 2, ..., 10)$ as $R_i = R_i^* m_{1i}$, then $MSE\left(\hat{\bar{T}}_i^*\right)$ becomes

$$MSE\left(\hat{\bar{T}}_i^*\right) = \sum_{h=1}^{k} W_h^2 \gamma_h \left(R_i^2 S_{\phi h}^2 + S_{yh}^2 \left(1 - \rho_{pbh}^2\right)\right), (i = 1, 2, ..., 10) \qquad (3.28)$$

where $R_1 = \frac{\bar{Y}}{P}$, $R_2 = \frac{\bar{Y}}{P + \beta_2(\phi)}$, $R_3 = \frac{\bar{Y}}{P + C_p}$, $R_4 = \frac{\bar{Y}}{P + \rho_{pb}}$,

$R_5 = \frac{\bar{Y}\beta_2(\phi)}{P\beta_2(\phi) + C_p}$ $R_6 = \frac{\bar{Y}C_p}{PC_p + \beta_2(\phi)}$, $R_7 = \frac{\bar{Y}C_p}{PC_p + \rho_{pb}}$, $R_8 = \frac{\bar{Y}\rho_{pb}}{P\rho_{pb} + C_p}$,

$R_9 = \frac{\bar{Y}\beta_2(\phi)}{P\beta_2(\phi) + \rho_{pb}}$, $R_{10} = \frac{\bar{Y}\rho_{pb}}{P\rho_{pb} + \beta_2(\phi)}$

The biases and mean square errors of the proposed estimators are summarized in the table below;

Table 3.2: Biases and Mean Square Errors of the Proposed Estimators

ESTIMATOR	BIAS	MEAN SQUARE ERROR
$\hat{\bar{T}}_1^*$	$\frac{\bar{Y}}{P^2}\sum_{h=1}^{k} W_h^2 \gamma_h S_{\phi h}^2$	$\sum_{h=1}^{k} W_h^2 \gamma_h \left(\left(\frac{\bar{Y}}{P} \right)^2 S_{\phi h}^2 + S_{yh}^2 \left(1 - \rho_{pbh}^2 \right) \right)$
$\hat{\bar{T}}_2^*$	$\frac{\bar{Y}}{\left(P + \beta_2(\phi) \right)^2}\sum_{h=1}^{k} W_h^2 \gamma_h S_{\phi h}^2$	$\sum_{h=1}^{k} W_h^2 \gamma_h \left(\left(\frac{\bar{Y}}{P + \beta_2(\phi)} \right)^2 S_{\phi h}^2 + S_{yh}^2 \left(1 - \rho_{pbh}^2 \right) \right)$
$\hat{\bar{T}}_3^*$	$\frac{\bar{Y}}{\left(P + C_p \right)^2}\sum_{h=1}^{k} W_h^2 \gamma_h S_{\phi h}^2$	$\sum_{h=1}^{k} W_h^2 \gamma_h \left(\left(\frac{\bar{Y}}{P + C_P} \right)^2 S_{\phi h}^2 + S_{yh}^2 \left(1 - \rho_{pbh}^2 \right) \right)$
$\hat{\bar{T}}_4^*$	$\frac{\bar{Y}}{\left(P + \rho_{pb} \right)^2}\sum_{h=1}^{k} W_h^2 \gamma_h S_{\phi h}^2$	$\sum_{h=1}^{k} W_h^2 \gamma_h \left(\left(\frac{\bar{Y}}{P + \rho_{pb}} \right)^2 S_{\phi h}^2 + S_{yh}^2 \left(1 - \rho_{pbh}^2 \right) \right)$
$\hat{\bar{T}}_5^*$	$\frac{\left(\beta_2(\phi) \right)^2 \bar{Y}}{\left(\beta_2(\phi) P + C_p \right)^2}\sum_{h=1}^{k} W_h^2 \gamma_h S_{\phi h}^2$	$\sum_{h=1}^{k} W_h^2 \gamma_h \left(\left(\frac{\beta_2(\phi) \bar{Y}}{\beta_2(\phi) P + C_p} \right)^2 S_{\phi h}^2 + S_{yh}^2 \left(1 - \rho_{pbh}^2 \right) \right)$
$\hat{\bar{T}}_6^*$	$\frac{C_p^2 \bar{Y}}{\left(C_p P + \beta_2(\phi) \right)^2}\sum_{h=1}^{k} W_h^2 \gamma_h S_{\phi h}^2$	$\sum_{h=1}^{k} W_h^2 \gamma_h \left(\left(\frac{C_p \bar{Y}}{C_p P + \beta_2(\phi)} \right)^2 S_{\phi h}^2 + S_{yh}^2 \left(1 - \rho_{pbh}^2 \right) \right)$
$\hat{\bar{T}}_7^*$	$\frac{C_p^2 \bar{Y}}{\left(C_p P + \rho_{pb} \right)^2}\sum_{h=1}^{k} W_h^2 \gamma_h S_{\phi h}^2$	$\sum_{h=1}^{k} W_h^2 \gamma_h \left(\left(\frac{C_p \bar{Y}}{C_p P + \rho_{pb}} \right)^2 S_{\phi h}^2 + S_{yh}^2 \left(1 - \rho_{pbh}^2 \right) \right)$
$\hat{\bar{T}}_8^*$	$\frac{\rho_{pb}^2 \bar{Y}}{\left(\rho_{pb} P + C_p \right)^2}\sum_{h=1}^{k} W_h^2 \gamma_h S_{\phi h}^2$	$\sum_{h=1}^{k} W_h^2 \gamma_h \left(\left(\frac{\rho_{pb} \bar{Y}}{\rho_{pb} P + C_p} \right)^2 S_{\phi h}^2 + S_{yh}^2 \left(1 - \rho_{pbh}^2 \right) \right)$

Table 3.2: Continued

$\hat{\hat{T}}_9^*$	$\frac{\left(\beta_2(\phi)\right)^2\bar{Y}}{\left(\beta_2(\phi)P+\rho_{pb}\right)^2}\sum_{h=1}^{k}W_h^2\gamma_h S_{\phi h}^2$	$\sum_{h=1}^{k}W_h^2\gamma_h\left(\left(\frac{\beta_2(\phi)\bar{Y}}{\beta_2(\phi)P+\rho_{pb}}\right)^2 S_{\phi h}^2+S_{yh}^2\left(1-\rho_{pbh}^2\right)\right)$
$\hat{\hat{T}}_{10}^*$	$\frac{\rho_{pb}^2\bar{Y}}{\left(\rho_{pb}P+\beta_2(\phi)\right)^2}\sum_{h=1}^{k}W_h^2\gamma_h S_{\phi h}^2$	$\sum_{h=1}^{k}W_h^2\gamma_h\left(\left(\frac{\rho_{pb}\bar{Y}}{\rho_{pb}P+\beta_2(\phi)}\right)^2 S_{\phi h}^2+S_{yh}^2\left(1-\rho_{pbh}^2\right)\right)$

3.7 EFFICIENCY COMPARISONS

Comparison of $\hat{\hat{T}}$ with $\hat{\hat{T}}_i^*$ $(i=1,2,...,10)$ is made as follows;

$$MSE\left(\hat{\hat{T}}_i^*\right)<MSE\left(\hat{\hat{T}}\right)$$

$$\sum_{h=1}^{k}W_h^2\gamma_h\left(R_i^2 S_{\phi h}^2+S_{yh}^2-\rho_{pbh}^2 S_{yh}^2\right)<\sum_{h=1}^{k}W_h^2\gamma_h\left(S_{yh}^2+R^2S_{\phi h}^2-2RS_{y\phi h}\right)$$

$$R_i^2\sum_{h=1}^{k}W_h^2\gamma_h S_{\phi h}^2-\sum_{h=1}^{k}W_h^2\gamma_h\rho_{pbh}^2 S_{yh}^2<R^2\sum_{h=1}^{k}W_h^2\gamma_h S_{\phi h}^2-2R\sum_{h=1}^{k}W_h^2\gamma_h S_{y\phi h}$$

Let $A=\sum_{h=1}^{k}W_h^2\gamma_h S_{\phi h}^2$, $B=\sum_{h=1}^{k}W_h^2\gamma_h\rho_{pbh}^2 S_{yh}^2$, $C=\sum_{h=1}^{k}W_h^2\gamma_h S_{y\phi h}$

Then we have

$$R_i^2 A-B<R^2A-2RC$$

$$R_i^2 A-R^2A-B+2RC<0$$

$$A\left(R_i^2 - R^2\right) - B + 2RC < 0$$

Where there are two conditions as follows;

(i) When $\left(R_i^2 - R^2\right) > 0$

$$A - \frac{B - 2RC}{R_i^2 - R^2} < 0$$

$$A < \frac{B - 2RC}{R_i^2 - R^2}$$

(ii) When $\left(R_i^2 - R^2\right) < 0$

$$A - \frac{B - 2RC}{R_i^2 - R^2} > 0$$

$$A > \frac{B - 2RC}{R_i^2 - R^2}$$

When either of these conditions is satisfied, the proposed estimators $\hat{\bar{T}}_i^* \left(i = 1, 2, \ldots, 10\right)$ will be more efficient than the traditional estimator $\hat{\bar{T}}$.

3.8 PROPERTIES OF THE PROPOSED ESTIMATORS

The following properties are obtained from theoretical study;

Asymptotically Unbiased:- The proposed estimators $\hat{\bar{T}}_i^*\ (i=1,2,...,10)$ are all biased for $Bias\left(\hat{\bar{T}}_i^*\right)\neq 0$. However, $\lim_{n\to\infty} E\left(\hat{\bar{T}}_i^*\right)=\bar{Y}$. This means that the proposed estimators approach unbiasedness as the sample size increases indefinitely.

Efficiency:- The discrepancies among the units of the population, the strong positive correlation between the study variable and auxiliary attribute and the normality distribution of the population paved way for estimators $\hat{\bar{T}}_i^*\ (i=1,2,...,10)$ to satisfy the conditions for their efficiency over traditional estimator $\hat{\bar{T}}$. That is,

$$MSE\left(\hat{\bar{T}}_i^*\right)< MSE\left(\hat{\bar{T}}\right)\ (i=1,2,...,10),\ \text{if}\ B_2(\phi)>0, C_p>0 \text{ and } \rho_{pb}>0.$$

Consistency:- The proposed estimators $\hat{\bar{T}}_i^*, (i=1,2,3,...,10)$ are consistent for $\lim_{n\to\infty} E\left(\hat{\bar{T}}_i^*\right)=\bar{Y}$ and $\lim_{n\to\infty} MSE\left(\hat{\bar{T}}_i^*\right)=0$. This implies that the estimate of the proposed estimators can reach exact value of population mean, $\bar{Y}$ of variable of interest by indefinitely increasing the sample size.

3.9 DETERMINATION OF SAMPLE SIZE

The guiding principle in the determination of sample size is to choose them in such manner to estimate with desire precision for the minimum cost or with maximum precision for a given cost (Raj and Chandhok, 1998).

The total cost of taken n observations from population of size N is given by;

$$C = \sum_{h=1}^{k} c_h n_h \tag{3.29}$$

Where c_h is the cost per unit in the h^{th} stratum and $n = \sum_{h=1}^{k} n_h$.

The mean square error of the proposed estimators $\hat{\hat{T}}_i^* \left(i = 1, 2, ..., 10\right)$ is given by;

$$MSE\left(\hat{\hat{T}}_i^*\right) \cong \sum_{h=1}^{k} W_h^2 \gamma_h \left(R_i^2 S_{\phi h}^2 + S_{yh}^2 \left(1 - \rho_{pbh}^2\right)\right),$$

$$\cong \sum_{h=1}^{k} W_h^2 \gamma_h S_{Dhi}^2 \tag{3.30}$$

Where $S_{Dhi}^2 = S_{yh}^2 \left(1 - \rho_{pbh}^2\right) + R_i^2 S_{\phi h}^2$

To determine the optimum value of n when cost function and mean square error are represented by equations (3.29) and (3.30) respectively, we defined the following function;

$$f\left(n_h\right) = MSE\left(\hat{\hat{T}}_i^*\right) + \mu_0 C \tag{3.31}$$

$$= \sum_{h=1}^{k} W_h^2 \gamma_h S_{Dhi}^2 + \mu_0 C$$

$$= \sum_{h=1}^{k} \left(\frac{1}{n_h} - \frac{1}{N_h}\right) W_h^2 S_{Dhi}^2 + \mu_0 \sum_{h=1}^{k} c_h n_h$$

$$f\left(n_h\right)=\sum_{h=1}^{k}\left(\frac{W_h S_{Dhi}}{\sqrt{n_h}}-\sqrt{\mu_0 c_h n_h}\right)^2 + terms\, independent\, of\; n_h \tag{3.32}$$

Been a fixed constant, the desired precision for minimum cost or maximum precision for a given cost can be obtained when each square term of equation(3.32) is zero, that is,

$$\frac{W_h S_{Dhi}}{\sqrt{n_h}}-\sqrt{\mu_0 c_h n_h}=0$$

$$n_h=\frac{W_h S_{Dhi}}{\sqrt{\mu_0 c_h}} \tag{3.33}$$

The mean square error of the estimators $\hat{\bar{T}}_i^*$ of population mean $\bar{Y}$ is minimum for a given cost C_0 or the cost of taking sample is minimum for a given value V_0 of $MSE\left(\hat{\bar{T}}_i^*\right)$ when

$$n \propto \frac{W_h S_{Dhi}}{\sqrt{c_h}} \tag{3.34}$$

and $\frac{1}{\sqrt{\mu_0}}$ is the constant of proportionality. The allocation of sample size to different stratum according to equation (3.33) is known as optimum allocation (Raj and Chandhok, 1998). From equation (3.34), we deduced the followings;

- $n_h \propto W_h = \frac{N_h}{N}, \Rightarrow n_h \propto N_h$. This implies that the larger the stratum size, the larger the sample size from that stratum.
- $n_h \propto S_{Dhi}$. This implies that the more heterogeneity within the stratum, the larger the sample size from that stratum.

- $n_h \propto \frac{1}{\sqrt{c_h}}$. This implies that the cheaper the cost of the stratum, the larger the sample size from that stratum.

When c_h (cost per unit), is the same from stratum to stratum, say c, equation (3.34) reduces to

$$n_h \propto W_h S_{Dhi} \tag{3.35}$$

and the constant of proportionality becomes $\frac{1}{\sqrt{\mu_N}}$, where $\mu_N = \mu_0 c_h$. The allocation of sample size to different stratum according to equation (3.35) is known as Neyman allocation (Raj and Chandhok, 1998).

If the allocation is based on only the stratum size called Proportional allocation (Raj and Chandhok, 1998), equation (3.35) reduces to

$$n_h \propto w_h \tag{3.36}$$

and the constant of proportionality becomes $\frac{1}{\sqrt{\mu_p}}$, where $\mu = \frac{\mu_N}{S_{Dhi}}$.

The constant of proportionality in the various allocations can be evaluated so as to satisfy the condition of fixed cost or fixed precision.

3.9.1 Constants of Proportionality for Fixed Cost

For a given cost C_0, we have

$$C = \sum_{h=1}^{k} c_h n_h = C_0 \tag{3.37}$$

The Constants of Proportionality for various allocations are now determined as follows;

Optimum allocation: The sample size for each stratum is given by equation (3.33) as

$$n_h = \frac{W_h S_{Dhi}}{\sqrt{\mu_0 c_h}}$$

Substitute this term into equation (3.37), we obtained

$$\frac{1}{\sqrt{\mu_0}} = \frac{C_0}{\sum_{h=1}^{k} W_h S_{Dhi} \sqrt{c_h}} \tag{3.38}$$

Equation (3.38) is the constant of proportionality. Substitute the constant into equation (3.34), we obtained;

$$n_h = \frac{W_h S_{Dhi} C_0}{\sqrt{c_h} \sum_{h=1}^{k} w_h S_{Dhi} \sqrt{c_h}} \tag{3.39}$$

Taking summation over equation (3.39) for $h = 1, 2, ..., k$, we now have

$$n = \frac{C_0 \sum_{h=1}^{k} \frac{W_h S_{Dhi}}{\sqrt{c_h}}}{\sum_{h=1}^{k} W_h S_{Dhi} \sqrt{c_h}} \tag{3.40}$$

Thus equations (3.39) and (3.40) give the number of sample size in different stratum and total sample size respectively under optimum allocation with maximum precision and fixed cost.

Neyman allocation: The sample size for each stratum is given by equation (3.34) as

$$n_h = \frac{W_h S_{Dhi}}{\sqrt{\mu_N}}$$

Substitute this term into equation (3.37), we obtained

$$\frac{1}{\sqrt{\mu_N}} = \frac{C_0}{\sum_{h=1}^{k} W_h S_{Dhi} c_h} \tag{3.41}$$

Equation (3.41) is the constant of proportionality. Substitute the constant into equation (3.35), we obtained;

$$n_h = \frac{W_h S_{Dhi} C_0}{\sum_{h=1}^{k} W_h S_{Dhi} c_h} \tag{3.42}$$

Taking summation over equation (3.42) for $h = 1, 2, ..., k$, we now have

$$n = \frac{C_0 \sum_{h=1}^{k} W_h S_{Dhi}}{\sum_{h=1}^{k} W_h S_{Dhi} c_h} \tag{3.43}$$

Thus equations (3.42) and (3.43) give the number of sample size in different stratum and total sample size respectively under Neyman allocation with maximum precision and fixed cost.

Proportional allocation: The sample size for each stratum is given by equation (3.35) as

$$n_h = \frac{W_h}{\sqrt{\mu_p}}$$

Substitute this term into equation (3.37), we obtained

$$\frac{1}{\sqrt{\mu_p}} = \frac{C_0}{\sum_{h=1}^{k} W_h c_h} \tag{3.44}$$

Equation (3.44) is the constant of proportionality. Substitute the constant into equation (3.38), we obtained;

$$n_h = \frac{W_h C_0}{\sum_{h=1}^{k} W_h c_h} \tag{3.45}$$

Taking summation over equation (3.45) for $h = 1, 2, ..., k$, we now have

$$n = \left(\frac{C_0 \sum_{h=1}^{k} W_h}{\sum_{h=1}^{k} W_h c_h} \right) = \frac{C_0}{\sum_{h=1}^{k} W_h c_h} \tag{3.46}$$

Thus equations (3.45) and (3.46) give the number of sample size in different stratum and total sample size respectively under proportional allocation with maximum precision and fixed cost.

3.9.2 Constants of Proportionality for Fixed precision

The MSE of the estimators under arbitrary allocation is given by

$$MSE\left(\hat{\bar{T}}_i^*\right) = \sum_{h=1}^{k} W_h^2 \gamma_h S_{Dhi}^2 \, , \left(i = 1, 2, ..., 10\right)$$

For any fixed precision, say V_0

$$V_0 = \sum_{h=1}^{k} \left(\frac{1}{n_h} - \frac{1}{N_h} \right) W_h^2 S_{Dhi}^2$$

$$\sum_{h=1}^{k} \frac{W_h^2 S_{Dhi}^2}{n_h} = V_0 + \frac{1}{N_h} \sum_{h=1}^{k} W_h^2 S_{Dhi}^2 \, , \left(i = 1, 2, ..., 10\right) \tag{3.47}$$

The Constants of Proportionality for various allocations are now determined as follows;

Optimum allocation: The sample size for each stratum under optimum allocation is given by equation (3.33). Substitute this into equation (3.47), we obtained the constant of proportionality as

$$\frac{1}{\sqrt{\mu_0}} = \frac{\sum_{h=1}^{k} W_h S_{Dhi} \sqrt{c_h}}{V_0 + \frac{1}{N} \sum_{h=1}^{k} W_h S_{Dhi}^2} \tag{3.48}$$

Substitute equation (3.48) into equation (3.33), we obtained

$$n_h = \frac{W_h S_{Dhi} \sum_{h=1}^{k} W_h S_{Dhi} \sqrt{c_h}}{\sqrt{c_h} \left(V_0 + \frac{1}{N} \sum_{h=1}^{k} W_h S_{Dhi}^2 \right)} \tag{3.49}$$

Taking summation over equation (3.49) for $h = 1, 2, ..., k$, we now have

$$n = \frac{\sum_{h=1}^{k} W_h S_{Dhi} \sqrt{c_h}}{\left(V_0 + \frac{1}{N} \sum_{h=1}^{k} W_h S_{Dhi}^2 \right)} \sum_{h=1}^{k} \frac{W_h S_{Dhi}}{\sqrt{c_h}} \tag{3.50}$$

Thus equations (3.49) and (3.50) give the number of sample size in different stratum and total sample size respectively under optimum allocation with fixed precision and minimum cost.

Neyman allocation: The sample size for each stratum under Neyman allocation is given by equation (3.34). Substitute this into equation (3.47), we obtained the constant of proportionality as

$$\frac{1}{\sqrt{\mu_N}} = \frac{\sum_{h=1}^{k} W_h S_{Dhi}}{V_0 + \frac{1}{N} \sum_{h=1}^{k} W_h S_{Dhi}^2} \tag{3.51}$$

Substitute equation (3.51) into equation (3.34), we obtained

$$n_h = \frac{W_h S_{Dhi} \sum_{h=1}^{k} W_h S_{Dhi}}{\left(V_0 + \frac{1}{N} \sum_{h=1}^{k} W_h S_{Dhi}^2 \right)} \tag{3.52}$$

Taking summation over equation (3.52) for $h = 1, 2, ..., k$, we now have

$$n = \frac{\left(\sum_{h=1}^{k} W_h S_{Dhi} \right)^2}{\left(V_0 + \frac{1}{N} \sum_{h=1}^{k} W_h S_{Dhi}^2 \right)} \tag{3.53}$$

Thus equations (3.52) and (3.53) give the number of sample size in different stratum and total sample size respectively under Neyman allocation with fixed precision and minimum cost.

Proportional allocation: The sample size for each stratum under proportional allocation is given by equation (3.35). Substitute this into equation (3.47), we obtained the constant of proportionality as

$$\frac{1}{\sqrt{\mu_p}} = \frac{\sum_{h=1}^{k} W_h S_{Dhi}^2}{V_0 + \frac{1}{N}\sum_{h=1}^{k} W_h S_{Dhi}^2} \tag{3.54}$$

Substitute equation (3.54) into equation (3.35), we obtained

$$n_h = \frac{W_h \sum_{h=1}^{k} W_h S_{Dhi}^2}{\left(V_0 + \frac{1}{N}\sum_{h=1}^{k} W_h S_{Dhi}^2\right)} \tag{3.55}$$

Taking summation over equation (3.55) for $h = 1, 2, ..., k$, we now have

$$n = \frac{\sum_{h=1}^{k} W_h S_{Dhi}^2}{\left(V_0 + \frac{1}{N}\sum_{h=1}^{k} W_h S_{Dhi}^2\right)} \tag{3.56}$$

Thus equations (3.55) and (3.56) give the number of sample size in different stratum and total sample size respectively under proportional allocation with fixed precision and minimum cost.

CHAPTER FOUR

EMPIRICAL STUDY

4.1 PRE-AMBLE

The information on 1500 students taken from Students Pre-Medical Registration, Usmanu Danfodiyo University, Sokoto (2011/2012 Session) was used as data for empirical study. The height of the student is the variable of interest and their gender was used as auxiliary attribute (Male=1 and Female=0). The stratification is based on the faculties. By using Neyman allocation (Cochran, 1977),

$$n_h = n\frac{N_h S_{yh}}{\sum_{h=1}^{k} N_h S_{yh}} \tag{4.1}$$

We have computed sample size in each stratum. The summary information on the empirical data is given below;

$y =$ Height of the students

$\phi =$ Gender of the students

$N = 1500,\ n = 300,\ \bar{Y} = 158.572,\ P = 0.71,\ \rho_{pb} = 0.541,\ C_p = 0.639,\ \beta_2(\phi) = 1.85$

Table 4.1: Parameters Used for Estimation of MSE in each Stratum

S/N	Faculty	Stratum size	Sample Size	Stratum Parameters
1	Agric.	96	18	$S^2_{yh} = 69.813$ $S^2_{\phi h} = 0.134$ $S_{y\phi h} = 1.562$ $\rho_{pbh} = 0.51$
2	Vet. Medicine	100	18	$S^2_{yh} = 61.328$ $S^2_{\phi h} = 0.196$ $S_{y\phi h} = 1.471$ $\rho_{pbh} = 0.424$
3	Education	288	58	$S^2_{yh} = 76.159$ $S^2_{\phi h} = 0.189$ $S_{y\phi h} = 1.964$ $\rho_{pbh} = 0.518$
4	Arts and Islamic Studies	198	43	$S^2_{yh} = 90.651$ $S^2_{\phi h} = 0.221$ $S_{y\phi h} = 2.452$ $\rho_{pbh} = 0.548$
5	Law	136	29	$S^2_{yh} = 85.334$ $\rho_{pbh} = 0.586$ $S^2_{\phi h} = 0.22$ $S_{y\phi h} = 2.538$
6	Col. Of Health Sci.	95	19	$S^2_{yh} = 77.099$ $S^2_{\phi h} = 0.234$ $S_{y\phi h} = 2.384$ $\rho_{pbh} = 0.561$
7	Social Science	96	20	$S^2_{yh} = 78.99$ $S^2_{\phi h} = 0.223$ $S_{y\phi h} = 2.707$ $\rho_{pbh} = 0.645$
8	Science	299	61	$S^2_{yh} = 77.54$ $S^2_{\phi h} = 0.241$ $\rho_{pbh} = 0.572$ $S_{y\phi h} = 2.472$
9	Management Sci.	192	34	$S^2_{yh} = 60.248$ $S^2_{\phi h} = 0.176$ $S_{y\phi h} = 1.393$ $\rho_{pbh} = 0.428$

4.2 RESULTS AND DISCUSSION

This section deals with the empirical analysis of the data used in this research work so as to ascertain the performances of the proposed estimators in comparison to traditional estimator.

Table 4.2: Relative Bias and Efficiency

Est.	Bias	R. B.	MSE	R. E.	P. R.	$R_i^2 - R^2$	$X = \frac{B-2RC}{R_i^2 - R^2}$	Cond. For efficiency
$\hat{\hat{T}}$	0.1403	100.0	25.919	100.0	223.34	-	-	-
$\hat{\hat{T}}_1^*$	0.1730	123.31	27.581	106.41	223.34	0	Undefined	Not Satisfied
$\hat{\hat{T}}_2^*$	0.0252	17.96	4.142	15.98	85.25	-42613.19<0	5.73×10^{-5}	Satisfied
$\hat{\hat{T}}_3^*$	0.0479	34.14	7.440	28.71	117.55	-36062.75<0	6.78×10^{-5}	Satisfied
$\hat{\hat{T}}_4^*$	0.0557	39.70	8.982	34.65	126.76	-33812.66<0	7.23×10^{-5}	Satisfied
$\hat{\hat{T}}_5^*$	0.0545	38.85	8.778	33.87	125.29	-34183.17<0	7.15×10^{-5}	Satisfied
$\hat{\hat{T}}_6^*$	0.0138	9.84	2.340	9.03	63.18	-45889.04<0	5.32×10^{-5}	Satisfied
$\hat{\hat{T}}_7^*$	0.0360	25.66	5.852	22.58	101.87	-39503.26<0	6.19×10^{-5}	Satisfied
$\hat{\hat{T}}_8^*$	0.0244	17.39	4.011	15.48	83.85	-42848.93<0	5.70×10^{-5}	Satisfied
$\hat{\hat{T}}_9^*$	0.0626	44.62	10.07	38.85	134.33	-31836.21<0	7.67×10^{-5}	Satisfied
$\hat{\hat{T}}_{10}^*$	0.0108	7.70	1.864	7.19	55.92	-46753.71<0	5.23×10^{-5}	Satisfied

The satisfactory conduct of an estimator that leads to its preference over the conventional estimator is its small MSE which increases the percentage relative efficiency of the estimator. The table below shows the estimators with their respective bias, MSE, relative bias, relative efficiency and satisfaction of conditions for efficiency.

From table 4.2, it is observed that the values of biases and MSE of all proposed estimators with exception of $\hat{\bar{T}}_1^*$, are less compare to the values of bias and MSE of the traditional estimator. Also the proposed estimators with exception of satisfied the conditions that lead to their better performance than the traditional estimator numerically.

Results obtained from the empirical analysis shows that the estimator $\hat{\bar{T}}_{10}^*$ which uses kurtosis of auxiliary attribute and point-biserial correlation of auxiliary attribute and variable of interest to improve the origin and weight respectively of the auxiliary proportion has the least bias and MSE among all the estimators considered in this research work. This implies that as the distribution of the auxiliary attribute approaches normality and degree of positive relationship between variable of interest and auxiliary attribute approaches unity, the proposed estimator $\hat{\bar{T}}_{10}^*$ is the most efficient.

The next estimator with high efficiency and less bias is $\hat{\bar{T}}_6^*$ which uses kurtosis and coefficient of auxiliary attribute to improve the origin and weight respectively of the auxiliary proportion. This implies that in the absence of point-biserial correlation, if the distribution of the auxiliary attribute approaches normality and the proportion of the units in the population possessing the attribute is large, the proposed estimator gives the most efficient estimate and preferred to estimators.

Estimators $\hat{\bar{T}}_8^*$ and $\hat{\bar{T}}_2^*$ which are the next two estimators with less bias and MSE, performed almost equally. The first estimator uses two parameters while the second uses one and this implies that with few of parameters considered, estimator which is less bias and more efficient than the traditional estimator can be obtained.

The remaining estimators $\hat{\bar{T}}_7^*, \hat{\bar{T}}_3^*, \hat{\bar{T}}_5^*, \hat{\bar{T}}_4^*, \hat{\bar{T}}_9^*$ are also less bias and more efficient than the traditional estimator and they are arranged in order of preference.

Generally, we observed that the proposed estimators which use some known value of population proportion performed better than the traditional estimator.

CHAPTER FIVE

SUMMARY, CONCLUSION AND RECOMMENDATION

5.1 SUMMARY

Ten ratio-type estimators are proposed in stratified random sampling using attribute as auxiliary information. The estimators were formed from the traditional estimator by linearly transforming the mean of study variable and used the known parameters values of auxiliary attribute (kurtosis, coefficient of variation and point-biserial correlation) to modified the population proportion. The bias and MSE of the proposed estimators were obtained up to first order approximation using Taylors' expansion. The bias and MSE of the proposed estimators were compared with that of traditional estimator. The efficiency comparison is done both theoretically and empirically. The proposed estimators were also compared and the order of their preference is obtained. The properties of the proposed estimators were also stated based on the theoretical study. Formulae for determine sample sizes under various allocations (Optimum, Neyman and Proportion) for fixed cost and desired precision when applying the proposed estimators were also obtained.

5.2 CONCLUSION

In this study, ratio estimators are proposed for stratified random sampling using auxiliary information on attribute. The ten estimators proposed, nine of which use some known parameters values of the population proportion perform better and less bias than the traditional estimator. In addition, procedures for sample determination for Optimum, Neyman and Proportional allocations were derived for the estimators proposed.

5.3 RECOMMENDATION

The estimators $\hat{\bar{T}}_2^*, \hat{\bar{T}}_3^*, \hat{\bar{T}}_4^*, \hat{\bar{T}}_5^*, \hat{\bar{T}}_6^*, \hat{\bar{T}}_7^*, \hat{\bar{T}}_8^*, \hat{\bar{T}}_9^*$ and $\hat{\bar{T}}_{10}^*$ are recommended for use in any practical situation involving heterogeneous population and for which efficient and less bias estimate of population mean $\bar{Y}$ is needed.

REFERENCES

Al odat, N. A. (2009): Modification in Ratio Estimator Using Rank Set Sampling. *European Journal of Scientific Researc.* **29** (2): 265-268.

Beale, E.M.L. (1962): Some uses of Computers in Operation Research. *Industrielle organization*. **31**: 51-52.

Bedi, P. K. (1996): Efficient Utilization of Auxiliary Information at Estimation Stage. *Biomet. Journal*. **38** (8): 973-976.

Chaudhuri, A. and Adrikari, A. K. (1979): Improving on the efficiencies of standard sampling Strategies through variate transformations in estimators for finite population means. *Tech. Reports*, ISI Calcutta.

Cochran, W. G. (1942): Sampling Theory when the Sampling Units are of Unequal Sizes. *Journ. American Statistical Association*. **37**: 199-212.

Cochran, W. G. (1977): *Sampling Techniques*. New York: John Wiley and Sons.

Dash, P., Dash, M., Mishra, S. and Pattanaik, P. S. (2011): Different Classes of Transformed Ratio and Product Estimators of Finite Population Mean in Stratified Sampling. *Euro. Journal of Social Sciences*. **19** (2): 227-232.

Deming, W. E. (1956): *Sample Design in Business Research*. New York: John Wiley and Sons.

Hansen, M. H. (1951): *Sample Survey Methods and Theory*. New York: John Wiley and Sons.

Hartley, H. O. and Ross, A. (1954): unbiased Ratio Estimates. *Nature*. **174**: 270-271.

Hurvitz, D. G. (1952): Sampling and Field procedures of the Pittsburg Morbidity Survey. *Public Health Reports*. **67**: 1003-1012.

Jessen, R. J. (1942): Statistical Investigation of a Sample Survey for obtaining Farm Fact. *Lowa Agric. Expt. Stat. Res.* Bulletin, 304.

Jessen, R. J. (1978): *Statistical Survey Techniques*. New York: John Wiley and Sons.

Jhajj, H. S., Sharma, M. K. and Grover, L. K. (2006): A family of Estimators of Popn. Mean Using Information on Auxiliary Attributes. *Pak. Journ. Stat*. **22** (1):43-50.

John, L. (2008): The Expected Value of a Point-biserial Correlation. *Rasch Measurement Transaction*. **22**(1): 1154

Kadilar, C. and Cingi, H. (2003): Ratio Estimators in Stratified Random Sampling. *Biometrical Journal*. **45** (2): 218-225.

Kadilar, C. and Cingi, H. (2005): A New Ratio Estimator in Stratified Random Sampling. *Comm. Statist. Theory and Methods*. **34**: 1-6.

Madow, W. G. (1953): On the Limiting Distribution of Estimates Based on Samples from Finite Universes. *Ann. Math. Stat.*, 535-545.

Mahalanobis, P.C. (1994): Recent Experiment in Statistical Sampling in the Indian Statistical Institute, *Journal of Royal Statistical Society*. **109**: 325-370.

Mclntyre, G. A. (1952): A Method of Unbiased Selective Sampling Ranked Set. Australian, *Journal of Agricultural Research*. **3**: 385-390.

Mohanty, S. and Das, M. N. (1971). Use of transformation in sampling, *J. Ind. Soc. Agr. Stat.* **23**(2): 83–87.

Murthy, M.N. (1967): *Sampling Theory and Methods*. Statistical Publishing Society, Calcutta, India.

Murthy, M.N. (1964): Product Method of Estimation. *Sankhya*. **26**: 69-74.

Naik, V. D. and Gupta, P. C. (1996): A note on estimation with known population proportion of an auxiliary character. *Journ.Ind. Soc. Agr. Stat.* **48** (2): 151-158.

Prasad, B. (1989): Some improved Ratio type estimators of population mean and ratio in finite population sample surveys. *Commun. Stat. Theor. Meth*. **18**(1): 379-392.

Raj, D. and Chandhok P. (1998): *Sample Survey Theory*. Narosa Publishing House, London.

Rao, P.S.R.S. (1981a): Estimation of the mean square of the ratio in Krewski D., Platek R., and Rao J.N.K., eds (1981). *Current Topics in Survey Sampling*, Academic Press: 305-315.

Rao, P.S.R.S. (1981b): Efficiencies of nine two-phase ratio estimators for the mean. *Journal of the American Stat. Association*. **76** (376): 434-442.

Rao, J.N.K, Webster, J. T.(1966): On two methods of bias reduction in the estimation of ratios. *Biometrika*. **53:** 571-577.

Rao, J.N.K. and Beagle L.D. (1967): A monte Carlo Study of some ratio estimators. *Sankhya*. **B29**: 47-56.

Reddy, V. N. and Rao, T. J. (1977). Modified pps method of estimation, *Sankhya. C* **39**: 185–197.

Reddy, V. N. (1974). On a transformed ratio method of estimation, *Sankhya. C* **36**: 59–70.

Samawi, H. M., Muttlak, H. A. (1996). Estimation of Ratio Using Rank set Sampling. *Biometrical Journal*. **38**(6), 753-764.

Searls, D.T. (1964): The utilization of known coefficient of variation in the estimation procedure. *Journal of American Statistical Association*. **59**: 1125-1226.

Searls, D. T. and Intarapanich, P. (1990): A note on an estimator for the variance that utilizes the Kurtosis. *The American Statistician*. **44** (4): 295-296.

Sen, A. R. (1978): Estimation of the population mean when the coefficient of variation is known. *Commun.Statist.,Theory- Meth.* **A7**: 657-672.

Shabbir, J. and Gupta, S. (2006): A new estimator of population mean in stratified sampling. *Comm. Statis. Theory and Method*. **35** (7): 1201-1209.

Singh, R., Chauhan, P., Sawan, N. and Smarandache, F. (2008): Ratio Estimators in Simple Random Sampling Using Information on Auxiliary Attribute. *Pak.Journ. Stat. Operation Research*. **4**: 47-53.

Singh, G. N. (2003): On the improvement of product method of estimation in sample surveys. *Jour. Ind. Soc. Agric. Statistics*. **56**(3): 267-275

Singh, H. P. and Tailor, R. (2003): Use of known correlation coefficient in estimating the finite Population mean. *Statistics in Transition*. **64**: 555-560.

Singh, H.P. and Upadhyaya, L.N. (1999): Use of transformed auxiliary variable in estimating the Finite Population Mean. *Biometrical Journal*. **41** (5): 627-636.

Sulkhatme, P. V. (1947): *Sampling Theory of Survey with Application*. Lowa State College Press, Ames, Lowa.

Singh, H. P., Tailor,, R. and Kakran M. S. (2004): An improved estimator of population means Using power transformation. *Jour. Ind. Soc. Agr. Stat. Math*. **25**: 51-55.

Singh, J., Pandey, B. N. and Hirano, K. (1973): On the utilization of a known coefficient of kurtosis in the estimation procedure of variance. *Annal 0f Inst. Stat.Math.* **25**: 51-55.

Sisodaiya, B. V. S. and Dwivedi, V. k. (1981): A modified ratio estimator using coefficient of variation of auxillary variable. *Jour. Ind. Soc. Agr. Statistics*. **33** : 13-18.

Srivastava, S. K. and Jhajj, H. S. (1981): A class of estimators of the population mean in Survey Sampling using auxiliary information. *Biometrika*. **68**: 341-343.

Srivenkatarama, T. and Tracy, D. S. (1979): On ratio and product methods of estimation in Sampling. *Statistical Neerlandica*. **33**: 27-35.

Srivenkatarama, T. and Tracy D. S. (1980): An alternative to ratio method in sample survey. *Annals of the institute of Statistical Mathematics*. **32**: 111-120.

Srivenkatarama, T. and Tracy D. S. (1981): Extending product method of estimation to positive correlation case in surveys. *The Australian Journal of Statistics*. **23**(1): 95-100.

Srivenkatarama, T. (1980): A dual to ratio estimator in sample survey, *Biometrika*. **67**(1): 199-204.

Srivenkataramana, T. (1978). Change of origin and scale in ratio and difference methods of estimation in sampling, *Canad. J. Stat.* **6**: 79–86.

Takahasi, K. and Wakimoto, K. (1968): "On unbiased estimates of population mean based on the sample stratified by means of ordering". *Ann. Intern. Mathem. Stat.* **20**: 1-31.

Tin, M. (1965): Comparison of some Ratio Estimators. *Journl of the American statistical Association.* **60**: 294-307.

Upadhayaya, L.N. and Singh H. P. (1984): On the estimation of the population mean with known Coefficient of variation. *Biometrical Journal*. **26**(8): 915-922.

Yates, F. (1948): Systematic Sampling. *Phil. Trans. Royal Society London*. **A241**: 345-377.

APPENDIX I

DATA

Table AI: Data for Faculty of Law

S/N	SEX	HEIGHT (cm)	S/N	SEX	HEIGHT(cm)	S/N	SEX	HEIGHT(cm)
1	M	166	38	M	155	75	F	146
2	M	165	39	M	158	76	M	156
3	F	154	40	F	146	77	M	159
4	F	150	41	M	160	78	M	163
5	M	162	42	M	160	79	M	170
6	F	150	43	M	161	80	M	171
7	M	155	44	M	168	81	M	160
8	F	146	45	F	146	82	M	165
9	M	165	46	M	160	83	M	165
10	M	156	47	M	162	84	M	156
11	M	167	48	F	151	85	M	155
12	M	160	49	M	157	86	F	145
13	F	158	50	M	143	87	F	148
14	F	145	51	F	146	88	F	145
15	F	146	52	M	178	89	F	150
16	F	147	53	M	155	90	M	165
17	M	176	54	F	150	91	M	160
18	M	168	55	M	173	92	F	149
19	M	155	56	M	149	93	M	163
20	F	155	57	F	140	94	M	160
21	M	151	58	M	160	95	F	170
22	M	168	59	F	157	96	F	150
23	M	166	60	M	161	97	M	155
24	F	155	61	M	165	98	M	178
25	M	175	62	M	152	99	M	161
26	M	156	63	M	157	100	M	160
27	M	150	64	F	152	101	F	142
28	M	171	65	M	155	102	M	174
29	F	158	66	M	149	103	M	158
30	M	157	67	F	141	104	M	155
31	M	163	68	M	164	105	M	155
32	F	160	69	M	145	106	F	160
33	M	162	70	M	151	107	M	169
34	F	155	71	M	161	108	F	149
35	M	174	72	F	142	109	F	162
36	M	159	73	M	169	110	F	145
37	F	158	74	M	150	111	F	150

Table AI: Continued

S/N	SEX	HEIGHT (cm)	S/N	SEX	HEIGHT(cm)	S/N	SEX	HEIGHT(cm)
112	M	162	121	M	160	129	M	152
113	M	159	122	M	172	130	M	158
114	F	150	123	M	170	131	M	179
115	M	168	124	F	145	132	M	168
116	M	162	125	M	155	133	M	173
117	F	148	126	M	153	134	F	152
118	F	154	127	M	176	135	M	183
119	M	156	128	M	162	136	M	167
120	M	170						

Table AII: Data for Faculty of Agricultural Science

S/N	SEX	HEIGHT (cm)	S/N	SEX	HEIGHT(cm)	S/N	SEX	HEIGHT(cm)
1	M	170	25	M	174	49	F	148
2	M	160	26	F	141	50	M	154
3	M	150	27	M	165	51	M	154
4	M	165	28	M	162	52	M	160
5	M	156	29	M	152	53	M	168
6	M	166	30	F	150	54	M	168
7	M	160	31	F	157	55	M	157
8	M	148	32	F	158	56	M	154
9	M	165	33	M	165	57	M	158
10	M	157	34	F	137	58	M	174
11	F	157	35	F	155	59	M	162
12	M	170	36	M	160	60	M	148
13	M	167	37	M	172	61	M	161
14	M	168	38	M	161	62	M	161
15	M	180	39	M	166	63	M	156
16	M	165	40	M	169	64	M	165
17	M	173	41	M	154	65	M	167
18	M	169	42	M	157	66	M	170
19	F	165	43	M	156	67	M	167
20	M	156	44	M	160	68	M	172
21	M	166	45	M	157	69	M	172
22	M	153	46	M	172	70	M	163
23	M	167	47	F	155	71	M	165
24	M	160	48	M	161	72	M	163

Table AII: Continued

S/N	SEX	HEIGHT (cm)	S/N	SEX	HEIGHT(cm)	S/N	SEX	HEIGHT(cm)
73	M	153	81	M	156	89	M	160
74	F	141	82	F	154	90	F	149
75	M	163	83	M	165	91	M	170
76	M	154	84	M	168	92	F	150
77	M	160	85	M	175	93	M	145
78	M	150	86	M	162	94	M	164
79	M	143	87	M	168	95	M	166
80	M	157	88	M	164	96	M	164

Table AIII: Data for Faculty of Vetnary Medicine

S/N	SEX	HEIGHT (cm)	S/N	SEX	HEIGHT(cm)	S/N	SEX	HEIGHT(cm)
1	M	158	26	M	160	51	F	145
2	M	160	27	M	174	52	M	155
3	M	163	28	M	160	53	F	159
4	M	170	29	M	155	54	F	157
5	M	148	30	F	152	55	M	160
6	F	155	31	M	161	56	M	160
7	M	158	32	M	168	57	M	166
8	M	160	33	M	160	58	F	145
9	M	160	34	F	156	59	M	168
10	F	150	35	M	160	60	M	160
11	M	168	36	F	165	61	F	164
12	M	165	37	M	154	62	M	152
13	F	160	38	M	155	63	M	165
14	M	157	39	F	142	64	M	163
15	F	140	40	M	167	65	M	156
16	M	156	41	M	150	66	M	153
17	M	154	42	M	154	67	M	157
18	F	140	43	F	147	68	M	155
19	M	165	44	M	155	69	M	165
20	M	148	45	M	162	70	M	158
21	F	150	46	M	165	71	M	152
22	M	155	47	M	159	72	F	150
23	F	152	48	F	154	73	M	165
24	M	161	49	M	164	74	M	153
25	M	160	50	M	162	75	M	173

Table AIII: Continued

S/N	SEX	HEIGHT (cm)	S/N	SEX	HEIGHT(cm)	S/N	SEX	HEIGHT(cm)
76	M	154	85	F	154	94	F	150
77	M	160	86	F	156	95	M	179
78	F	144	87	M	160	96	F	164
79	M	165	88	M	161	97	M	170
80	F	165	89	M	170	98	M	179
81	M	160	90	M	169	99	F	169
82	M	162	91	M	155	100	M	169
83	M	160	92	M	155			
84	M	165	93	M	166			

Table AIV: Data for Faculty of Education

S/N	SEX	HEIGHT (cm)	S/N	SEX	HEIGHT(cm)	S/N	SEX	HEIGHT(cm)
1	M	161	25	M	166	49	M	165
2	M	164	26	M	175	50	M	147
3	M	161	27	M	165	51	M	159
4	M	168	28	M	161	52	M	167
5	M	165	29	M	155	53	F	147
6	M	159	30	M	174	54	M	167
7	M	168	31	M	160	55	M	142
8	M	160	32	M	167	56	M	166
9	M	157	33	M	156	57	M	153
10	M	161	34	M	148	58	M	163
11	M	162	35	M	163	59	M	157
12	F	151	36	M	160	60	M	168
13	M	150	37	F	157	61	M	155
14	F	153	38	M	160	62	M	167
15	M	170	39	M	163	63	M	159
16	M	156	40	M	141	64	M	165
17	M	164	41	M	162	65	M	172
18	M	161	42	M	152	66	M	161
19	M	165	43	M	154	67	F	146
20	F	160	44	M	163	68	M	161
21	M	161	45	M	163	69	M	165
22	M	166	46	M	161	70	M	155
23	M	175	47	M	162	71	M	160
24	M	155	48	M	150	72	M	161

Table AIV: Continued

S/N	SEX	HEIGHT (cm)	S/N	SEX	HEIGHT(cm)	S/N	SEX	HEIGHT(cm)
73	F	145	112	M	177	151	M	160
74	M	165	113	F	150	152	F	145
75	M	157	114	M	166	153	F	145
76	M	163	115	M	165	154	M	160
77	M	161	116	M	160	155	M	155
78	M	171	117	F	150	156	M	163
79	M	165	118	F	135	157	M	165
80	M	166	119	F	146	158	M	170
81	M	164	120	F	160	159	F	166
82	M	170	121	F	165	160	M	166
83	F	157	122	F	151	161	M	150
84	M	154	123	F	140	162	M	170
85	M	158	124	F	150	163	M	153
86	M	161	125	M	166	164	M	165
87	M	155	126	F	146	165	M	150
88	F	150	127	M	150	166	M	168
89	M	143	128	M	177	167	F	153
90	M	165	129	M	164	168	F	155
91	M	157	130	M	158	169	M	161
92	F	150	131	M	156	170	M	156
93	M	147	132	F	157	171	M	170
94	M	164	133	F	146	172	M	153
95	F	150	134	F	150	173	M	150
96	M	155	135	F	150	174	F	155
97	M	176	136	F	165	175	M	161
98	M	164	137	M	168	176	M	155
99	F	155	138	M	175	177	F	155
100	M	161	139	M	154	178	M	162
101	M	163	140	M	172	179	M	167
102	M	145	141	M	163	180	F	155
103	M	160	142	M	171	181	F	150
104	M	151	143	M	169	182	F	167
105	M	160	144	M	155	183	M	173
106	M	147	145	M	160	184	M	160
107	M	147	146	M	167	185	M	150
108	M	152	147	M	129	186	M	164
109	F	140	148	M	166	187	M	161
110	M	165	149	M	157	188	F	157
111	F	167	150	M	158	189	F	150

Table AIV: Continued

S/N	SEX	HEIGHT (cm)	S/N	SEX	HEIGHT(cm)	S/N	SEX	HEIGHT(cm)
190	M	151	223	F	155	256	M	154
191	M	168	224	M	166	257	F	158
192	M	163	225	M	168	258	F	158
193	M	160	226	M	164	259	M	170
194	F	150	227	F	140	260	F	158
195	M	150	228	M	160	261	F	147
196	M	165	229	M	166	262	M	156
197	F	153	230	M	159	263	M	168
198	M	164	231	M	173	264	M	165
199	M	164	232	M	161	265	M	170
200	M	147	233	M	161	266	M	145
201	F	138	234	M	157	267	M	166
202	M	155	235	F	152	268	M	140
203	M	162	236	M	165	269	M	161
204	M	156	237	M	160	270	F	155
205	M	166	238	M	165	271	M	149
206	F	142	239	M	155	272	M	168
207	M	163	240	M	160	273	M	160
208	F	135	241	M	161	274	M	167
209	M	158	242	F	172	275	M	163
210	M	166	243	M	158	276	F	142
211	M	155	244	F	140	277	F	142
212	M	170	245	F	146	278	M	155
213	M	162	246	M	166	279	F	160
214	M	165	247	M	164	280	F	149
215	M	162	248	F	146	281	M	165
216	M	168	249	M	163	282	M	169
217	M	149	250	M	158	283	M	170
218	M	156	251	M	145	284	M	178
219	M	165	252	M	167	285	M	167
220	F	143	253	M	169	286	M	156
221	F	153	254	F	158	287	M	158
222	M	150	255	F	164	288	M	171

Table AV: Data for College of Health Science

S/N	SEX	HEIGHT (cm)	S/N	SEX	HEIGHT(cm)	S/N	SEX	HEIGHT(cm)
1	M	165	33	M	155	65	M	153
2	F	151	34	M	169	66	M	150
3	M	167	35	F	145	67	M	152
4	M	157	36	F	145	68	M	147
5	M	164	37	F	150	69	M	155
6	M	156	38	M	162	70	M	151
7	M	158	39	M	160	71	F	154
8	M	164	40	F	163	72	M	172
9	M	165	41	M	160	73	M	167
10	M	173	42	F	154	74	M	154
11	F	158	43	M	166	75	F	155
12	M	159	44	M	156	76	F	161
13	M	167	45	F	150	77	M	166
14	M	150	46	F	148	78	M	164
15	M	163	47	F	146	79	F	142
16	F	143	48	F	154	80	F	156
17	M	154	49	M	168	81	F	150
18	F	152	50	M	160	82	M	165
19	F	148	51	M	161	83	M	151
20	M	158	52	M	157	84	F	147
21	F	148	53	M	170	85	M	151
22	M	150	54	F	146	86	F	156
23	M	155	55	F	147	87	F	160
24	F	140	56	M	146	88	M	172
25	F	152	57	M	161	89	M	175
26	M	163	58	M	169	90	M	156
27	M	170	59	M	155	91	M	159
28	M	166	60	F	135	92	M	173
29	M	162	61	M	172	93	F	168
30	F	156	62	M	160	94	M	170
31	F	148	63	M	157	95	M	174
32	M	152	64	F	150			

Table AVI: Data for Faculty of Social Science

S/N	SEX	HEIGHT (cm)	S/N	SEX	HEIGHT(cm)	S/N	SEX	HEIGHT(cm)
1	M	163	33	M	160	65	F	177
2	M	157	34	F	154	66	M	153
3	M	158	35	M	150	67	F	170
4	M	156	36	M	161	68	M	165
5	M	168	37	F	148	69	F	160
6	M	167	38	F	145	70	F	150
7	M	170	39	F	146	71	M	152
8	M	156	40	F	151	72	F	152
9	M	162	41	M	175	73	M	164
10	M	170	42	M	153	74	M	170
11	M	165	43	M	160	75	M	167
12	M	158	44	F	145	76	F	160
13	F	156	45	M	170	77	M	175
14	M	155	46	M	170	78	M	160
15	M	153	47	M	160	79	M	171
16	M	160	48	M	170	80	M	176
17	F	154	49	M	163	81	M	161
18	M	170	50	F	154	82	M	166
19	M	163	51	M	153	83	M	153
20	M	167	52	M	170	84	M	164
21	M	171	53	M	166	85	F	163
22	F	160	54	M	165	86	M	167
23	M	156	55	M	171	87	F	155
24	M	167	56	M	157	88	M	167
25	F	158	57	M	171	89	F	154
26	M	161	58	M	150	90	F	143
27	M	175	59	F	147	91	M	180
28	F	153	60	M	152	92	M	165
29	F	161	61	F	142	93	M	163
30	F	150	62	M	170	94	F	145
31	F	152	63	F	141	95	F	148
32	M	152	64	M	160	96	F	153

Table AVII: Data for Faculty of Science

S/N	SEX	HEIGHT (cm)	S/N	SEX	HEIGHT(cm)	S/N	SEX	HEIGHT(cm)
1	F	155	40	M	178	79	M	155
2	M	158	41	M	155	80	M	164
3	F	147	42	F	164	81	M	161
4	F	155	43	F	165	82	M	167
5	F	158	44	F	164	83	F	145
6	F	148	45	F	156	84	M	170
7	M	165	46	M	175	85	M	151
8	M	170	47	M	155	86	M	165
9	M	168	48	M	173	87	M	163
10	M	148	49	M	166	88	M	163
11	F	150	50	M	165	89	M	168
12	M	173	51	M	159	90	M	169
13	F	163	52	M	162	91	M	165
14	M	156	53	M	167	92	M	151
15	M	162	54	M	160	93	M	176
16	M	164	55	F	159	94	M	170
17	M	143	56	M	170	95	M	165
18	M	165	57	F	154	96	F	153
19	F	163	58	F	145	97	F	146
20	F	156	59	M	160	98	M	161
21	F	156	60	M	155	99	M	160
22	M	164	61	M	157	100	M	163
23	F	155	62	M	157	101	F	143
24	M	158	63	F	142	102	M	160
25	F	153	64	M	162	103	M	160
26	F	155	65	F	164	104	M	171
27	M	160	66	M	160	105	F	154
28	M	155	67	F	180	106	F	180
29	M	165	68	M	160	107	F	145
30	M	165	69	M	160	108	M	156
31	F	142	70	M	155	109	F	161
32	M	154	71	M	161	110	M	161
33	M	164	72	M	165	111	M	172
34	M	162	73	M	165	112	F	139
35	M	155	74	F	147	113	M	151
36	F	160	75	M	162	114	F	146
37	M	168	76	M	155	115	M	164
38	M	184	77	F	153	116	M	160
39	M	179	78	M	168	117	M	161

Table AVII: Continued

S/N	SEX	HEIGHT (cm)	S/N	SEX	HEIGHT(cm)	S/N	SEX	HEIGHT(cm)
118	F	154	157	F	152	196	M	162
119	M	170	158	F	150	197	M	165
120	F	164	159	M	155	198	M	170
121	M	180	160	M	172	199	M	165
122	M	177	161	M	162	200	F	138
123	F	145	162	M	166	201	M	154
124	M	161	163	F	155	202	M	160
125	M	167	164	M	145	203	M	165
126	F	151	165	F	153	204	F	154
127	M	162	166	F	153	205	F	162
128	M	163	167	F	144	206	M	150
129	M	159	168	M	157	207	M	167
130	F	146	169	F	145	208	F	145
131	F	164	170	M	156	209	M	165
132	M	158	171	M	165	210	F	152
133	M	155	172	M	168	211	F	151
134	M	155	173	M	163	212	M	170
135	M	145	174	M	160	213	F	152
136	M	170	175	F	155	214	M	156
137	F	162	176	F	153	215	M	165
138	M	169	177	M	165	216	F	145
139	M	164	178	F	148	217	M	165
140	M	155	179	M	152	218	M	159
141	M	162	180	M	158	219	M	155
142	M	155	181	M	152	220	F	145
143	M	160	182	M	145	221	M	165
144	F	145	183	M	145	222	M	157
145	F	140	184	M	153	223	F	147
146	F	145	185	M	166	224	M	172
147	F	148	186	F	145	225	F	149
148	F	139	187	M	165	226	F	159
149	F	149	188	M	168	227	M	155
150	M	156	189	M	152	228	M	169
151	M	159	190	M	155	229	M	164
152	F	147	191	F	146	230	F	140
153	M	155	192	M	162	231	F	145
154	M	164	193	M	165	232	M	153
155	M	155	194	M	146	233	F	156
156	M	157	195	F	144	234	F	153

Table AVII: Continued

S/N	SEX	HEIGHT (cm)	S/N	SEX	HEIGHT(cm)	S/N	SEX	HEIGHT(cm)
235	M	168	257	F	158	279	F	150
236	M	162	258	M	164	280	M	152
237	M	163	259	F	147	281	M	156
238	M	161	260	F	157	282	F	149
239	M	153	261	F	158	283	M	155
240	F	152	262	F	149	284	M	156
241	M	165	263	F	149	285	F	146
242	M	180	264	M	165	286	F	148
243	F	147	265	F	150	287	M	155
244	M	162	266	M	162	288	M	158
245	F	150	267	M	164	289	F	141
246	M	161	268	M	162	290	M	160
247	M	145	269	M	166	291	M	155
248	M	160	270	M	160	292	M	158
249	M	171	271	M	162	293	F	150
250	M	166	272	M	120	294	M	149
251	M	163	273	M	155	295	F	158
252	F	143	274	M	155	296	M	167
253	M	169	275	M	160	297	M	154
254	F	147	276	M	165	298	F	150
255	F	147	277	M	156	299	F	145
256	F	156	278	F	156			

Table AVIII: Data for Faculty of Management Science

S/N	SEX	HEIGHT (cm)	S/N	SEX	HEIGHT(cm)	S/N	SEX	HEIGHT(cm)
1	M	156	13	M	161	25	M	170
2	M	155	14	F	160	26	M	150
3	M	163	15	M	160	27	F	168
4	F	156	16	M	166	28	M	160
5	M	160	17	M	145	29	M	161
6	M	172	18	M	169	30	F	145
7	F	156	19	F	156	31	M	157
8	M	156	20	M	160	32	M	161
9	M	155	21	M	172	33	M	167
10	M	161	22	M	163	34	M	165
11	M	154	23	F	165	35	M	175
12	M	160	24	M	168	36	M	156

Table AVIII: Continued

S/N	SEX	HEIGHT (cm)	S/N	SEX	HEIGHT(cm)	S/N	SEX	HEIGHT(cm)
37	M	164	76	M	165	115	M	168
38	M	156	77	M	160	116	M	151
39	M	164	78	F	153	117	M	163
40	M	152	79	M	168	118	M	170
41	F	150	80	M	170	119	M	160
42	M	162	81	M	170	120	F	153
43	M	174	82	M	154	121	M	146
44	M	161	83	M	154	122	F	153
45	M	164	84	M	158	123	M	161
46	M	160	85	M	158	124	M	143
47	M	163	86	M	159	125	M	153
48	F	152	87	F	176	126	M	168
49	M	171	88	M	172	127	M	165
50	M	167	89	M	169	128	M	160
51	M	170	90	M	162	129	M	160
52	F	155	91	M	161	130	M	172
53	M	165	92	F	144	131	M	168
54	M	163	93	M	168	132	M	156
55	M	162	94	M	170	133	M	160
56	M	156	95	F	149	134	M	164
57	F	153	96	M	155	135	M	156
58	M	167	97	F	140	136	M	154
59	F	156	98	M	167	137	F	144
60	F	145	99	M	160	138	M	160
61	F	146	100	F	144	139	M	150
62	M	151	101	F	165	140	M	154
63	M	167	102	M	153	141	F	149
64	M	170	103	M	155	142	F	156
65	M	153	104	F	165	143	F	154
66	M	160	105	F	155	144	M	165
67	M	155	106	M	160	145	M	161
68	M	156	107	M	160	146	M	162
69	M	163	108	M	166	147	M	153
70	M	163	109	M	171	148	M	163
71	M	157	110	M	154	149	F	155
72	F	155	111	F	164	150	M	150
73	M	167	112	M	167	151	F	140
74	M	176	113	M	167	152	M	162
75	M	164	114	F	148	153	F	143

Table AVIII: Continued

S/N	SEX	HEIGHT (cm)	S/N	SEX	HEIGHT(cm)	S/N	SEX	HEIGHT(cm)
154	M	158	167	M	160	180	M	164
155	M	155	168	M	160	181	M	166
156	M	155	169	F	157	182	M	166
157	M	156	170	F	150	183	M	153
158	M	174	171	F	162	184	M	160
159	M	168	172	M	170	185	M	161
160	M	151	173	M	154	186	M	161
161	M	176	174	M	151	187	M	168
162	M	162	175	M	162	188	M	159
163	M	160	176	M	154	189	M	157
164	M	156	177	M	168	190	M	168
165	F	147	178	F	158	191	M	171
166	F	147	179	M	144	192	M	157

Table AIX: Data for Faculty of Arts and Islamic Studies

S/N	SEX	HEIGHT (cm)	S/N	SEX	HEIGHT(cm)	S/N	SEX	HEIGHT(cm)
1	M	171	23	M	157	45	M	146
2	M	159	24	M	163	46	F	158
3	M	160	25	M	155	47	M	163
4	M	180	26	M	164	48	M	166
5	M	160	27	M	157	49	M	145
6	M	162	28	M	160	50	F	171
7	M	174	29	M	165	51	M	160
8	M	171	30	M	160	52	M	157
9	M	165	31	M	170	53	M	153
10	M	161	32	F	152	54	M	165
11	M	159	33	M	156	55	M	150
12	F	146	34	F	150	56	M	164
13	M	152	35	F	151	57	F	159
14	M	158	36	M	157	58	M	163
15	M	143	37	F	154	59	F	150
16	F	145	38	M	161	60	F	153
17	M	154	39	M	156	61	M	162
18	F	153	40	M	151	62	M	167
19	M	161	41	M	160	63	F	147
20	M	157	42	M	170	64	M	159
21	M	164	43	F	156	65	M	163
22	M	163	44	M	169	66	M	159

Table AIX: Continued

S/N	SEX	HEIGHT (cm)	S/N	SEX	HEIGHT(cm)	S/N	SEX	HEIGHT(cm)
67	M	168	106	M	164	145	F	159
68	M	167	107	M	160	146	M	169
69	M	168	108	M	170	147	M	165
70	F	164	109	F	147	148	F	146
71	F	167	110	M	155	149	F	145
72	F	156	111	M	155	150	M	173
73	F	158	112	M	175	151	F	151
74	F	153	113	F	147	152	F	153
75	M	166	114	F	153	153	F	153
76	M	156	115	M	163	154	M	178
77	M	172	116	F	145	155	M	102
78	F	146	117	F	147	156	F	155
79	F	157	118	M	146	157	M	161
80	M	171	119	M	160	158	F	141
81	M	160	120	M	168	159	F	143
82	M	165	121	M	161	160	F	153
83	M	160	122	F	144	161	F	150
84	M	154	123	M	147	162	M	166
85	M	157	124	M	149	163	F	153
86	M	160	125	M	162	164	M	167
87	M	166	126	M	164	165	M	167
88	F	155	127	F	148	166	M	165
89	F	102	128	M	162	167	M	161
90	F	147	129	F	150	168	F	163
91	M	148	130	F	154	169	M	160
92	F	141	131	M	155	170	M	160
93	M	154	132	M	165	171	M	159
94	M	164	133	M	147	172	M	141
95	M	165	134	M	155	173	F	160
96	M	160	135	M	156	174	M	161
97	M	154	136	M	165	175	M	159
98	F	155	137	M	153	176	M	163
99	M	165	138	M	169	177	M	161
100	F	155	139	M	176	178	F	141
101	M	165	140	M	144	179	F	158
102	M	154	141	F	145	180	F	140
103	M	149	142	F	140	181	F	157
104	F	148	143	F	150	182	F	160
105	F	150	144	F	150	183	F	146

Table AIX: Continued

S/N	SEX	HEIGHT (cm)	S/N	SEX	HEIGHT(cm)	S/N	SEX	HEIGHT(cm)
184	M	154	189	M	167	194	F	147
185	M	160	190	M	166	195	M	161
186	M	174	191	F	150	196	M	147
187	M	156	192	M	164	197	M	173
188	M	164	193	M	165	198	F	141

APPENDIX II

SOURCE CODE

```
/**
 * @(#)MeanSquareError.java
 *
 * MeanSquareError application
 *
 * @author
 * @version 1.00 2012/11/28
 */
 import java.io.*;
 import java.util.*;
 import java.text.*;


public class MeanSquareError {
	static BufferedReader keybrd= new BufferedReader( new 
InputStreamReader(System.in));
    static DecimalFormat df=new DecimalFormat("0.000");
    public static void main(String[] args) throws IOException {
    double totalpopn, samplesize,popnProp,popnMean,corr,kurtosis,cv;
    int strata,noPara;
    System.out.print("Enter the Number of Strata in the Population=");
    strata=Integer.parseInt(keybrd.readLine());
    System.out.print("Enter the Number of pairs of two parameters=");
    noPara=Integer.parseInt(keybrd.readLine());
```

```
double[] list1= new double[strata];
double[] list2= new double[strata];
double[] list3= new double[strata];
double[] list4= new double[strata];
double[] list5= new double[noPara];
double[] list6= new double[noPara];
double[] list7= new double[noPara];
double[] list8= new double[noPara];
double[] list9= new double[strata];
double[] list10= new double[strata];
double[] list11= new double[noPara];
double list12;
double list13;
    System.out.print("Enter the population Size=");
    totalpopn=Double.parseDouble(keybrd.readLine());
    System.out.print("Enter the Sample Size=");
    samplesize=Double.parseDouble(keybrd.readLine());
    System.out.print("Enter the Population Proportion=");
    popnProp=Double.parseDouble(keybrd.readLine());
    System.out.print("Enter the Population Mean=");
    popnMean=Double.parseDouble(keybrd.readLine());
    System.out.print("Enter the population correlation=");
    corr=Double.parseDouble(keybrd.readLine());
    System.out.print("Enter the population kurtosis=");
    kurtosis=Double.parseDouble(keybrd.readLine());
```

```
System.out.print("Enter the population coefficient of variation=");
cv=Double.parseDouble(keybrd.readLine());
System.out.println(" ");
System.out.println("In this Section You Provide The Variances of Variable Of Interest for all Strata");
readVariances(list1);
System.out.println(" ");
System.out.println("In this Section You Provide The Variances of Auxiliary Attribute for all Strata");
readVariances(list2);
System.out.println(" ");
System.out.println("In this Section You Provide The Covariances for all Strata");
readVariances(list3);
System.out.println(" ");
System.out.println("In this Section You Provide The Correlation for all Strata");
readCorr(list4);
System.out.println(" ");
System.out.println("In this Section You Provide The Fisrt Parameters (m1) for all Strata");
readParram(list5);
System.out.println(" ");
System.out.println("In this Section You Provide The Second Parameters (m2) for all Strata");
readParram1(list6);
System.out.println(" ");
System.out.println("In this Section You Provide The sample sizes for all Strata");
readSample(list9);
```

```
        System.out.println(" ");
        System.out.println("In this Section You Provide The population sizes for all Strata");
        readPop(list10);
        list7=getBias(list5,list6,list2,list9,list10,popnMean,popnProp,totalpopn);
        list11=getRatio(list5,list6,popnProp,popnMean);
        list8=getMSE(list11,list1,list2,list9,list10,list4,totalpopn);
        list12=trdMSE(list1,list2,list3,list9,list10,popnMean,popnProp,totalpopn);
        list13=trdBias(list2,list3,list9,list10,popnMean,popnProp,totalpopn);
        printResult(list7,list8,list11,list12,list13);
    }
    public static void readVariances ( double[] variance) throws IOException{
        double var;
        for(int i=0; i<variance.length; i++){
            System.out.print("Enter the Variances or Covariance of Stratum "+(i+1));
            var=Double.parseDouble(keybrd.readLine());
            variance[i]=var;
        }
    }
    public static void readCorr ( double[] cor) throws IOException{
        double rho;
        for(int i=0; i<cor.length; i++){
            System.out.print("Enter the Correlation of Stratum "+(i+1));
            rho=Double.parseDouble(keybrd.readLine());
            cor[i]=rho;
```

```
        }
}
public static void readParram ( double[] parr) throws IOException{
        double par;
        for(int i=0; i<parr.length; i++){
                System.out.print("Enter the first Parameter for Stratum "+(i+1));
                par=Double.parseDouble(keybrd.readLine());
                parr[i]=par;
        }
}
public static void readParram1 ( double[] parr) throws IOException{
        double par;
        for(int i=0; i<parr.length; i++){
                System.out.print("Enter the second Parameter for Stratum "+(i+1));
                par=Double.parseDouble(keybrd.readLine());
                parr[i]=par;
        }
}
public static void readSample ( double[] parr) throws IOException{
        double par;
        for(int i=0; i<parr.length; i++){
                System.out.print("Enter the sample size for Stratum "+(i+1));
                par=Double.parseDouble(keybrd.readLine());
                parr[i]=par;
        }
```

```
    }
    public static void readPop ( double[] parr) throws IOException{
        double par;
        for(int i=0; i<parr.length; i++){
            System.out.print("Enter the Population size for Stratum "+(i+1));
            par=Double.parseDouble(keybrd.readLine());
            parr[i]=par;
        }
    }
    public static void printResult ( double[] list7,double[] list8,double[] list11,double
list12,double list13){
        System.out.println("-------------------------------------------------------------------------
---------------------");

        System.out.println("Estimator"+"\t"+"Bias"+"\t"+"Rel.Bias"+"\t"+"MSE"+"\t"+"
Rel. Eff."+"\t");
        System.out.println("-------------------------------------------------------------------------
---------------------");
        for(int i=0; i<list7.length; i++){

        System.out.println("T"+(i+1)+"\t"+list7[i]+"\t"+(list7[i]/list13)+"\t"+list8[i]+"\t"+
(list8[i]/list12));
        }
    }
    public static double[] getRatio ( double[] list5,double[] list6,double prop,double
popnMean){
        double[] ratio=new double[list5.length];
        double r;
```

```
        for(int i=0; i<list5.length; i++){
            r=(list5[i]*popnMean)/(list5[i]*prop+list6[i]);
            ratio[i]=r;
        }
        return ratio;
    }
    public static double[] getBias ( double[] list5,double[] list6,double[] list2,double[]
list9,double[] list10,double popnMean,double prop,double popnsize) {
        double[] bias=new double[list5.length];
        double[] tau=new double[list5.length];
        double t,b,sum1=0;
        for(int i=0; i<list5.length; i++){
            t=list5[i]/(list5[i]*prop+list6[i]);
            tau[i]=t;
        }
        for(int i=0; i<list9.length; i++)
            sum1+=(popnsize/list9[i])*(popnsize/list9[i])*(1/list10[i]-
1/list9[i])*list2[i];
        for(int i=0; i<list5.length; i++){
            b=sum1*popnMean*tau[i]*tau[i];
            bias[i]=b;
        }
        return bias;
    }
    public static double[] getMSE ( double[] ratio,double[] list1,double[] list2,double[]
list9,double[] list10,double[] list4,double popnsize) {
```

```
        double[] mse=new double[ratio.length];
        for(int i=0; i<ratio.length; i++){
                double sum2=0;
                for(int j=0; j<list9.length; j++){
                sum2+=(popnsize/list9[j])*(popnsize/list9[j])*(1/list10[j]-
1/list9[j])*(ratio[i]*ratio[i]*list2[j]+(list1[j]*(1-list4[j])));
                }
                mse[i]=sum2;
        }
        return mse;
    }
    public static double trdMSE (double[] list1,double[] list2,double[] list3,double[]
list9,double[] list10,double popnMean,double prop,double popnsize) {
        double mse1=0;
        for(int i=0; i<list9.length; i++)
                mse1+=(popnsize/list9[i])*(popnsize/list9[i])*(1/list10[i]-
1/list9[i])*(list1[i]+(popnMean/prop)*(popnMean/prop)*list2[i]+2*(popnMean/prop)*list
3[i]);
        return mse1;
    }
    public static double trdBias ( double[] list2,double[] list3,double[] list9,double[]
list10,double popnMean,double prop,double popnsize) {
        double bias1;
        double sum1=0;
        for(int i=0; i<list9.length; i++)
                sum1+=(popnsize/list9[i])*(popnsize/list9[i])*(1/list10[i]-
1/list9[i])*(popnMean/prop*list2[i]-list3[i]);
```

```
        bias1=sum1/prop;
        return bias1;
    }
}
```