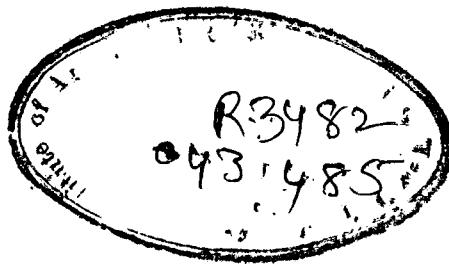


V 95

COMPARISION OF VARIOUS ESTIMATORS IN SAMPLING THROUGH MONTE-CARLO METHODS

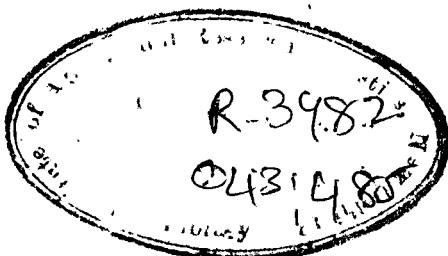
SRINIVASA BHAGAVAN



INSTITUTE OF AGRICULTURAL RESEARCH STATISTICS
(INDIAN COUNCIL OF AGRICULTURAL RESEARCH)
NEW DELHI

COMPARISON OF VARIOUS ESTIMATORS IN SAMPLING
THROUGH MONTE CARLO METHODS

SRIRIVASA MEAGAVAN



Dissertation submitted in fulfilment of the
requirements of Post-Graduate Diploma
in Agricultural Statistics of the
Institute of Agricultural
Research Statistics
New Delhi

INSTITUTE OF AGRICULTURAL RESEARCH STATISTICS

(I.O.A.R.)

NEW DELHI 12

1974

ACKNOWLEDGEMENTS

I have great pleasure in expressing my deep sense of gratitude to Dr.S.H.Banwani, Statistician-cum-Associate Professor, I.A.R.S., New Delhi, for his superb and unique guidance, keen interest and constant encouragement throughout the course of this investigation and preparation of the thesis.

I am grateful to Dr.D.Singh, Director, I.A.R.S., for providing adequate facilities and encouragement for the piece of research work done in the thesis.

I am greatly indebted to Sh.G.S.Pillai and Sh.K.Kher for their help in analysis of the data on electronic computer.

Lastly but not the least, I must thank my colleagues B.S.Kamesh, V.K.Gupta and P.Thangarajan for their kind co-operation in thesis writing.

I am also thankful to the Indian Council of Agricultural Research for providing financial assistance in the form of a fellowship.

(SRINIVASA BRAHMAN)

CONTENTS

CHAPTER	PAGE
INTRODUCTION	
I	MONTE CARLO METHOD AND REVIEW OF LITERATURE
II	DESCRIPTION OF POPULATION UNDER STUDY
III	COMPARISON OF SAMPLE MEAN AND RATIO ESTIMATORS
IV	COMPARISON OF SAMPLE MEAN AND REGRESSION ESTIMATORS
V	COMPARISON OF RATIO AND REGRESSION ESTIMATORS
VI	STUDY ON THE FREQUENCY DISTRIBUTION OF THE ESTIMATES OBTAINED BY THE SAMPLE MEAN, RATIO AND REGRESSION METHODS OF ESTIMATION
SUMMARY AND DISCUSSION	
APPENDIX	
REFERENCES	

I N T R O D U C T I O N

Generally theoreticians deduce conclusions from postulates, whereas experimentalists infer conclusions from observations. For those who care to hunt around and stretch a point experimental mathematics is as old as hills. The old testament (1 Kings vii 23 and 2 chronicles IV 2) refers to an early mathematical experiment on the numerical value of the constant . In this case the apparatus consisted of King Solomon's temple and the experimenter observed its columns to be about three times as great in girth as breadth. It would be nice to think that he inferred that this was a universal property of circular objects; but the text does not say so. The present study comprises application to sample survey techniques of that branch of experimental mathematics which is concerned with experiments on random numbers. This type of experimentation has been known by Monte-carlo methods.

The idea behind in these "Monte-carlo" studies is to postulate a model and then to construct one or more structures within the model by assigning specific numerical values to the parameters. In this manner an artifical population is generated which has the given features as the postulated popn. In order to judge the relative performance

of different estimators of the parameters of this population which in a way a sample from the super population, large number of repeated samples of different sizes are selected. For this purpose estimate of the population mean and variances were computed. The resulting frequency distributions of different estimators of the parameters based on varying sample sizes are studied along with the true values of the parameters. These frequency distributions give indication of the comparable efficiencies of different estimation procedures.

In the present study, four populations comprising of 200 units each were generated in such a manner that the auxiliary variate and study variate follows bivariate normal distribution, with pre-assigned values of means and standard deviations of the two variates. The generated population has got the correlation co-efficients (ρ) between the auxiliary and study variates ranging between 0.29 to 0.91. Actually the super bivariate population was selected had the value of ρ ranging between 0.3 to 0.9. In one of the populations, some of the study variates were changed so as to introduce a slight skewness and departure from normality. From each of the four populations of size 200 sampling units each, 200 independent samples of varying sizes ranging between 10 to 30 were selected at random. On the basis

of each sample, the population total was estimated by the three procedures of estimation viz., sample mean, ratio and regression, using the corresponding value of the auxiliary variates in the case of ratio and regression estimators.

In the absence of exact formulae of variance and mean square errors of ratio and regression estimators the present study is aimed at comparing relative performances of these procedures of estimation with sample mean through monte-carlo methods and finding out their frequency distribution for varying sample sizes. The study inter-alia suggests the validity of comparisons between three estimators based on approximate formulae of the variances of ratio and regression estimators as given in the standard books. These studies also enable us to see as to how far the approximate formulae of variances of ratio and regression estimators enables to compute accurately the confidence intervals of given confidence co-efficient. In addition to this study on likelihood of making a wrong decision when comparing ratio versus sample mean estimators based on a single sample.

CHAPTER - I

MONTE-CARLO METHOD AND REVIEW OF LITERATURE

The name and systematic development of Monte-carlo method dates from 1944. There were however a number of isolated and undeveloped instances on much earlier occasions as far back as in the second half of the 19th century when number of people performed experiments in which they threw a needle in a haphazard manner on a board ruled with parallel straight lines in order to infer the value of π , by observing the number of intersections between the needle and the lines. In fact one would be interested to know as to what is Monte-carlo method. The method may briefly be described as device of studying an artificial stochastic model of a physical or mathematical process. The original Von-Neumann concept of Monte-carlo method seems to have been that Monte-carlo specifically designated the use of random sampling procedures for treating deterministic mathematical problems.

Recently, there have been a number of applications of these methods for solving complicated partial differential equations and these methods which utilize various probability techniques and sampling procedures have found their application in various fields. These methods have been lately applied in the theory of econometrics, where it has been possible to compare different methods of estimating the parameters of a given economic model, specified by a

set of structural equations. The earliest use of Monte-carlo methods in solving statistical problems was perhaps made by "Student" who inferred the possible shape of the frequency distributions of 't' statistic prior to its mathematical derivation. Student drew a large number of samples from a normal population and worked out the distribution of the ratio of the mean to its standard deviation by plotting the frequency distribution of this ratio. J.N.K. Rao and Boosle (1967) carried out Monte-carlo studies for investigating the small sample efficiencies of different ratio estimators of the population ratio, \bar{Y}/\bar{X} assuming linear regression of \bar{Y} on \bar{X} and normal distribution of deviations. Recently Rao and Bayless (1969,70) have compared different unequal probabilities sampling estimators -ratio could be shown to reduce to Midzuno system Sen scheme(1952) of varying probabilities selection.

Monte-carlo studies in sample survey methodology by Mahajan(1972) for comparing sample mean, ratio and regression estimator~~s~~ was perhaps first of its kind. He generated ten hypothetical populations each of size 200 units satisfying the model

$$Y_i = \beta X_i + e_i$$

$$\text{where } V(e_i / X_i) = \gamma X_i^\lambda; \lambda \geq 1, \gamma > 0$$

$$\text{and } i = 1, 2, 3, \dots, n$$

corresponding to different values of " λ " varying between 1.50 to 4.00, the resultant population seems to follow

beta distribution. In this study it was demonstrated that ratio has an edge over regression estimator. However, a small draw back in this study is that the comparisions are based on only 25 samples which is rather small for these types of study.

J.N.K.Rao (1969) studied the behaviour of various types of ratio estimators experically. He used several sets of live data which represent a wide variety of populations of varying sizes ranging from 7 to 270 units, the co-efficient of variation of the auxilary variate ranging from 0.14 to 1.19 and the correlation co-efficient between the study and the auxilary variate ranging from 0.535 to 0.995. Accordingly he had drawn all the $\binom{N}{n}$ possible samples or 2000 samples whichever is smaller from a given set of data and compared simple ratio estimator with varying types(ten in number) of unbiased ratio estimators and computed their mean square errors and biases.

A.H.Manwani(1973) compared the variances of ratio and regression estimators through Monte-carlo method. He used several sets of actual field data taken from sample surveys on fruits and vegetables and live-stock, so as to demonstrate the method of comparing the efficiencies of ratio, regression and sample mean estimators and showed that comparision of efficiencies of ratio and regression estimators on the basis of approximate formulae of variance is likely to lead to wrong conclusions. These empirical studies on actual data indicate that the statement "Ratio is always inferior to regression or at the most has the same precision as that of regression" is not always

true. # Also the statement in the standard books to the effect that the regression estimator will always be more efficient than sample mean estimator when $\rho > 0$ is not always true. It has demonstrated that regression estimator could be very much inferior to sample mean estimator even though the value of ρ was positive.

CHAPTER - II

DESCRIPTION OF POPULATION UNDER STUDY

In sample surveys, the comparison of ratio and regression methods of estimation with sample mean estimator is generally made by assuming a regression model between the study variate (Y) and the auxiliary variate (X) specified by the equations.

$$Y_i = \alpha + \beta X_i + e_i$$

$$E(e_i / X_i) = 0, E(e_i e_j / X_i X_j) = 0, i \neq j$$

$$E(e_i^2 / X_i) = \sigma^2, i, i, j = 1, 2, 3, \dots, n$$

On the basis of this model it has been suggested that unless $\alpha=0$ ratio will always be less efficient as compared to regression. The two will be equally efficient when $\alpha=0$. As compared to sample mean, ratio is suggested to be more efficient in case $\rho > \pm C_x / C_y$, where ρ is the correlation coefficient between the study and the auxiliary variates and C_x and C_y are the co-efficient of variations in the auxiliary variate and study variate respectively. Also regression is said to be always more efficient than sample mean for $\rho > 0$.

Since the exact formulae of variances of both ratio and regression estimators are not easy to work out, the above conclusions are based on approximate formulae which

are correct to the first order of approximation of $1/n$. Perhaps it is assumed that when both the auxiliary variate (X) and the study variate (Y) follow normal distribution the contributions of the terms of the order $\geq 1/2$ and above to the variance s^2 would be negligible and the conclusion regarding the relative performance of the three estimators will be valid. For the present study, four populations each of size 200 units have been drawn from a bivariate normal population.

and $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$.

In otherwords, it is assumed that both X and Y variables are normally distributed with means and and variances respectively and X and Y are linearly related with correlation co-efficient given by ρ . The four finite populations generated for the purpose of the presentng study are drawn from the super-population with frequency distribution given by (1) by assigning arbitrary values of 50.0, 12.5, 200.0 & 100.0 to $\mu_1, \sigma_1, \mu_2 & \sigma_2$ respectively and four values of 0.3, 0.5, 0.7 & 0.9 to ρ . The choice for these values of the parameters of the population was made in such a manner that in each of the four population, generated corresponding to four values of ρ , the condition

$f > \frac{1}{2} c_x / c_y = 0.25$ (a pre-assigned value)

is satisfied. Under the above condition, ratio estimator has been stated to be more efficient than sample mean estimator.

In the actual sample survey data, we rarely come across negative values of study and auxiliary variates. Hence while drawing a finite population from the super-population(1) the program for generating the populations values was written in such a fashion that the probability of getting negative values in the sample was less than 0.01. As a matter of fact a total of 210 units was selected instead of 200. It was found that in none of the four populations samples corresponding to four values of β the number of negative values were more than ten in number. These negative values were replaced by positive one's from the additional drawings of 10 units. Thus each of the four resultant population consisted of 200 values all positive for both X & Y variates and these values are given in the appendix.

In order to study the effect of slight departure from normality on the comparison of various estimators, in one of the population (at Sl.No.2 with $\beta=0.5$) five values were arbitrarily changed thereby introducing a moderate skewness in that population. The means, standard deviations, co-efficient of variations, skewness ($\gamma_1 = \mu_3/\sqrt{\mu_2^3}$), kurtosis ($\gamma_2 = \frac{\mu_4}{\mu_2} - 3$) etc., from each of the generated population are given in Table 2.1.

For a normal curve the skewness and kurtosis should be zero but in practical situations we rarely come across exact values what we needed. In the three populations the values of γ_1 & γ_2 range between 0.0025 to 0.4203 and 0.0034 to 0.2563 respectively. The values of Charlier's

measure of skewness which is defined by

$$Sk(ch) = \frac{f(x_0) - \varphi(o)}{\varphi(o)}$$

Where $f(x_0)$ is the value of ordinate at mean of the normal frequency function(f) followed by the population under consideration and $\varphi(o)$ the value of the ordinate at mean of the normal distribution function.

have also been given in the Table 2.1. It can be seen that $Sk(ch)$ is very small for all the generated population except for the population-II, suggesting thereby that these generated populations are distributed normally. The values of kappa given in the last column of Table 2.1. are said to be near about zero except for population-II which has kappa value slightly higher. Hence we can conclude that generated populations are nearly normal except slight skewness in population-II.

We have thus generated four populations corresponding to four different values of ρ each being distributed normally except for the population-II wherein we have got a moderate skewness. The correlation co-efficient(ρ), regression co-efficient (β), and other important parameters of the above four populations are given in Table 2.2.

It will be seen from the Table 2.2 that the value of α in regression equation of Y on X is not zero. Hence the condition for ratio estimator to be equally efficient as regression is not satisfied. Accordingly for each of

the four population, we expect the following in regard to the relative efficiencies of three estimators commonly used in sample survey estimation.

(i) Since $\rho > \frac{C_x}{C_y} > 0$, both ratio and regression estimators are expected to be more efficient than sample mean.

(ii) Since $\alpha = 0$ for none of the four populations ratio estimator is expected to be biased though the extent of relative bias is likely to be small, as also it will be less efficient than regression.

In the following chapters we will examine how far the statements hold good by comparing the efficiencies of three estimators through empirical investigations called Monte-carlo method, since the formulae for actual variances of both ratio and regression estimators are not available.

TABLE 21

INTERCORRELATION COEFFICIENTS OF INFLUX AND STOCK VARIATES.

TYPE OF THE VARIATE	NAME NO	MEAN	STANDARD DEVIATION	CO-EFF OF CORR. OF INFLUX OR STOCK	INFLUX OR STOCK OF NUMBER OF BUSINESSES	
					NUMBER OF BUSINESSES OF GROWTH	NUMBER OF BUSINESSES OF SHRINKAGE
ANNUAL VARIATE	1.	50.2264	12.5283	0.2434	0.1527	-0.1723
	2.	49.5999	12.2123	0.2697	0.1718	-0.1138
	3.	49.8479	12.0407	0.2415	0.0003	0.3387
X	4.	50.0163	12.3510	0.2569	0.2075	0.0877
STOCH. VARIATE	1.	214.9260	52.8214	0.4364	0.0023	0.0034
	2.	305.1277	101.0260	0.4926	0.6515	2.1169
	3.	202.4512	69.8212	0.4437	0.4203	0.2563
Y	4.	204.3714	97.2911	0.4761	0.2167	0.2328

TABLE 32

BASIC STATISTICAL CHARACTERISTICS OF THE FOUR POPULATIONS UNDER STUDY

POP NO	COEFFICIENT OF CORRELATION	χ^2 (per cent) ^a	REGRESSION COEFFICIENT (%)	BESTFITTED REGRESSION EQUATION	BESTFITTED REGRESSION EQUATION (mm/sec)	BESTFITTED REGRESSION EQUATION (mm/sec)	BESTFITTED REGRESSION EQUATION (mm/sec)
1.	0.2938	0.2053	2.1773	103.6379	0.032566	4.3203	57.526138
2.	0.4655	0.2536	3.7945	19.5866	0.025910	4.1108	50.250170
3.	0.6140	0.2719	6.8045	- 27.0503	6721.7760	4.0614	4801.7823
4.	0.9120	0.2677	6.9230	- 162.6377	1773.7177	4.0253	4716.1161

- The value of χ^2 (per cent) was found as 0.25 while applying the super bivariate normal population.

as bestfit sum of squares from the fit line : $(\hat{y} - y_i)^2$

the deviation from the linear regression = $(y_i - \hat{y}_i)^2$

CHAPTER - III

COMPARISION OF SAMPLE MEAN AND RATIO ESTIMATORS

In this chapter, we are mainly interested in studying the bias in ratio estimators and to see how far the criterion $\rho > \frac{C_x}{C_y}$ helps in judging the relative efficiencies of the two estimators. In addition to this, study on likelihood of making a wrong decision when comparing ratio versus sample mean estimators based on a single sample.

The relative performance of the ratio versus sample mean estimator has been evaluated on the basis of approximate formula of variance,

$$V(\hat{T}_R) = V(\hat{T}) - 2 \text{COV}(\hat{x}, \hat{T}) + R^2 V(\hat{x}) \dots \dots \dots (3.1)$$

which is correct to the first order of approximation in $1/n$ assuming that the contribution of the terms of order $1/n^2$ and above is negligible.

In the absence of the exact formula of variance of ratio estimator, the exact variance has been estimated from 200 repeated samples of different sizes at random out of the totality of ${}^N C_n$ samples in the entire sample.

For this purpose, from the populations which have been generated in chapter-II satisfying the required assumption of bivariate normal distribution with pre-assigned values of ρ ranging between 0.3 to 0.9, 200 repeated samples of

varying sizes ranging from 10 to 30 are selected. The corresponding X-values have been considered as auxiliary variates for the use of ratio estimate. The selection of the sample at each drawing is made independent of the previous drawings. Each of these 200 samples provide estimate of the population mean also variance of the estimate as per the approximate formula (3.1) could be estimated.

3.1. Standard formulas of estimators and variances used in sample mean and ratio method of estimation.

Let y_1, y_2, \dots, y_n be the units selected in a sample of size n and x_1, x_2, \dots, x_n be the corresponding auxiliary variates.

Let \bar{y}_n be the sample mean of the y_i 's

\bar{x}_n be the sample mean of the x_i 's

($i=1, 2, \dots, n$)

\bar{Y}_N be the population mean of the Y_i 's

\bar{X}_N be the population mean of the X_i 's

($i=1, 2, \dots, N$)

and let $R = \bar{Y}_N / \bar{X}_N$

then the sample mean provides unbiased estimate of the population mean with the variance of the estimate given by

$$V(\bar{y}_{s.m.}) = \frac{N-n}{N,n} S_y^2$$

where S_y^2 is the population mean square of Y's and the estimate of the variance is given by

$$V(\bar{y}_{s.m.}) = \frac{N-n}{N,n} S_y^2 \quad \dots \dots \dots (3.2)$$

where s_y^2 is the sample mean square of y's after the

also the ratio estimate of \bar{Y}_H is given by

$$\bar{Y}_R = \left\{ \frac{\bar{Y}_H}{\bar{X}_H} \right\}^{\frac{1}{2}} \bar{X}_H = R_H \bar{X}_H$$

with variance

$$V(\bar{Y}_R) = \frac{n-n}{n(n-1)} (S_Y^2 + R_H^2 S_X^2 - 2R_H S_{XY})$$

where S_Y^2 , S_X^2 are the population mean squares of Y and X respectively and S_{XY} is their mean product

The estimate of $V(\bar{Y}_R)$ is given by

$$V(\bar{Y}_R) = \frac{n-n}{n(n-1)} \left\{ \sum_{i=1}^n y_i^2 + R_H^2 \sum_{i=1}^n x_i^2 - 2R_H \sum_{i=1}^n x_i y_i \right\} \dots (3.3)$$

As described in the beginning of this chapter, the formula (3.1) of variance of \bar{Y}_H is approximate, correct to the first order of $1/n$. The derivation of true variance is however a difficult task. The estimates of true variances of both the estimators could be worked out with the aid of Monte-carlo methods, which may briefly be described as follows.

From the given populations, 200 repeated samples of sizes varying between 10 to 30 sampling units are selected in independent sets consisting of 50 samples each. Accordingly, we get 200 sample mean and ratio estimates from the above 200 samples. Mean of these sample estimates provides unbiased estimate of the population mean-if the estimator is an unbiased one. The mean square between the sample estimates gives estimate of true variance of the frequency distribution of the

estimator(variance of the estimator). In case the estimator is biased(in case of ratio) the mean of the 200 ratio estimate will provide a biased estimate with estimate of the mean square error given by

$$EST(MSE_R) = EST[V(\bar{Y}_R)] + EST(\text{Bias})^2 \quad \dots \dots (3.4)$$

where estimate of the square of bias could be approximately substituted by $(\hat{\bar{Y}}_R - \bar{Y}_R)^2$

The methodology adopted for the calculation of the estimates of the exact variances for both sample mean and ratio and the mean square error of ratio estimator can be briefly described as follows.

Let $\bar{y}_1, \bar{y}_2, \dots, \bar{y}_s, \dots, \bar{y}_{200}$ be the mean of y 's for each of the 200 samples, all with mean given by \bar{y} and let $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_s, \dots, \bar{x}_{200}$ be the mean of the corresponding x 's for the 200 values. Then the estimate of the variance for sample mean estimate is given by

$$v(\bar{y}_{s.m.}) = \frac{1}{200-1} \sum_{s=1}^{200} (\bar{y}_s - \bar{y})^2$$

the ratio estimate of \bar{Y}_R for the 200 samples are given by

$$\bar{y}_{s(R)} = \left[\frac{\bar{y}_s}{\bar{x}_s} \right] \bar{x}_R \quad (s=1, 2, \dots, 200)$$

and the estimate of variance and mean square errors for the ratio estimate are given by

$$v(\bar{y}_R) = \frac{1}{200-1} \sum_{s=1}^{200} (\bar{y}_{s(R)} - \bar{y}_R)^2$$

where $\bar{y}_R = \frac{1}{200} \sum_{s=1}^{200} \bar{y}_{s(R)}$

$$MSE(\bar{Y}_R) = v(\bar{Y}_R) + b(\bar{Y}_R)^2$$

$$\text{where } b(\bar{Y}_R) = (\bar{Y}_R - \bar{Y}_N)$$

3.2 Study of bias in ratio estimator.

The estimate of relative bias in ratio estimator could be given by

$$\text{Rel(Bias)}_R = \frac{|\bar{Y}_R - \bar{Y}_N|}{\bar{Y}_N} \times 100 \quad \dots\dots\dots(3.5)$$

To get an idea about the extent of bias, the relative bias in ratio estimate for all the four populations for varying sample sizes have been presented in Table 3.1. It is found that the biases ranges from 0.01 to 1.24%. It is also found that the degree of bias decreases with the increase in sample size. In the absence of formula for obtaining the estimate of bias better than the one given above, the picture presented in the Table 3.1 is sufficient to indicate that the amount of bias in ratio estimator is negligible.

3.3 Comparison between sample mean and ratio estimator through variances and relative efficiencies.

The true variance of an estimator could be estimated by second moment of the frequency distribution of the estimator generated by the repeated sample drawings. Table 3.2 gives estimates of true variances of the two estimators based on sample sizes of 10, 15, 20, 25 & 30 for each of the four populations. From the table it is seen that the

T A B L E - 3.1

Relative bias in ratio estimator for varying
sample size (figures in percentage)

Sample size	Population			
	I (0.29)	II (0.46)	III (0.65)	IV (0.91)
10	1.08	0.34	0.15	0.16
15	0.48	1.07	0.10	0.08
20	0.50	0.30	0.35	0.32
25	0.33	0.27	0.02	0.33
30	0.57	0.60	0.88	0.02

(Figures in bracket are correlation co-efficient
between study and auxiliary variates)

T A B L E - 3.2

Variances of the sample mean and ratio estimators
and the mean square error of ratio estimator for
varying sample sizes for each of the four population

POPULATION - I

$$(\rho = 0.2908, \pm \sigma_x/\sigma_y = 0.2858)$$

Sample size	$v(\bar{y}_{s.m})$	$v(\bar{y}_R)$	$MSE(\bar{y}_R)$
10	623.4510	941.6248	946.9590
15	456.1952	665.1817	666.2288
20	363.2361	511.2543	512.4043
25	258.8815	384.2221	384.7146
30	202.5506	289.0541	290.9687

POPULATION - II

$$(\rho = 0.4585, \pm \sigma_x/\sigma_y = 0.2536)$$

Sample size	$v(\bar{y}_{s.m})$	$v(\bar{y}_R)$	$MSE(\bar{y}_R)$
10	1066.7980	732.9768	733.4541
15	728.7134	626.2741	631.0880
20	550.6492	377.3690	377.7352
25	427.4051	326.7152	327.0221
30	314.8065	244.8249	246.3228

..continued

Table 3.2 continued...

POPULATION - III
 $(\rho = 0.6440, \frac{\sigma_x}{\sigma_y} = 0.2719)$

Sample size	$v(\bar{Y}_{s,n})$	$v(\bar{Y}_R)$	$MSE(\bar{Y}_R)$
10	706.6736	370.3522	370.4396
15	550.6492	284.8599	284.9020
20	367.0556	184.8832	187.7820
25	309.9691	152.6272	152.6293
30	276.3077	148.5302	151.6943

POPULATION - IV
 $(\rho = 0.9150, \frac{\sigma_x}{\sigma_y} = 0.2697)$

Sample size	$v(\bar{Y}_{s,n})$	$v(\bar{Y}_R)$	$MSE(\bar{Y}_R)$
10	847.2107	233.8700	233.9732
15	517.1954	147.1174	147.1425
20	386.3782	103.4563	103.8812
25	296.8085	82.7405	83.1825
30	265.0045	67.0945	67.0958

sample mean estimator has variance which is consistently less than that of ratio in the case of first population for which $\rho = 0.2902$ which is slightly greater than $\frac{1}{2} C_x/C_y$. Thus even though, the criterion $\rho > \frac{1}{2} C_x/C_y$ holds good, we see that ratio is inferior to sample mean. We also see that the sample mean estimator has variance which is consistently greater than that of ratio for the remaining three populations which has got $\rho > 0.4540$. Hence we can conclude that the criterion $\rho > \frac{1}{2} C_x/C_y$ for ratio to be superior over sample mean is not true for the values of ρ near about $\frac{1}{2} C_x/C_y$ and small values of ρ . The difference $(\rho - \frac{1}{2} C_x/C_y)$ should be positive and sufficiently large for ratio to be superior to sample mean. We also observe that the variance of the sample mean estimator decreases with the increase in sample size approximately in proportion to $(\frac{1}{n} - \frac{1}{N})$, However it is noticed that the decrease in variance of ratio with increase in sample size is faster than that for the sample mean estimator.

The relative efficiency of ratio as compared to sample mean estimator based on mean square errors have been presented in Table 3.3. It is seen from the above table that efficiency is less than 71% for the first population and more than 100% for $\rho > 0.2902$ and there is a increasing trend in effect with the increasing values of ρ . It is also significant to note that efficiency does not appreciably show increasing trend with increase in the sample size. From the figures given in

FIGURE 2
PERCENTAGE EFFICIENCY (η) OR RATIO AS COMPARED TO SAMPLE SIZE BASED ON
THE SQUARE METERS.

Efficiency in Percentage for varying sample size.

SAMPLE SIZE NO	CON. COP.IMP (β^2)	Efficiency in Percentage for varying sample size.				
		10	15	20	25	30
I	0.2902	65.84	69.46	70.39	67.29	69.71
II	0.4383	163.45	112.67	145.77	130.70	127.60
III	0.6340	190.77	473.08	195.47	203.09	182.15
IV	0.9120	322.10	351.50	371.24	366.82	324.95

$$\dagger \quad \eta = (\bar{Y}(\bar{q}_0)/\bar{Y}(q_0)) \quad \times \quad 100$$

the Table 3.3 we find that in population which follows bivariate normal distribution, the following holds good.

(i) when values of ρ the correlation co-efficient between the study and the auxiliary variate lies in the range $0.30 > \rho > \frac{1}{2}C_x/C_y$

the sample mean estimator is more efficient than ratio. Loss in efficiency will be of the order of 30% irrespective of the sample size.

(ii) In population, where the values of ρ lies in the range $0.30 < \rho > \frac{1}{2}C_x/C_y$

the ratio estimator will be superior to sample mean estimator, the gain in efficiency increasing with the value of ρ . The gain in efficiency will be of the order of 70% for ρ around 0.50 and of the order of 290% or even more when $\rho=0.91$. The efficiency does not vary with sample size.

(iii) Even in the presence of slight skewness in the population -II where the value of ρ is 0.4585 which is greater than $\frac{1}{2}C_x/C_y$

the ratio estimator is superior over sample mean estimator.

3.4 Study on likelihood of making a wrong decision when comparison of ratio versus sample mean estimator is based on a given sample.

Generally, statistician has only one sample available for his study. The sample values are available both for auxiliary and study characters. The problem faced by the

statistician is to take decision about the right type of the estimator for building up an efficient estimate of the population mean. Natural practice is to suggest the type of estimator by looking at the information provided by the sample data. For example when the sample mean value of $\hat{f} > \frac{\hat{C}_x}{\hat{C}_y}$ where \hat{C}_x, \hat{C}_y are the co-efficient of variation for X and Y respectively, one may recommend ratio estimator in preference to sample mean estimator. This decision about the choice of an appropriate estimator on the basis of binary parameters estimated from a given sample is likely to go wrong. It will be of interest to calculate the probability-hereafter to be likelihood of such wrong decision. The choice of the estimator in all the four populations under study, the first order condition of $f > \frac{C_x}{C_y}$ for ratio to be superior estimator than sample mean is satisfied. However, we have already seen that in the first population where $f=0.30$, this condition is not sufficient since sample mean is consistently (for all sample sizes) better than ratio estimator. Hence so far as the first population is concerned, the criterion $r > \frac{C_x}{C_y}$, where r is the sample correlation co-efficient between x & y values and C_x/C_y is sample ratio of C_x/C_y , is not of much help. However for the other three populations where $f > \frac{C_x}{C_y}$ also implies that ratio is consistently better than sample mean, we would estimate the probability of wrong decision. One estimate of this probability is the proportional number of samples in which $r < \frac{C_x}{C_y}$. The Table 3.4 gives relevant values.

T A B L E - 3.4

Estimated probability of wrong decision about the choice between ratio and sample mean on the basis of criterion $r > \frac{1}{2} c_x/c_y$ calculated from estimators for a single sample of varying sizes

Population	Sample sizes				
	10	15	20	25	30
II	0.210	0.203	0.180	0.160	0.155
III	0.110	0.050	0.039	0.015	0.010
IV	0.000	0.000	0.000	0.000	0.000

T A B L E - 3.5

Estimated values of parameters of the equation $p = a e^{-bx}$

Population	β	a	b	R^2 (in %ges)
II	0.4585	0.1024	0.1787	80.72
III	0.6440	3.9540	0.9919	87.02

From the values given in the Table 3.4 it is not significant to note that in the population in which auxiliary character is highly correlated with the study character and the frequency distribution of the two variates is bivariate normal, the choice of estimator based on a single sample of even a moderate-size of 10 units will lead to almost surely correct decision. However for the populations with values of ρ between 0.5 to 0.9, the likelihood of wrong decision will be of the order 0.01 to 0.21 depending upon the sample size and the magnitudes of correlation co-efficient. Gupta(1973) established the relationship between the likelihood of wrong decision and sampling fraction. The two functions suggested by him were,

$$p = ae^{-bf} \dots\dots\dots(1)$$

$$P = af^{-b} \quad \dots\dots\dots (ii) \text{ for } a, b > 0$$

where p = likelihood of wrong decision in the choice of an estimator.

f = sampling fraction in percentage ($\times 100$)

The functions (i) and (ii) were fitted to values in Table 3.4 for the populations II & III, where ρ values are 0.4585 and 0.6440 respectively. It was observed that for function(i) the value of R^2 , the multiple correlation co-efficient for the above said populations was very high viz., 0.81 and 0.87. For the other function (ii) however the value of R^2 was observed to be 0.64 and 0.73. Hence we can conclude that the fit described by the first function is better than the second function. The same was also

found to be true in the case of populations described by Gupta which followed Beta type frequency curve. The least square estimates of the constants a & b of the function $p = ae^{-bx}$ are presented in Table 3.5.

CHAPTER - IV

COMPARISON OF SAMPLE MEAN AND REGRESSION ESTIMATORS

In the previous chapter, we have compared the sample mean and ratio estimators. Here we consider the comparison of sample mean and regression estimators. As in the previous case, here also the correlation co-efficient (ρ) plays an important role. From standard texts, we learn that regression is superior to sample mean if $\rho > 0$. We will examine this statement for the populations described in chapter II which satisfy the required assumption of bivariate normal distribution and were generated so as to have varying values of ρ ranging from 0.3 to 0.9. In fact all the four generated populations has got positive correlation. As before here also we are interested in the bias in regression estimators, comparison of sample mean and regression estimators.

The relative performance of the regression estimator is usually evaluated on the basis of approximate formulae of variance.

$$V(\hat{Y}) = \left(\frac{n-2}{n \cdot n} \right) S_y^2 (1-\rho^2) \dots \dots \dots \quad (4.1)$$

This formula is correct to the first order of approximation in $\frac{1}{n}$ assuming that the contribution of the terms of order $\frac{1}{n^2}$ and above is negligible. For the comparison sake, from the populations which have been generated in chapter II satisfying the required assumption of bivariate normal distribution with pre-assigned values of ρ ranging between 0.3 to 0.9, 200 repeated samples of

varying sizes ranging from 10 to 30 are selected. The corresponding X-values have been considered as the auxiliary variates for the use of regression estimate. The selection of the sample at each drawing is made independent of the previous drawings. Each of these 200 samples provide estimate of the population mean, variance of the estimate and other important parameters.

4.1 Standard formulae of estimators and variances used in regression method of estimation

Let y_1, y_2, \dots, y_n be the units selected in a sample of size n and x_1, x_2, \dots, x_n be the corresponding auxiliary variates.

Let \bar{y}_n be the sample mean of the y_i 's

\bar{x}_n be the sample mean of the x_i 's

$$(i=1, 2, \dots, n)$$

\bar{Y}_N be the population mean of the Y_i 's

\bar{X}_N be the population mean of the X_i 's

$$(i=1, 2, \dots, N)$$

then the regression estimate of the population mean is

$$\bar{Y}_{reg} = \bar{y}_n + b(\bar{X}_N - \bar{x}_n)$$

$$\text{where } b = \frac{\sum_{i=1}^n (y_i - \bar{y}_n)(x_i - \bar{x}_n)}{\sum_{i=1}^n (x_i - \bar{x}_n)^2}$$

$$v(\bar{Y}_{reg}) = \frac{N-n}{Na(n-1)} S_y^2 (1-\rho^2)$$

where S_y^2 is the population mean square of Y

and estimate of the variance is given by

$$v(\bar{Y}_{reg}) = \frac{N-n}{Na(n-1)} \left\{ \frac{\sum_{i=1}^n (y_i - \bar{y}_n)^2 - \frac{\{\sum_{i=1}^n (y_i - \bar{y}_n)(x_i - \bar{x}_n)\}^2}{\sum_{i=1}^n (x_i - \bar{x}_n)^2}}{\sum_{i=1}^n (x_i - \bar{x}_n)^2} \right\} . \quad (4.2)$$

The methodology adopted for the calculation of the estimate of the exact variance and the mean square error of the regression estimator can be briefly be described as follows.

Let $\bar{y}_1, \bar{y}_2, \dots, \bar{y}_s, \dots, \bar{y}_{200}$ be the mean of y 's for each of the 200 samples with mean given by \bar{y} and let $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_s, \dots, \bar{x}_{200}$ be the mean of the corresponding x 's for the 200 samples.

$$\text{then } \bar{y}_{s(\text{reg})} = \bar{y}_s + b(\bar{x}_N - \bar{x}_s)$$

$$\text{where } b = \frac{\sum_{s=1}^n (\bar{x}_s - \bar{y}_s)(x_s - \bar{x}_s)}{(x_s - \bar{x}_s)^2} \quad (s=1, 2, \dots, 200)$$

$$\text{and } v(\bar{y}_{\text{reg}}) = \frac{1}{200-1} \sum_{s=1}^{200} (\bar{y}_{s(\text{reg})} - \bar{y}_{\text{reg}})^2$$

$$\text{where } \bar{y}_{\text{reg}} = \frac{1}{200} \sum_{s=1}^{200} \bar{y}_{s(\text{reg})}$$

$$\text{MSE}(\bar{y}_{\text{reg}}) = v(\bar{y}_{\text{reg}}) + \{ b(\text{reg}) \}^2$$

$$\text{where } b(\text{reg}) = (\bar{y}_{\text{reg}} - \bar{y})$$

4.2 Study of bias in regression estimator.

The estimate of relative bias in regression estimator could be given by

$$\text{Rel(Bias)}_{\text{reg}} = \frac{|\bar{y}_{\text{reg}} - \bar{y}_N|}{\bar{y}_N} \times 100 \quad \dots \dots \dots \quad (4.3)$$

To get an idea about the extent of bias, the relative bias in regression estimate for all the four populations for varying sample sizes are presented in Table 4.1. It is found that the biases ranges from 0.01 to 1.75%.

It is also found that the degree of bias decreases with the increase in sample size. In the absence of formula for obtaining the estimate of bias, better than the one given above, the picture presented in Table 4.1 is sufficient to indicate that the extent of bias in regression estimator is negligible.

4.3 Comparison between sample mean and regression estimator through variances and relative efficiencies

The true variance of an estimator could be estimated by the second moment of the frequency distribution of the estimator generated by the repeated sample drawings. Table 4.2 give estimates of true variances of the two estimators based on sample sizes of 10,15,20,25 & 30 units for each of the four populations. From the table 4.2 it is seen that in the case of the first population with $\rho = 0.29$ the sample mean estimator has variance which is consistently less than (for all sample sizes) that of regression estimator. Eventhough $0 < \rho < 0.30$ we see that regression is inferior to sample mean. We also see that the sample mean estimator has variance which is consistently greater than that of regression for the remaining three populations with the correletion co-efficient greater than 0.4540. Hence we can conclude that the criterion $\rho > 0$ for regression to be superior over sample mean estimator is not always true, but the case will yield fruitful results if ρ is sufficiently large.

We also observe that the variance of the sample mean estimator decreases with the increase in sample size approximately in proportion to $(\frac{1}{n} - \frac{1}{N})$, however it is noticed that the decrease in variance of regression with

T A B L E - 4,1

Relative bias in regression estimator for
varying sample size (Figures in percentage)

Sample size	POPULATION			
	I (0.29)	II (0.46)	III (0.65)	IV (0.91)
10	1.07	0.61	1.75	0.31
15	0.29	1.36	0.49	0.19
20	0.23	0.19	1.49	0.53
25	0.01	0.55	0.41	0.32
30	0.62	0.45	0.89	0.01

TABLE - 4.2

Variances of the sample mean and regression estimators and mean square error of regression estimator for varying sample sizes for each of the four population

POPULATION - I
($\rho = 0.2908$, $\frac{\sigma_x}{\sigma_y} = 0.2858$)

Sample size	$v(\bar{y}_{s.m})$	$v(\bar{y}_{reg})$	$MSE(\bar{y}_{reg})$
10	625.4510	942.5775	947.8261
15	456.1952	615.9975	616.3979
20	365.2361	436.6483	436.9476
25	258.8815	321.0747	321.0755
30	202.5506	247.3526	249.1286

POPULATION - II
($\rho = 0.4985$, $\frac{\sigma_x}{\sigma_y} = 0.2936$)

Sample size	$v(\bar{y}_{s.m})$	$v(\bar{y}_{reg})$	$MSE(\bar{y}_{reg})$
10	1066.7980	745.1671	746.7582
15	728.7134	624.6047	632.3431
20	550.6492	423.9109	424.0590
25	427.4051	358.0418	359.3307
30	314.8069	244.3642	245.2166

Continued...

Table 4.2 continued..

POPULATION - III ($\rho = 0.6440$, $t\sigma_x/\sigma_y = 0.2697$)			
Sample size	$v(\bar{y}_{\text{sum}})$	$v(\bar{y}_{\text{reg}})$	$MSE(\bar{y}_{\text{reg}})$
10	706.6738	427.8542	440.4602
15	550.0844	353.1182	354.1204
20	367.0556	189.9682	199.1085
25	309.9691	156.7796	157.4801
30	276.3077	144.5539	147.8054

POPULATION - IV ($\rho = 0.9130$, $t\sigma_x/\sigma_y = 0.2697$)			
Sample size	$v(\bar{y}_{\text{sum}})$	$v(\bar{y}_{\text{reg}})$	$MSE(\bar{y}_{\text{reg}})$
10	847.2107	152.6346	153.0322
15	517.1954	88.3370	88.4824
20	386.3782	54.1942	55.3567
25	298.8085	44.8021	45.2361
30	265.0045	32.8435	32.8437

TABLE - 4.3

**MEASURING IMPACT (%) OR INACCURACY AS COMPARED TO SIMPLE MEAN
BASED ON REAL SQUINT MEAS.**

ROW NO.	(P)	INFLATION IN Percentages For various sample sizes				
		10	15	20	25	30
I	0.2908	115.70	114.01	82.97	60.43	51.59
II	0.4203	142.35	113.24	129.05	110.24	120.20
III	0.6410	160.44	125.24	104.35	96.83	106.94
IV	0.9120	202.62	201.52	197.98	186.13	196.57

$$\dagger \quad E = [\bar{y}(\bar{\psi}_{S,m}) / MSE(\bar{\psi}_{reg})] \times 100$$

increase in sample size is much faster than that for the sample mean estimator.

The relative efficiency of regression as compared to sample mean estimator based on mean square error have been presented in Table 4.3. It is seen from the above table that efficiency of regression estimator is less than 82% in the population ~ I . For the remaining three populations wherein $\rho > 0.4585$ efficiency of regression estimator is greater than 100%. There is an increasing trend in effect with the increasing values of ρ . It is also significant to note that efficiency does not appreciably show increasing trend with in the sample size. This shows that like ratio, the efficiency of regression estimator versus sample mean estimator does not depend on sample size.

From the figures given in Table 4.3. we find that in porulation which follows bivariate normal distributionthe following holds good.

(i) when values of ρ - the correlation co-effcient between the study and the auxilary variate lies in the range $0.30 > \rho > 0$

- the sample mean estimator is more efficient than regression. Loss in efficiency being of the order of 20 to 35%

(ii) In population, where the value of ρ lies in the range $0.30 < \rho > 0$

- the regression estimator will be superior to

sample mean estimator, the gain in efficiency will be of the order 70% for ρ around 0.50 and of the order of 700% or even more when $\rho=0.91$.

(iii) Even in the presence of slight skewness in the population-II where the value of ρ is 0.4585 which is greater than zero.

-the regression estimator is superior over sample mean estimator.

CHAPTER - V

COMPARISON OF RATIO AND REGRESSION ESTIMATORS

In the previous two chapters we have compared sample mean with ratio and sample mean with regression estimators. In this chapter we would like to compare ratio with regression estimator. For this purpose the mean square errors already obtained are described in Table 5.1. The methodology adopted in computing the variances has already been discussed in the previous two chapters. We have already seen that the biases in both ratio and regression estimators are negligible. From the Table 5.1. we see that regression is superior than ratio for the first population for all the sample sizes except for sample size 10. Infact we have already seen that when ρ is round about 0.3 both the methods are inefficient and regression has got an edge over ratio for $n > 10$. For the populations II & III where the values of ρ are 0.4585 and 0.6440 respectively, we observe that ratio is consistently superior over regression for all the sample sizes except for the sample size 50. For the fourth population which has $\rho = 0.9150$ regression seems to be highly superior over ratio. Hence we can conclude that for smaller values of ρ (as small as 0.3) although both the methods are inefficient, regression is better than ratio. For ρ ranging from 0.3 to 0.7 ratio seems to be somewhat better than regression for smaller sample sizes and as the value of ρ approaches 1 regression comes out to be best method of estimation.

T A B L E - 5.1

Mean square errors of the ratio and regression estimators for varying sample sizes for each of the four population

POPULATION - I

($\beta=0.2908$, $\frac{\sigma_x}{\sigma_y} = 0.2858$)

Sample size	MSE(\bar{Y}_R)	MSE(\bar{Y}_{reg})
10	946.9590	947.8261
15	666.2288	616.5979
20	512.4043	436.9476
25	384.7146	321.0755
30	290.5687	249.1286

POPULATION - II

($\beta=0.4585$, $\frac{\sigma_x}{\sigma_y}=0.2536$)

Sample size	MSE(\bar{Y}_R)	MSE(\bar{Y}_{reg})
10	733.4541	747.7582
15	631.0880	632.3431
20	577.7352	424.0590
25	527.0221	359.5307
30	246.3228	245.2166

Contd.

Table 5.1 contd.

POPULATION - III
 $(\rho = 0.6440, \frac{\sigma_x}{\sigma_y} = 0.2719)$

Sample size	MSE(\bar{y}_R)	MSE(\bar{y}_{reg})
10	370.4396	440.4602
15	284.9020	354.1204
20	197.7820	199.1085
25	152.6293	157.4801
30	151.6943	147.8054

POPULATION - IV
 $(\rho = 0.9130, \frac{\sigma_x}{\sigma_y} = 0.2697)$

Sample size	MSE(\bar{y}_R)	MSE(\bar{y}_{reg})
10	233.9732	193.0322
15	147.1425	88.4824
20	103.8812	55.3567
25	83.1825	45.2361
30	67.0958	32.8457

TABLE 5-2

PERCENTAGE INCREASE IN CUPRINO TO PLATO
BASED ON FAIR SQUARE MEASURE.

FORM No.	ρ	Increase in percentage for various sample sizes				
		10	15	20	25	30
I	0.2902	99.23	100.00	117.27	119.82	116.63
II	0.4553	98.43	99.00	98.73	97.51	100.43
III	0.6640	88.10	80.45	94.31	90.52	102.59
IV	0.9020	153.88	105.20	107.45	102.53	104.29

$$\uparrow \theta = \frac{\ln(\frac{P_1}{P_0})}{\ln(\frac{T_1}{T_0})} \times 100$$

We observe that the variances of both the methods decreases with the increase in sample size. However, the decrease in variance of regression with increase in sample size is much faster than that for the ratio method of estimation.

It is seen from the table 5.2 that efficiencies are just above 100% in the case of the first population except for the sample size of 10 units out of 200. In populations II and III where the β values are 0.4585 and 0.6440 respectively, the ratio estimator seems to be more efficient than regression estimator except for fairly large sample size of 50 units out of 200. In population-IV we see that regression estimator is superior than ratio estimator. The gain in efficiency is of the order 50 to 100%. As in the previous case efficiency does not shows any sort of trend with the sample size.

CHAPTER - VI
.....

STUDY ON THE FREQUENCY DISTRIBUTION OF THE ESTIMATES
.....
OBTAINED BY SAMPLE MEAN, RATIO AND REGRESSION METHODS
.....
OF ESTIMATION
.....

In this chapter we are mainly interested in studying the frequency distribution of the estimates of the population mean obtained by the three methods of estimation viz., sample mean, ratio and regression. Studies have also been made to examine how far the approximate formula of variance of ratio and regression estimates are accurate for computing confidence interval of a given size. In addition to this studies have also been made to compare the above said three methods of estimation on the basis of extent of concentration of the estimate around the population mean.

For this purpose from the populations described in chapter - II, 200 independent samples were drawn for varying sample sizes 10, 15, 20, 25 & 30 units. From the above 200 samples, estimates of the population mean, and the four moments were computed for the sake of studying the frequency distribution of these estimates in relation with the true value. The generated populations as already mentioned follow approximately normal distribution. Hence we expect the estimates obtained from the 200 samples should also behave approximately normal distribution. This behaviour can be well studied by the measure of

kurtosis and Charlier measure of skewness and these are given by

$$\text{Kurtosis } (\beta_2) = \beta_2 - 3$$

$$\text{Charlier measure of skewness } (\kappa_1) = \left\{ \frac{\mu_3}{\mu_2} - \frac{\mu_4}{\mu_2^2} \right\}$$

where $\beta_1 = \mu_3/\mu_2^{3/2}$ $\beta_2 = \mu_4/\mu_2^2$

μ_1, μ_2, μ_3 & μ_4 are the central moments of the different estimates about their respective mean values.

For a normal distribution the above mentioned two measures should be zero. In addition to the above two measures the Pearson Kappa (κ) should also be zero. The expression for the κ , is given by

$$\kappa_1 = \frac{\beta_1(\beta_2 + 3)^2}{4(2\beta_2 - 3\beta_1 - 6)(4\beta_2 - 3\beta_1)}$$

These measures have been computed for each of the frequency distribution underlying three procedures of estimation for varying sample sizes taken from each of the four populations under study. The figures are given in Table 6.1.

In practical situations we cannot obtain exact values which we require. From the Table 6.1 it is seen that the value of three measures are around zero. Hence we can say that the estimates obtained follow approximately normal distribution.

TABLE 61.

CHARTED MEASURES OF SKINNESS, MEASURES OF KURTOSIS AND THE MEASURE KAPP_A
FOR VARYING SAMPLE SIZES FOR THE 3 METHODS OF ESTIMATION FOR THE BOVINE POPULATION
POPULATION I (000-2902)

SAMPLE SIZE	S.E. S.E.	\bar{Y}_1^* Ratio	\bar{Y}_2 Ratio	\bar{Y}_3 Ratio	Kappa Ratio	Kappa Ratio	Kappa Ratio		
10	-0.0193	-0.0093	0.0203	-0.3709	-0.2709	0.1602	-0.0326	-0.0210	0.0156
15	-0.0093	-0.0723	-0.0513	0.0228	-0.5611	-0.3522	-0.2835	-0.0163	-0.0708
20	-0.0559	-0.0260	-0.0404	0.5225	-0.1703	-0.3207	-0.0079	-0.0836	-0.0079
25	-0.0141	-0.0293	-0.0053	-0.0332	-0.1716	-0.0096	-0.0073	-0.1162	-0.2272
30	-0.0261	-0.0203	-0.0112	-0.2722	-0.2253	-0.3231	-0.0015	-0.0011	-0.0081

TABLE 61 (Continued)

POPULATION XI ($\rho = 0.453$)

SAMPLE SIZE	\bar{S}_M	γ_1^* ratio	Mean	SD ₁	γ_2^* ratio	Mean	SD ₂	R ratio	Mean	SD _R
10	-0.0163	0.0203	0.3128	-0.1231	0.5010	0.3392	-0.0383	-0.0382	-0.0753	
15	0.0392	0.0122	0.0107	0.3762	0.1531	0.1631	0.1547	0.3716	0.4164	
20	-0.0021	-0.0169	-0.0170	-0.0360	-0.1494	-0.1232	-0.2015	-0.0358	-0.0256	
25	0.0252	0.0121	0.0324	0.2037	0.7265	0.5221	0.0220	0.1619	0.2497	
30	0.0523	0.0237	0.0515	0.7510	0.8357	0.6433	0.0337	0.6769	0.6557	

TABLE 6.1 (Continued)

PENETRATION III ($\delta = 0.640$)

SAMPLE SIZE	B.H.	Ratio	B.H.	Ratio	B.H.	Ratio	B.H.	Ratio	B.H.	Ratio
10	-0.0279	0.0053	0.0346	-0.1529	0.0819	0.4570	-0.1232	0.4600	0.0035	
15	-0.0510	0.0638	0.0770	-0.3935	-0.3617	0.6165	-0.0113	-0.0076	0.0003	
20	-0.0341	0.0231	0.0169	-0.2250	0.1837	0.1437	-0.0535	0.0162	0.1159	
25	0.0666	-0.0532	-0.0501	0.5239	-0.3460	-0.4259	0.0013	-0.0217	-0.0151	
30	0.0070	-0.0114	-0.0235	0.0564	-0.0763	-0.1603	0.0003	-0.3763	-0.0221	

TABLE 5.1 (Continued)

REGRESSION II ($\mu = 0.3120$)

NUMBER ITEM	R.E.	BETAS	R.E.	BETAS	R.E.	BETAS	R.E.	BETAS	R.E.	BETAS
10	-0.0135	-0.0287	0.0038	-0.0121	-0.0247	0.0216	-0.0232	-0.0753	0.0250	
15	0.0373	0.0206	0.0557	0.0385	0.0193	0.0600	0.0287	0.0369	0.0203	
20	-0.0176	-0.0175	0.0126	-0.0123	0.0153	0.0103	-0.0137	-0.2773	0.0000	
25	0.0028	-0.0267	0.0042	0.0029	-0.1930	0.0000	0.0030	-0.0146	0.0027	
30	0.0029	-0.0168	-0.0273	-0.0203	-0.0324	-0.2167	-0.0206	-0.0144	-0.0026	

6.2 To examine how far the approximate formulae of variance of ratio and regression estimates are accurate for computing confidence interval of a given size

The estimate of the variance enables the statistician to compute confidence interval of the population parameter (mean or total). The confidence interval for the population mean \bar{Y}_R or \bar{Y}_{reg} for a given size say α is

$$\bar{Y} \pm t_{n-1, \alpha/2} (s.d) \dots \dots \dots \dots \dots \dots \dots \quad (6.1)$$

where $t_{n-1, \alpha/2}$ is the t-value for $(n-1)$ df at α^{th} level and s.d is the standard deviation of the mean.

By the confidence interval of size α , we mean that the computed interval will cover the unknown population mean \bar{Y}_H with probability $(1-\alpha)$ where $(1-\alpha)$ is known as the confidence co-efficient. In otherwords, it implies that, in case we select 100 repeated samples of a given size, and in each case compute confidence interval of size α as given by (6.1), then out of 100 samples only 100 intervals will not cover the population mean while the rest $100(1-\alpha)$ will contain the population mean. That is to say if we use exact formulae for the s.d we can expect the probability that the computed interval will cover the population mean to be $(1-\alpha)$.

In the absence of exact formulae of s.d in the case of ratio and regression estimators as also both \bar{Y}_R and \bar{Y}_{reg} not following exactly normally distributed it is quite likely that the probability that the computed interval will contain the population mean may deviate from $(1-\alpha)$.

TABLE 6-2. ESTIMATED PROBABILITY THAT THE ESTIMATED INCOME
EXCEEDS THE POPULATION MEAN.

ESTIMATE 10.

$(\text{Var}_Y) = 0.0001$				$(\text{Var}_Y) = 0.0003$			
β	Sale	Ratio	Prob.	β	Sale	Ratio	Prob.
0.25	73.0	62.5	0.0	24.5	81.0	62.5	0.0
0.50	55.0	62.5	0.0	12.0	82.5	62.5	0.0
0.60	47.0	62.5	0.0	9.5	83.0	62.5	0.0
0.75	39.0	62.5	0.0	3.0	83.0	62.5	0.0
<u>ESTIMATE 11</u>				<u>ESTIMATE 12</u>			
$(\text{Var}_Y) = 0.00025$	$(\text{Var}_Y) = 0.00035$	$(\text{Var}_Y) = 0.00025$	$(\text{Var}_Y) = 0.00035$	β	Sale	Ratio	Prob.
0.25	74.0	56.5	27.0	25.0	82.0	62.5	0.0
0.50	56.0	56.5	36.0	12.5	82.5	62.5	0.0
0.60	48.0	56.5	45.5	9.0	83.0	62.5	0.0
0.75	40.0	56.5	24.5	3.5	83.5	62.5	0.0
<u>ESTIMATE 13</u>				<u>ESTIMATE 14</u>			
$(\text{Var}_Y) = 0.0001$	$(\text{Var}_Y) = 0.0003$	$(\text{Var}_Y) = 0.0001$	$(\text{Var}_Y) = 0.0003$	β	Sale	Ratio	Prob.
0.25	55.5	50.5	21.5	22.0	83.5	62.5	0.0
0.50	37.5	50.5	33.0	12.0	84.0	62.5	0.0
0.60	30.5	50.5	37.5	9.5	84.5	62.5	0.0
0.75	23.5	50.5	24.5	3.0	85.0	62.5	0.0

TABLE NO. 22

$(\log_e) = 67.000$

$(\log_e) = 94.000$

ρ	Ball	Bottle	Beaker	Ball	Bottle	Beaker
0.20	71.5	82.5	65.0	86.0	82.5	81.5
0.45	63.0	63.5	60.0	62.0	67.0	67.0
0.64	62.0	70.0	63.0	64.0	71.0	64.0
0.91	63.0	71.0	72.0	74.0	94.0	94.0

TABLE NO. 23

$(\log_e) = 67.275$

$(\log_e) = 94.250$

0.20	73.5	81.5	61.5	85.0	81.5	80.5
0.45	67.5	77.5	66.0	75.0	82.0	87.0
0.64	65.0	65.0	62.0	65.0	71.0	64.0
0.91	63.5	72.0	70.5	72.0	94.0	94.0

For this purpose from the 200 sample mean, ratio and regression estimates, confidence intervals of two different sizes are computed for all the 200 estimates. The sizes α_1 & α_2 of the confidence intervals are so chosen that $t_{n-1, \alpha_1/2}$ & $t_{n-1, \alpha_2/2}$ takes rounded values 1 & 2 respectively, for the sake simplicity.

Table 6.2 gives the percentage probabilities that the computed interval contains the population mean. It will be seen from the table that the confidence intervals based on ratio estimators by using approximate formulae of its standard deviation are quite satisfactory. The probability of the confidence interval of given size increases with the value of f . Thus for all sample sizes for the value of $f=0.29$ the probability is less than the confidence co-efficient. While for $f=0.91$ the realised probabilities consistently(for all sample sizes) higher than the co-efficient of the confidence interval. It is the same with the case of regression estimator, also.

6.2 Comparision between sample mean, ratio and regression methods of estimation on the basis of extent of concentration of the estimate around the population parameter.

While comparing the three estimators viz., sample mean ratio and regression on the basis of variance criterion (as estimated from large number of repeated samples) it was seen that sample mean was superior than either ratio or regression method of estimation for the populations where f is of order 0.3 or less. Ratio estimates has got an edge over regression estimators when f ranges from 0.3 to 0.9. However for the population where f is of order 0.9 or more regression is found to be definitely

TABLE 4.2

PERCENTAGE OF THE SAMPLE ESTIMATES (OUT OF 200) THAT ARE IN ACCEPTABLE RANGE
ACROSS PORTFOLIO ALLOCATIONS (CONTINUED).

PORTFOLIO ALLOCATION 1.2%

PORTFOLIO	Method	5 Percent			10 Percent			20 Percent		
		10	20	30	10	20	30	10	20	30
Portfolio 1	Sample	15.5	17.9	23.5	20.5	20.3	23.5	19.5	22.5	27.0
	Ratio	10.0	16.0	24.5	23.5	20.5	27.5	18.0	21.0	26.0
	Range	15.0	17.0	22.0	20.0	23.0	23.0	18.0	21.0	26.5
Portfolio 2	Sample	16.5	20.0	27.5	25.0	27.0	27.0	20.0	22.0	22.5
	Ratio	15.0	21.0	23.0	20.5	20.0	22.0	18.0	20.0	20.0
	Range	15.0	19.5	22.5	20.0	20.5	21.5	18.0	20.0	20.5
Portfolio 3	Sample	14.0	25.0	24.5	22.5	45.0	40.0	50.0	55.0	77.0
	Ratio	24.0	37.0	62.0	42.0	52.0	62.0	72.0	82.0	87.0
	Range	21.0	36.5	53.0	42.5	52.5	52.5	52.5	52.5	67.5
Portfolio 4	Sample	23.0	25.0	40.0	23.0	40.0	43.5	49.5	49.5	59.0
	Ratio	23.5	33.0	34.5	47.0	60.5	72.0	72.0	93.5	94.0
	Range	23.0	31.0	63.5	60.0	65.0	65.0	62.0	63.5	66.0

Differences within brackets denotes correlations coefficients.

superior to both sample mean and regression.

In order to judge the relative merits of those estimates, a different criterion, based on the extent of concentration of estimated values around the true value of the population mean has been adopted. In other words, likelihoods of an estimate lying within different ranges of concentration around the population mean have been worked out for the three estimators. The different ranges of concentration considered are $\pm 2.5\%$, $\pm 5\%$, and $\pm 10\%$ around the population mean.

Accordingly, the estimated values of population mean obtained on the basis of three estimators, have been classified into different ranges of concentration around the population mean. The figures given in Table 6.3 gives the percentage likelihood of an estimate lying within different ranges of concentration around the population mean and for sample sizes 10, 20 and 30 units selected from the generated populations.

To judge the relative performance of the estimators on the basis of this ground, the following methodology is adopted.

- (a) If an estimator provides estimate such that likelihood of the estimate lying in different ranges of concentration around the population parameter(mean) is consistently larger than the one provided by another estimator, then the former estimator will be regarded as superior to the latter.
- (b) If an estimator provides estimate such that likelihood of the estimate lying in smaller ranges of concentration around the population parameter(mean), say $\pm 2.5\%$, is larger than

the one provided by the other estimator, but for larger ranges, say +5%, $\pm 10\%$. the other estimator has a greater likelihood of lying in these ranges, then an estimator which has greater values of likelihood for smaller ranges, is said to have an edge over the other estimator.

On the basis of this, it can be seen for the population-I where the f value is 0.2902 that even for a sample size of 10 units , sample mean is better than both ratio and regression. Also for sample size 30 units likelihood of sample mean estimate lying around the smaller ranges is still higher showing once again that sample mean is much superior to both ratio and regression estimator. However for the other populations where the f values lies between 0.3 and 0.9 it can be seen from the Table 6.3 that ratio has got an edge over regression estimator. Also due to extra amount of computational labour involved in regression estimator, ratio may be suggested in preference to regression estimator where f lies between 0.3 and 0.9. But for the populations $f > 0.9$, regression estimator is found to be most superior method of estimation.

SUMMARY AND DISCUSSION

Generally, sample survey statistician is faced with the problem as to how he should make the best use of available information on some auxiliary variable which is highly correlated with the variable under study. The types of population which we come across in surveys for estimating the production or extent of cultivation of agricultural commodities generally follow a linear regression relationship on auxiliary character X . Hence in this thesis the population was assumed to be coming from a bivariate normal distribution. Four populations were generated for four different values of β ranging between 0.3 and 0.9 for given mean and given standard deviations. The X -observations were treated as auxiliary variates and y -observations were treated as study variate.

The present study was aimed at finding out the suitable method of estimation when auxiliary variates are available. The study was carried out through Monte-carlo methods which envisage the simulation of populations which have the given pattern of relationship between the study and auxiliary variates for some reasonably fixed values of parameters. From each of the generated population consisting of 200 units, corresponding to the four varying values of β , 200 independent sample (four sets of 50 each) for varying sample sizes were selected. Each of these samples provided estimates of population total by using three different methods of estimation along with the corresponding estimates of the variances. One estimate

of the bias in the estimate of the population mean estimated through ratio and regression methods of estimation is provided by taking the difference between the average value of the estimates taken over 200 samples and the true value of the population mean. Similarly the mean square of 200 sample values of the estimates of the population mean provides estimate of the true value of the variance of an estimate of a given parameter. The present study has revealed the following results,

(i) So far as the amount of bias in estimating the population mean through ratio and regression estimator is concerned, the study has revealed that its value is appreciably small being less than 2% of the mean for all the population.

(ii) Comparing the performances of the ratio and regression methods of estimation with the sample mean for estimating the population mean on the basis of mean square errors it has been established that when f is of the order 0.3 or less then sample mean is preferable over the ratio and regression method of estimation since the true variance as estimated by the mean square between the estimates of the population mean provided by 200 independent samples, is consistently smaller in the case of sample mean than in the case of either ratio or regression method of estimation. For the population - I even when the χ^2 criterion $f > \frac{4}{X} / C_Y$ hold good the study has revealed that sample mean is superior than both the ratio & regression method of estimation.

(iii) Comparing ratio and regression method of estimation it is found that the regression is superior than ratio method of estimation when β is of the order 0.9 or more . For the intermediate values of β ranging from 0.3 to 0.9 ratio estimator has got an edge over the regression estimator for such populations as discussed in this thesis. The second population which is slightly skewed has not violated the expected results.

(iv) Study on likelihood of making a wrong decision when comparision of ratio versus sample mean estimator is based on a single sample has revealed that for the second and third populations the probability of committing a wrong dicision (p) and the sampling fraction (f) is best related by the exponential function of the form $p = e^{-bf}$ for $a, b > 0$. The value of R^2 , the multiple correlation co-efficient varied between 0.81 and 0.87.

(v) Effort was also made to study the frequency distributions followed by estimates of the mean obtained by the three methods of estimation. It has been found that the estimates so obtained follows approximately normal distribution although there was a slight skewness in the second population this has not violated the behaviour of the estimates following a normol distribution.

(vi) In studying the behaviour of the approximate formulae using the confidence intervals it is observed that these methods will underestimate the probability to a very little extent in the absence of exact formulae. This clearly shows that the estimate of the true variance are more or less accurate.

(vii) To examine further the comparison of the above said three methods of estimation a deeper study was made on the basis of the extent of concentration of individual sample mean, ratio and regression estimates provided by 200 samples around the population mean. On the basis of this study it has been established that with small sample sizes, the use of regression estimates is warranted only if the correlation between the study and auxiliary character (f) is assumed to be of the order of 0.9 and above. The use of ratio estimate is warranted for f in between 0.3 and 0.9. Perhaps for large sample sizes of 30 units and above the use of regression estimator might be warranted even in case f is slightly less than 0.9. The use of sample mean estimator is warranted when f is of order 0.3 or less.

A P P E N D I X
.....

Sl. No.	POPULATION.-I		POPN.-II		POPN.-III			POPN.-IV	
	X	Y	X	Y	X	X	Y	X	Y
1.	36.93	106.12	50.91	184.52	49.99	237.36	56.23	278.31	
2.	40.05	202.73	49.40	179.75	49.03	226.57	70.94	384.22	
3.	37.88	13.30	52.43	202.20	49.46	170.15	51.12	168.25	
4.	77.99	274.20	56.92	105.47	46.04	47.93	32.28	85.18	
5.	21.61	168.90	58.91	242.32	46.48	322.17	48.97	203.66	
6.	73.77	248.34	43.32	100.47	55.55	215.90	50.44	76.86	
7.	39.23	90.09	62.60	191.57	55.38	120.23	53.65	305.38	
8.	32.67	231.64	42.84	389.91	37.69	165.30	38.25	78.31	
9.	59.18	468.89	61.54	335.25	58.71	285.49	54.61	257.74	
10.	27.63	81.66	55.09	210.69	42.64	144.40	71.73	413.50	
11.	58.26	295.51	68.21	297.61	52.49	210.14	53.25	180.17	
12.	24.85	80.48	51.42	310.61	68.54	252.41	66.03	343.85	
13.	54.04	175.59	50.52	65.13	50.70	189.17	58.35	297.92	
14.	48.00	94.07	43.73	176.84	40.63	229.47	48.76	142.53	
15.	46.16	168.62	57.58	237.57	49.36	290.14	58.83	273.50	
16.	52.14	212.48	61.49	200.14	43.23	157.74	56.04	296.27	
17.	73.44	274.44	49.80	228.75	56.41	297.77	70.83	347.94	
18.	65.69	341.04	28.09	4.37	47.18	107.84	47.77	283.15	
19.	55.20	196.66	53.38	149.93	52.45	262.66	59.18	267.45	
20.	30.18	15.32	42.96	274.97	56.21	298.85	29.02	96.98	
21.	70.10	372.07	64.90	286.57	39.50	246.95	56.18	273.24	
22.	32.40	232.64	64.73	348.27	51.04	306.08	50.13	146.68	
23.	51.57	322.19	71.81	405.21	49.98	145.00	48.68	195.48	
24.	34.41	61.40	33.57	198.10	43.44	6.37	71.96	322.16	
25.	57.62	340.31	48.79	246.49	57.80	238.97	66.29	327.04	
26.	48.91	119.64	51.22	134.48	43.14	152.44	32.63	92.39	
27.	54.79	119.30	63.64	332.92	63.11	378.98	38.93	104.56	
28.	36.54	271.21	66.13	157.34	72.76	286.52	48.61	158.12	
29.	41.88	284.99	67.03	344.96	61.27	220.73	53.05	177.58	
30.	54.93	274.87	37.57	10.85	34.58	71.21	38.15	101.07	
31.	61.84	295.20	47.86	156.39	54.63	221.35	67.54	308.05	
32.	47.17	91.21	45.54	189.86	49.22	242.95	44.74	119.85	
33.	65.62	202.16	32.21	237.07	67.88	271.19	50.62	55.93	
34.	59.76	195.52	39.29	230.29	57.16	369.16	45.81	169.18	

35.	51.20	292.79	61.14	248.64	74.57	371.78	42.37	165.95
36.	52.26	54.81	54.90	224.57	40.66	20.17	36.51	69.95
37.	57.16	201.82	65.54	166.16	57.22	235.14	37.04	103.68
38.	29.90	68.29	39.16	139.64	42.42	165.04	42.12	198.02
39.	38.52	173.58	44.87	127.81	58.82	142.24	65.18	309.12
40.	71.51	223.15	39.78	118.69	32.89	194.03	60.86	285.96
41.	46.14	360.13	98.12	192.20	49.90	247.46	50.99	178.57
42.	49.88	257.32	29.20	119.55	64.66	264.98	55.60	268.92
43.	35.44	289.06	37.30	65.34	50.99	259.63	43.56	181.83
44.	40.31	291.14	26.87	166.62	68.89	331.90	42.69	220.07
45.	64.40	353.20	40.07	95.95	48.16	258.99	48.54	163.22
46.	69.27	358.54	63.35	319.77	53.34	275.31	59.33	221.95
47.	45.90	151.43	37.50	149.80	51.49	309.73	59.65	252.63
48.	38.35	15.33	38.29	87.40	36.66	84.53	44.97	207.21
49.	76.06	479.72	40.77	282.11	62.39	189.40	44.73	189.23
50.	67.34	279.98	52.19	162.45	26.79	111.52	25.47	35.67
51.	40.19	233.22	24.66	131.39	53.15	221.12	71.57	247.32
52.	80.15	263.53	42.95	236.70	45.48	121.55	27.28	64.59
53.	42.93	184.66	77.80	318.40	32.71	97.97	49.57	149.95
54.	47.35	109.86	46.10	172.37	28.88	106.21	24.82	58.24
55.	37.38	126.64	40.73	296.92	37.43	190.55	49.37	165.27
56.	51.26	268.04	59.95	357.56	40.92	198.24	29.70	50.83
57.	46.58	294.33	48.60	245.65	52.49	195.01	60.54	203.71
58.	31.12	166.84	22.96	33.82	22.93	63.09	61.36	338.41
59.	52.83	141.68	27.11	85.02	60.35	84.74	45.70	198.34
60.	46.25	272.97	52.38	91.71	59.05	186.29	36.51	59.15
61.	45.81	291.98	63.23	294.23	35.63	228.58	57.30	246.13
62.	58.37	194.65	42.20	214.97	47.63	108.60	66.77	339.20
63.	77.57	331.82	38.99	309.80	59.11	283.05	31.11	22.52
64.	54.47	11.42	52.17	407.37	50.25	148.34	41.97	175.63
65.	59.79	139.71	67.59	391.65	54.08	142.86	70.41	343.40
66.	54.62	184.42	57.35	199.90	54.27	175.05	44.68	149.61
67.	43.81	325.51	49.20	144.98	56.91	319.01	57.03	314.57
68.	48.20	212.93	66.75	310.85	53.26	271.99	45.62	207.26
69.	50.32	167.55	50.51	175.10	45.60	80.90	63.85	341.92
70.	42.64	247.06	72.11	323.90	60.44	328.37	43.48	109.32
71.	27.50	123.61	57.50	152.19	52.69	246.56	47.49	168.85
72.	24.08	228.47	64.28	209.36	75.57	276.69	43.89	126.24

73.	58.39	493.52	46.04	204.94	53.95	229.80	36.94	137.44
74.	26.71	137.51	57.41	303.44	43.07	73.02	34.20	59.36
75.	53.62	318.55	43.10	72.02	54.37	260.37	32.94	117.18
76.	37.37	253.83	54.16	210.59	66.70	277.37	35.75	113.40
77.	60.56	204.85	39.53	145.18	25.49	115.70	61.84	348.24
78.	59.11	325.78	43.82	199.98	69.89	329.67	21.45	11.02
79.	50.12	222.68	35.86	108.56	34.06	54.40	38.21	230.01
80.	44.41	203.47	29.63	43.42	74.97	201.46	43.10	180.10
81.	50.15	238.68	44.77	72.62	33.32	182.04	43.52	135.37
82.	53.18	230.52	53.47	207.68	39.75	201.54	38.82	113.06
83.	69.48	214.38	33.95	102.01	44.63	94.35	51.47	197.39
84.	52.92	92.02	62.76	326.31	74.54	460.24	36.80	71.01
85.	38.78	48.74	57.71	152.01	55.01	237.77	55.72	186.94
86.	44.82	274.22	28.41	9.01	42.41	156.77	65.49	274.84
87.	52.28	204.85	30.41	89.25	75.43	406.60	49.36	240.57
88.	56.50	216.34	57.04	230.63	64.74	305.40	54.02	252.85
89.	30.11	230.44	48.98	84.54	46.11	202.84	40.65	156.21
90.	52.87	274.09	55.70	189.56	50.93	271.79	69.52	333.22
91.	26.77	20.13	50.78	314.23	55.98	256.17	54.61	245.68
92.	49.12	153.71	41.89	105.59	62.94	196.95	44.39	180.73
93.	50.02	320.98	53.66	245.76	46.27	226.14	39.80	134.87
94.	43.21	206.62	38.54	315.23	74.01	445.90	71.60	382.75
95.	54.60	291.80	45.30	185.47	54.34	321.59	51.39	197.78
96.	56.57	188.66	43.19	108.72	45.94	267.57	42.74	144.11
97.	39.25	353.67	45.57	97.17	40.41	143.03	44.95	174.50
98.	68.62	262.01	45.21	196.37	49.94	164.79	50.07	244.30
99.	54.24	239.60	59.70	141.49	48.85	229.88	56.23	182.19
100.	36.04	280.02	55.13	150.15	55.52	239.92	40.35	77.48
101.	39.56	362.65	39.24	99.34	60.86	160.40	38.22	110.47
102.	54.14	238.67	54.08	129.42	12.67	40.51	51.59	246.66
103.	55.71	95.48	55.99	406.65	50.31	224.73	36.11	113.32
104.	61.93	315.93	35.85	282.25	38.79	52.71	43.27	235.03
105.	49.30	228.88	53.61	189.81	37.94	415.42	27.57	41.36
106.	30.37	394.38	32.22	83.22	59.36	166.28	43.72	75.71
107.	60.86	231.12	35.84	121.22	400.11	194.51	51.98	181.80
108.	30.24	192.13	7.80	67.66	60.33	247.48	67.29	312.28
109.	64.09	183.46	46.39	182.01	34.59	113.78	28.94	100.61
110.	54.24	182.42	45.90	253.68	32.38	198.60	37.98	112.61

111.	64.26	314.83	47.90	228.24	57.78	193.57	60.26	254.21
112.	55.72	316.57	33.91	79.19	48.19	158.88	40.82	153.87
113.	47.82	66.08	39.91	57.62	57.16	187.65	60.71	251.13
114.	49.10	262.25	44.32	235.19	44.43	204.62	57.21	266.77
115.	43.29	214.47	54.80	194.01	31.03	158.53	77.01	378.96
116.	50.06	222.64	35.49	238.32	49.29	130.48	52.72	259.53
117.	39.50	27.90	49.17	112.54	41.05	82.43	53.03	238.76
118.	67.13	224.64	50.96	278.66	51.07	163.07	60.80	270.84
119.	53.37	211.20	58.14	344.22	49.58	139.90	47.11	155.39
120.	77.86	322.70	49.19	164.51	63.27	284.46	50.55	175.46
121.	48.77	196.79	53.84	285.28	40.99	149.59	51.87	257.26
122.	23.71	270.74	52.35	232.87	44.44	157.02	82.17	467.27
123.	55.27	205.58	28.69	135.62	59.95	174.67	22.91	9.52
124.	69.23	113.37	36.72	14.07	48.36	242.20	37.58	160.13
125.	47.44	229.29	52.14	88.15	59.80	331.93	34.72	169.64
126.	46.91	55.66	56.83	220.92	57.98	179.48	50.30	196.18
127.	41.68	250.40	48.52	272.18	34.81	48.54	43.57	166.23
128.	42.83	84.58	44.00	108.88	52.39	128.14	70.54	379.41
129.	66.69	225.67	59.16	307.41	64.57	199.59	32.91	90.40
130.	57.78	275.28	49.49	220.85	44.57	105.20	94.20	514.58
131.	44.66	217.91	57.18	329.58	62.65	233.71	54.96	274.60
132.	38.12	229.35	50.90	276.31	52.38	154.85	83.13	305.57
133.	37.50	15.33	30.34	318.46	39.70	186.48	55.62	224.02
134.	53.40	316.61	35.78	369.56	33.167	129.79	37.94	97.48
135.	56.86	232.77	38.70	202.91	59.80	201.46	40.37	192.60
136.	43.86	146.57	71.62	285.88	38.86	29.21	50.58	211.15
137.	59.97	257.51	35.62	157.61	45.64	234.17	52.27	285.94
138.	79.62	255.30	55.68	336.79	55.91	237.26	58.22	223.69
139.	52.18	238.94	47.24	215.73	34.52	53.08	24.58	30.06
140.	41.35	291.50	49.71	245.11	55.69	274.33	75.57	397.58
141.	62.68	202.19	44.14	272.26	63.08	260.59	67.72	292.30
142.	52.22	367.67	55.95	239.52	65.14	325.99	50.75	263.99
143.	41.39	103.31	35.75	181.72	42.76	66.86	62.18	212.82
144.	51.41	320.34	26.08	181.43	38.86	202.25	30.29	18.34
145.	51.04	309.81	60.92	396.28	92.74	446.77	56.74	267.45
146.	43.23	37.24	44.58	271.16	18.44	42.71	59.47	283.86
147.	23.51	175.76	61.97	178.95	58.10	230.98	47.97	178.61
148.	48.62	220.70	67.63	228.44	48.87	224.14	47.93	177.78

149.	78.70	155.69	61.86	150.72	67.46	233.35	49.53	184.55
150.	44.17	109.26	38.53	259.85	54.26	95.33	47.24	209.82
151.	62.68	203.70	40.45	77.78	46.75	237.66	67.82	354.59
152.	68.24	230.18	61.30	267.00	49.51	146.12	68.48	324.49
153.	50.69	164.02	53.65	156.91	65.09	300.29	63.08	312.53
154.	63.97	166.71	57.81	165.14	40.42	58.08	41.88	91.14
155.	58.06	239.01	51.30	165.34	55.08	363.62	57.94	284.11
156.	49.49	352.12	53.82	158.98	63.34	236.41	53.45	216.65
157.	44.86	108.03	48.45	244.82	54.40	236.60	42.86	102.32
158.	59.89	125.89	47.16	162.07	50.19	117.38	60.99	359.38
159.	39.42	84.60	62.14	250.99	40.88	154.07	56.81	226.07
160.	56.09	168.76	68.77	338.02	61.36	186.11	855.56	75.31
161.	38.99	249.66	29.69	29.77	63.44	320.89	73.40	402.17
162.	78.65	221.46	65.03	353.65	61.00	246.78	57.90	172.13
163.	53.72	121.71	57.55	434.03	44.64	263.45	42.73	102.26
164.	45.55	310.02	34.41	177.35	65.12	239.56	69.28	366.57
165.	37.51	345.40	50.73	62.55	40.55	72.86	59.79	252.77
166.	41.28	280.03	57.57	128.81	49.54	206.23	65.68	302.07
167.	59.55	408.38	44.26	6.97	54.48	257.46	46.23	222.13
168.	61.96	46.88	35.99	714.92	47.65	247.64	36.69	103.88
169.	37.24	143.20	58.88	179.18	52.16	107.38	52.53	194.96
170.	37.08	151.38	41.58	86.33	44.15	118.93	58.82	278.33
171.	61.25	264.59	67.58	316.98	36.80	148.67	95.26	122.90
172.	52.96	216.19	45.71	508.89	40.99	149.59	48.10	227.36
173.	52.90	195.86	38.89	207.65	33.90	93.12	37.79	114.50
174.	43.47	266.13	33.11	319.77	44.33	129.72	59.46	350.22
175.	52.71	76.80	37.19	88.91	56.25	200.94	40.69	136.79
176.	38.79	285.93	38.90	154.41	59.23	473.22	41.01	209.58
177.	41.70	268.99	50.58	252.39	64.58	429.50	50.44	193.31
178.	45.64	351.74	69.02	351.39	41.97	143.04	62.09	282.89
179.	58.87	271.79	51.24	192.54	45.90	152.44	43.48	163.24
180.	41.25	173.58	53.78	244.23	36.36	162.70	52.85	210.63
181.	61.04	183.30	17.21	168.17	66.07	279.83	36.80	127.16
182.	39.60	223.84	29.53	6.16	31.03	202.84	30.23	75.14
183.	49.11	324.10	67.80	351.16	45.95	113.77	53.97	228.59
184.	46.98	191.40	47.37	210.27	43.27	155.11	59.72	361.26
185.	35.04	192.48	40.52	53.82	26.47	142.87	48.29	139.68
186.	62.32	245.21	50.36	207.86	52.37	289.92	54.42	178.64

187.	59.10	165.85	62.21	333.93	41.25	251.09	44.51	209.01
188.	27.30	173.72	33.34	235.10	40.07	196.09	38.59	137.40
189.	38.92	105.86	33.42	157.28	47.26	126.32	55.89	225.22
190.	39.82	322.98	65.03	132.78	46.48	186.54	57.01	159.98
191.	47.49	179.30	64.95	306.09	54.93	117.42	39.73	148.69
192.	52.31	213.02	56.40	181.64	43.54	157.77	64.32	267.51
193.	60.14	190.84	74.03	275.01	46.21	127.55	44.60	158.67
194.	44.88	80.69	43.03	187.80	35.70	223.42	30.13	53.41
195.	51.04	286.45	51.79	309.51	61.91	208.61	48.88	249.78
196.	78.09	193.94	78.26	333.23	26.37	193.21	32.17	28.76
197.	64.10	144.73	52.86	168.01	43.56	280.77	71.56	395.88
198.	56.86	111.61	64.48	249.89	63.56	319.17	46.69	253.50
199.	56.42	202.39	46.96	156.10	54.61	111.61	56.53	182.90
200.	80.04	339.88	42.35	139.23	44.90	119.59	59.86	290.49

R E F E R E N C E S
.....

- COCHRAN W.G. (1972) "Sampling Techniques" Wiley Eastern Private Ltd. New Delhi.
- ELDERTON. W.P. (1927) "Frequency curves and correlation" Charles & Edwin Layton, London.
- GUPTA. V.K. (1973) "Some empirical studies of estⁿ. procedures in sample surveys on minor crops" Unpublished thesis for Diploma in Agricultural Statistics of IARS New Delhi.
- HAMMERSLEY. J.M & HANDSCOMB. D.C. (1965) "Monte-carlo methods" Methuen and co., Ltd., London.
- HARTLEY. H.O. & RAO. J.N.K. (1962) "Sampling with unequal prob's and WOR" Ann. Math. Stats. 33
- H.A. MEYER (1954) "Symposium on Monte-carlo methods" John Wiley & Sons. Inc. New York.
- Cramer. H. "Mathematical methods of Statistics" Asim publishing house.
- JOHNSTON. J (1960) "Econometric methods" J. McGraw-Hill Book company Inc. New York.
- KENDALL. M.G (1948) "The Advanced theory of Statistics Vol I Charles Griffin & Co. Ltd., London.
- RAO. J.N.K., LeNelle D. BEEGLE (1967) "A Monte-carlo study of some ratio estimators" Sankhya 29, 47-56.
- Sukhatme. P.V. & Sukhatme B.V. (1970) "Sampling theory of Surveys with application" Iowa State University Anne, IOWA, USA
- MANWANI. A.H (1973) "Theory of Estimation of multiple characters in sample survey" Unpublished thesis for Ph.D in Ag. St. I.A.R.S. New Delhi
- MAHAJAN. R.K (1972) "Certain Monte-carlo studies in sample surveys" Unpublished thesis for Diploma in Agricultural Statistics of IARS New Delhi.

.....