A CLASS OF RATIO COMPOSITE ESTIMATORS

Stephen M. Woodruff, Bureau of Labor Statistics

Certain ratio composite estimators are investigated and a general form for writing their variance is derived. The form of the variance permits relatively easy computation of optimal values for the parameters in the sampling scheme. The computations are carried through in detail for a special case of composite estimation using data from sampling on two occasions. In this special case, criteria for the use of composite estimation verses sampling on the latest occasion is derived. Some applications and extensions of these results are also discussed and an Empirical study using a computer simulation is presented.

Key Words: Composite estimator; ratio estimator; Simple Random Sampling

1. INTRODUCTION

Current survey estimates can often be considerably improved upon by the use of survey results from the preceeding time periods. The extra care devoted to survey design when data for more than one occasion is to be used will often be rewarded with pleasing reductions in mean square error. This article is a summary of some findings on the use of composite estimation on ratio estimators with ratio adjustments. A special case of sampling on 2 occasions is examined in detail and an extension of the "2 occasion" results are investigated for sampling on more than 2 occasions. An empirical study of these estimators is described in the last section.

The composite estimator that is studied herein is mentioned briefly in several references, notably Hansen, Hurwitz and Madow, Wolter, Cochran, Woodruff and Eckler, but, nowhere could I find any detailed exposition devoted explicitly to it. In particular, the optimization of the sampling design with respect to sample overlap as well as weighting is analyzed using a linearization procedure that yields surprisingly compact results.

2. THE ESTIMATOR AND SAMPLING SCHEME

It is desired to study the properties of composite estimators of the form:

$$\hat{R}_h = \alpha_h (r'_{h,h-1}/r'_{h-1,h})\hat{R}_{h-1} + (1-\alpha_h)r_h \text{ for } h = 3,4,5\ldots \quad (2.1)$$

where $\hat{R}_2 = \alpha_2 (r'_{2,1}/r'_{1,2})r_1 + (1-\alpha_2)r_2$

$0 \leq \alpha_i \leq 1$ for $i = 2,3,4\ldots\ldots\ldots\ldots$,

$r'_{i,j} = \bar{y}'_{i,j}/\bar{x}'_{i,j}$ and $r_i = \bar{y}_i/\bar{x}_i$,

$\bar{y}'_{i,j}$ is the sample mean for the Y characteristic at time i based on the overlaped sample between time i and time j ($|i - j| = 1$). $\bar{x}'_{ij}$ is analogously defined.

$\bar{y}_i$ ($\bar{x}_i$) is the sample mean for the Y(X) characteristic at time i based on the entire sample at time i.

Initially (at time $i = 1$) a simple random sample (SRS) of size n is selected and from these n units a SRS of size $\lambda_1 n$ ($0 < \lambda_1 < 1$) is selected for carryover to time $i = 2$.

At time $i = 2$ a sample of size n is constituted from the $\lambda_1 n$ units previously selected plus a new SRS of size $(1 - \lambda_1)n$ selected from the remaining units in the case of without replacement sampling. From these n units at time $i = 2$, a subsample of size $\lambda_2 n$ is selected for carryover to time $i = 3$. At time $i = 3$ a sample of size n is constructed from these $\lambda_2 n$ units plus a new sample of size $(1 - \lambda_2)n$ from the remaining units just as was done at the preceeding step. This process is continued until time h. It is assumed throughout that the population variance of X and Y is constant over time.

The quantity that I wish to estimate is $R_h = \bar{Y}_h/\bar{X}_h$ where $\bar{Y}_h$ is the population mean for the Y characteristic at time h and $\bar{X}_h$ the population mean for the X characteristic at time h.

3. THE VARIANCE OF THE ESTIMATOR

$\hat{R}_h$ is a function of $\bar{x}_i$, $\bar{y}_i$, $\bar{x}'_{i,j}$ and $\bar{y}'_{i,j}$ for i and j ranging between 1 and h. Expanding $\hat{R}_h$ in a Taylor Series about the corresponding population means and ignoring terms of order 2 and higher it is found, the variance of this linearized quantity, which is denoted by $\tilde{R}_h$, is of the form given in Theorem 1.

Theorem 1: Var $(\tilde{R}_h) = \text{trace}(\Sigma F)$
where $\Sigma = E(ZZ')$

and $Z = (Z_1, Z_2, \ldots Z_n)$ where $Z_i = (\bar{Y}_h/\bar{X}_h)((y_i/\bar{Y}_i) - (x_i/\bar{X}_i))$

y_i (x_i) is the Y value (X value) at time i of a single unit selected at random from the population (the selection probability is distributed uniformly over the entire population). The random vector, Z , is thus distributed as the outcome of the random selection of one unit from which $x_1\ x_2 \ldots x_h\ y_1\ y_2 \ldots y_h$ are observed.

F is an h x h matrix of constants that are themselves functions of the sampling design. That is, F is a function of $\lambda_1\ \lambda_2 \ldots \lambda_{h-1} \ldots \alpha_2\ \alpha_3 \ldots \alpha_h$ n and N the size of the universe.

Proof of Theorem 1.1: The linearization, $\tilde{R}_h$, of $\hat{R}_h$ reduces to constant terms plus terms of the form:

$L_k(\bar{Y}_h/\bar{X}_h)((\bar{y}_{ki}/\bar{Y}_i) - (\bar{x}_{ki}/\bar{X}_i))$ where $\bar{y}_{ki}$ & $\bar{x}_{ki}$

are sample means from a SRS at time i of some size, not necessarily n. L_k is a constant depending on the sampling design; $k = 1,2,\ldots,m$ where m is usually greater than h. This is done in section 4 for $\tilde{R}_2$ and by mathematical induction is easily intended to $\tilde{R}_h$ for $h > 2$.

The variances of the terms $L_k\ (\bar{Y}_h/\bar{X}_h)((\bar{y}_{ki}/\bar{Y}_i) - (\bar{x}_{ki}/\bar{X}_i))$ and the covariances between them are then expressed in terms of the variances and covariances of the components of Z to give:

$$\text{Var}(\tilde{R}_h) = \sum_{i=1}^{h} \sum_{j=1}^{h} f_{ij}\,\text{Cov}(Z_i\,Z_j) = \text{trace}(\Sigma F)$$

Q.E.D.

$F = F(\alpha_2\ \alpha_3 \ldots \alpha_h\ \lambda_1 \ldots \lambda_{h-1}\ n, N)$ is a matrix valued function of these design parameters. To select these design parameters so the Var $(\tilde{R}_h)$ is minimized it is necessary to solve the system:

$$\operatorname{trace}(\Sigma F\alpha_i) = 0$$
$$\operatorname{trace}(\Sigma F\lambda_i) = 0 \qquad (3.2)$$

$i = 1,2,\ldots,h$

where $Fa = (f_{ij}a) = \left(\frac{\partial f_{ij}}{\partial a}\right)$

In the case of sampling on 2 occasions from an infinite universe these equations admit a simple solution. This case is discussed below.

4. SAMPLING ON TWO OCCASIONS

To simplify notation in this case let $\alpha_2 = \alpha$ and $\lambda_1 = \lambda$.

To estimate $R_2 = (\bar{Y}_2/\bar{X}_2)$ where $\bar{Y}_2$ is the population mean for the Y characteristic at time 2 and similarily for X_2 (i.e. $Y_2 = \frac{1}{N}\sum_{i=1}^{N} Y_i$).

$$\text{let } \hat{R}_2 = \alpha \frac{\bar{y}_2'/\bar{x}_2'}{\bar{y}_1'/\bar{x}_1'} \cdot \frac{u\bar{y}_1'' + m\bar{y}_1'}{u\bar{x}_1'' + m\bar{x}_1'} + (1-\alpha)\frac{u\bar{y}_2'' + m\bar{y}_2'}{u\bar{x}_2'' + m\bar{x}_2'}$$

where:

1) $\bar{y}_2'$ is the sample mean for the $\lambda n = m$ units at time 2, which were retained from the 1st occasion for the Y characteristic and similarily for $\bar{x}_2'$.
ii) $\bar{y}_1'$ is the sample mean at time 1 for the Y characteristic based on the m retained units and similarly for $\bar{x}_1'$.
iii) $\bar{y}_1''$ is the sample mean for the Y characteristic at time 1 based on the $(1-\lambda)n=u$ units not in the retained portion of the sample at time 1 and similarly for $\bar{x}_1''$.
iv) $\bar{y}_2''$ is the sample mean based on the $(1-\lambda)n=u$ units sampled anew at time 2 for characteristic Y and similarly for $\bar{x}_2''$.
v) α is a real number between 0 and 1.

Hence: $\hat{R}_2 = \alpha\,(r_2'/r_1')\,r_1 + (1-\alpha)r_2$

where $r_2 = (u\bar{y}_2'' + m\bar{y}_2')/(u\bar{x}_2'' + m\bar{x}_2')$, $r_i' = \bar{y}_i'/\bar{x}_i'$

If N is sufficiently large we can assume independence between the nonmatched samples at time 2 and time 1 (i.e. $\bar{x}_2''$ & $\bar{x}_1''$ are independent).

Expanding $\hat{R}_2$ in a Taylor Series about expected values and ignoring the terms of order 2 and higher you get:

$$\begin{aligned}
\tilde{R}_2 = {} & (\bar{Y}_2/\bar{X}_2) + (\bar{y}_2' - \bar{Y}_2)(1/\bar{X}_2)(\alpha + (1-\alpha)(m/n)) \\
& - (\bar{x}_2' - \bar{X}_2)(\bar{Y}_2/\bar{X}_2^2)(\alpha + (1-\alpha)(m/n)) \\
& - (\bar{y}_1' - \bar{Y}_1)(\alpha(u/n))(\bar{Y}_2/\bar{X}_2\bar{Y}_1) \\
& + (\bar{x}_1' - \bar{X}_1)(\alpha(u/n))(\bar{Y}_2/\bar{X}_2\bar{X}_1) \\
& + (\bar{y}_1'' - \bar{Y}_1)(\alpha(u/n))(\bar{Y}_2/\bar{X}_2\bar{Y}_1) \\
& - (\bar{x}_1'' - \bar{X}_1)(\alpha(u/n))(\bar{Y}_2/\bar{X}_2\bar{X}_1) \\
& + (\bar{y}_2'' - \bar{Y}_2)((1-\alpha)(u/n))(1/\bar{X}_2) \\
& - (\bar{x}_2'' - \bar{X}_2)((1-\alpha)(u/n))(\bar{Y}_2/\bar{X}_2^2)
\end{aligned}$$

The variance of this expression can be computed to get trace (ΣF).

$$F = T \cdot \begin{bmatrix} 1 & (\lambda - \frac{\lambda}{\alpha} - 1) \\ (\lambda - \frac{\lambda}{\alpha} - 1) & 1 + (1/T) \end{bmatrix} \qquad (4.1)$$

where $T = (1/n)(\alpha^2(1-\lambda)/\lambda)$

In this case the system (2.2) reduces to

$$\operatorname{trace}(\Sigma F_\lambda) = 0$$
$$\operatorname{trace}(\Sigma F_\alpha) = 0 \qquad (4.2)$$

The solution of this system that minimizes Var $(\tilde{R}_h)$ is:

opt $\lambda = (1/2)(1-2\text{ opt }\alpha)/(1-\text{opt }\alpha)$ and

opt $\alpha = (1/2) - (1/2)\,(\mathrm{Var}(Z_2 - Z_1)/(\mathrm{Var}(Z_2) + \mathrm{Var}(Z_1)))^{1/2}$

It is noted, at this point, that this function relating the optimal value for λ to the optimal value for α is it's own inverse function. Thus there is a duality, of sorts, between the optimal weight for combining 2 estimators into one and the optimal overlap for "merging" of the 2 samples. This value of λ is called the optimal overlap. The ratio of this minimum variance to the variance of the usual linearized estimate of R_2 based on a SRS without replacement of size n at time 2 is:

$$(V\ \min/V\ \mathrm{SRS}) = 1 + (\alpha^2/(1-2\alpha))(1 + (\mathrm{Var}(Z_1)/\mathrm{Var}(Z_2)) - (2\,\mathrm{Cov}(Z_1Z_2)/\mathrm{Var}(Z_2))(1 + (\lambda/\alpha) - \lambda))$$

If Z_1 and Z_2 are positively correlated then the composite estimator provides an improvement and this ratio can be written as:

$$1 - \left(\frac{1}{2} - \frac{1}{2}H\right)^2 \left[\frac{\mathrm{Var}(Z_1) + \mathrm{Var}(Z_2)}{\mathrm{Var}(Z_2)}\right]$$

where $$H = \left[\frac{\mathrm{Var}(Z_1 - Z_2)}{\mathrm{Var}(Z_1) + \mathrm{Var}(Z_2)}\right]^{\frac{1}{2}}$$

The random variables Z_1 and Z_2 upon which this sampling design depends can be related, via their variance/covariance structure, to the 2 estimates which form the composite estimator. The variance of Z_2 is simply the variance of the linearization of r_2. $Z_1^2 \doteq (\bar{Y}_2/\bar{X}_2)\times(\bar{X}_1/\bar{Y}_1)\times(\bar{y}_1/\bar{x}_1) + K$, where K is a constant. Thus z_1 looks alot like the adjusted estimator $(r_{2,1}'/r_{1,2}')\,r_1$.

Note that this variance ratio is a function of

$\delta = (\mathrm{Var}(Z_1)/\mathrm{Var}(Z_2))^{1/2}$ and

$\rho = \rho\,(Z_1 Z_2)$, the correlation between Z_1 and Z_2.

$$\text{Hence } \frac{V\min}{V\,\mathrm{srs}} = 1 - \left[\frac{1}{2} - \frac{1}{2}H(\delta,\rho)\right]^2 \left[\frac{1}{\delta^2} + 1\right]$$

where $H(\delta,\rho) = (1+\delta^2 - 2\rho\delta)/(1+\delta^2)$

Table 1 gives values of this ratio for various values of ρ and values of δ^2 near unity. The nature of Z_1 and Z_2, as described on the previous page, would make one expect δ^2 to lie near one.

5. APPLICATIONS

This type of estimator was originally studied for the special case of data on 2 occasions discussed in section 3. This estimator is intended for possible use in strata which are quite large relative to the sample size so that the results derived in section 3 (assuming an infinite universe) would remain applicable.

In the case of using data from 3 or more occasions, numerical methods may be necessary in order to solve the equations (3.2) for optimal sampling parameters. An alternate method that might be reasonably adequate would be to iterate the results derived in section 3 for data on 2 occasions. Set $\lambda_0 = 1.0$ in order to estimate $E(Z'Z)$, where $Z = (Z_0, Z_1)$, from a sample selected in year zero and retained in year one. λ_1 would be estimated from this historical data then for α_2 use

$$\alpha_2 = (1/2)-(1/2)(\mathrm{Var}(Z_1-Z_2)/(\mathrm{Var}(Z_1)+\mathrm{Var}(Z_2)))^{1/2}$$

and for λ_2 use $\lambda_2 = (1/2)((1-2\alpha_2)/(1-\alpha_2))$

and so on for the third, fourth, fifth and succeeding occasions. This is:

$$\alpha_{h-1} = \frac{1}{2}-\frac{1}{2}\left[\frac{\mathrm{Var}\,(Z_{h-2}-Z_{h-1})}{\mathrm{Var}\,(Z_{h-2})+\mathrm{Var}\,(Z_{h-1})}\right]^{\frac{1}{2}}$$

and $\lambda_{h-1} = (1/2)((1-2\alpha_{h-1})/(1-\alpha_{h-1}))$

for $h = 4, 5,$ ______

(NOTE: estimate α_h from overlap between time h & h-1)

A conservative estimate of the improvement over SRS at occasion h of this composite estimation scheme would be $1-((1/2)-(1/2)H)^2\,((1/\delta^2)+1)$

where $\delta^2 = (\mathrm{Var}(Z_{h-1})/\mathrm{Var}(Z_h))^{1/2}$

and $H = ((1+\delta^2-2\rho\delta)/(1+\delta^2))^{1/2}$

with $\rho = \rho\,(Z_{h-1}, Z_h)$.

6. EMPIRICAL ANALYSIS

In order to test the theory presented here, 22 distinct populations (data sets) with 5 years (years 0 to 4) of information on 3000 units were constructed and samples of size 100 were selected. The zeroth year's sample was retained in year one (i.e. $\lambda_0 = 1.0$) in order to estimate $E(Z'Z)$ where $Z = (Z_1, Z_2)$. These estimates are what is referred to in section 5. as historical data, from which λ_1, the overlap between year one and year two, is estimated. For year two a composite estimate was made using the procedure detailed in section 4 for sampling on 2 occasions; λ_2 being estimated from the overlap sample between time one and two. Composite estimates for year three and year four were obtained using the leapfrog approach outlined in section 5. These estimates were then compared to the true population values using a relative mean square error measure described below. The usual ratio estimator, r_h, was also computed at each time h=1,2,3,4 and its relative mean square error was computed. This experiment was then replicated 29 times on each of the 22 different populations. For each of the 22 different populations and for each of years 2, 3, and 4 the relative mean square error was estimated using the 29 replications for both the usual ratio estimator and the composite estimator. The ratio of these measures was then computed and tabulated below. In addition, the number of times that the composite estimator was closer to the true value than the usual estimator in 29 replications is tabulated for each of the populations and years.

The first 11 data sets were constructed using Model One for eleven different values of P. Model One is as follows:

$$x_{ij} = \alpha_x + \beta_x \cdot i + \gamma_x f_{xij} + (1-\gamma_x) f_{xj}$$

$$y_{ij} = \alpha_y + \beta_y \cdot i + \gamma_y f_{yij} + (1-\gamma_y) f_{yj}$$

i=1, 2, 3, 4, 5 (year)
j=1, 2, --- 3000 (establishment)

$f_{yij} = N(0,1)$ if i=1
$f_{xij} = N(0,1)$

1 VARIANCE RATIOS

CORRELATION	δ^2				
ρ	.8	.9	1.0	1.1	1.2
.1	.9988	.9988	.9987	.9986	.9986
.2	.9951	.9947	.9944	.9942	.9939
.3	.9882	.9874	.9867	.9860	.9855
.4	.9775	.9759	.9746	.9734	.9723
.5	.9620	.9594	.9571	.9551	.9533
.6	.9402	.9361	.9325	.9293	.9265
.7	.9095	.9032	.8977	.8930	.8888
.8	.8652	.8555	.8472	.8402	.8342
.9	.7949	.7792	.7662	.7557	.7472
1.0	.6180	.5597	.5	.5098	.5185

$$f_{y_{ij}} = f_{y_{i-1,j}} + N(0,P) \qquad \text{if } i > 1$$
$$f_{x_{ij}} = f_{x_{i-1,j}} + N(0,P)$$
$$f_{y_j} = N(0,5) \qquad f_{x_j} = N(0,1.5) \quad f_{yj} \;\&\; f_{xj}$$

generated for each establishment, j=1,2,--- 3000 and α_x β_x γ_x α_y β_y and γ_y are constants $\alpha_y = 40$, $\beta_y = 2$, $\gamma_y = .3$, $\alpha_x = 6$, $\beta_x = .8$, $\gamma_x = .25$ were used. The N(a,b) were generated using the Normal random number generator at the National Institutes of Health Computer Center. Four random numbers were generated for each pair (i, j) , j=1, ... 3000 and i=1,2,...,5 and all of these 60,000 random numbers are assumed to be independent.

The second 11 data sets were constructed using Model Two for eleven different values of q. Model Two is as follows:

$$x_{ij} = 8 + f_{xj} + f_{x_{ij}}$$
$$y_{ij} = 40 + 3 \cdot i + f_{yj} + f_{yij}$$
$$i = 1,2,\ldots,5$$
$$j = 1,2,\ldots,3000$$

$$f_{xj} = N(0,0.8)$$
$$f_{yj} = (49/8)\, f_{xj} + N(0, \tfrac{49}{50})$$

$$\left.\begin{array}{l} f_{yij} = N(0, 40/q) \\ f_{xij} = N(0, 8/q) \end{array}\right\} \quad i = 1$$

$$\left.\begin{array}{l} f_{yij} = (1/\sqrt{i}) \cdot f_{y_{i-1,j}} + \sqrt{(i-1)/i} \cdot N(0, 40/q) \\ f_{xij} = (1/\sqrt{i}) \cdot f_{x_{i-1,j}} + \sqrt{(i-1)/i} \cdot N(0, 8/q) \end{array}\right\} \quad i > 1$$

As in Model One, four random numbers were generated for each pair (i, j) and all of these 60,000 random numbers are assumed to be independent.

The relative mean square error for each year and each of the 22 populations is measured by:

$$(1/m) \sum_{j=1}^{m} ((\text{COMP})_j - \text{ACTUAL})^2/\text{ACTUAL}^2$$
$$\& \qquad (1/m) \sum_{j=1}^{m} ((\text{RATIO})_j - \text{ACTUAL})^2/\text{ACTUAL}^2$$

where m is the number of times this experiment was replicated (m=29); $(\text{COMP})_j$ is the value of the composite estimator on the j^{th} replication and $(\text{RATIO})_j$ is the value of the ratio estimator on the j^{th} replication. "ACTUAL", is the true population value of the parameter being estimated. The number that is tabulated in table two and three is the relative mean square error for the composite estimator divided by the relative mean square error for the ratio estimator. For each of the 22 populations, the number of times that the composite estimator was superior to the usual ratio estimator is tabulated as the integer directly under the ratio of mean square errors.

7. Conclusions

In Model I it is curious to note that the leapfrog approach outlined in section five worked very well in general the first year it is applied (year 3) but deteriorates after that (year 4). This is probably because the weights differ more from their optimum values when they are estimated as for a two year composite estimator, when instead, the estimator uses four years of data. This suggests that this type of estimator could be profitably used with three years of data. That is, $\hat{R}_h$ could be:

$$\hat{R}_h = \alpha_h (\alpha_{h-1}(r_{h-1,h-2}/r_{h-2,h-1})\, r_{h-2} + (1-\alpha_{h-1}) r_{h-1})(r_{h,h-1}/r_{h-1,h}) + (1-\alpha_h) r_h$$

TABLE 2.

MODEL I

P	ρ	Year 2	Year 3	Year 4
.1	.999	.475	.327	.700
		17	21	20
.2	.999	.490	.285	.802
		18	23	23
.4	.997	.542	.328	.788
		17	22	24
.6	.993	.560	.309	.821
		18	20	21
1.0	.981	.658	.402	.856
		16	22	23
2.0	.931	.790	.666	.854
		18	17	20
3.0	.862	.944	.848	.865
		15	16	18
4.0	.788	1.063	.941	.964
		16	14	24
5.0	.716	1.163	.908	.996
		12	16	16
6.0	.651	1.062	.976	.991
		13	10	20
7.0	.593	1.050	.888	.761
		15	14	16

TABLE 3.

MODEL II

Q	ρ	Year 2	Year 3	Year 4
.1	.713	.997	1.055	1.069
		17	10	10
.2	.714	.988	1.057	1.039
		15	10	13
.4	.715	.962	1.09	.930
		16	11	14
.8	.719	.970	1.012	.890
		16	13	11
1.5	.732	.974	.974	.782
		15	14	13
4.0	.802	.968	1.169	.683
		17	15	21
7.0	.878	.992	.971	.999
		18	19	14
10	.923	.892	.919	.996
		16	15	21
20	.975	.723	.666	.814
		18	19	24
40	.993	.625	.344	.799
		18	22	18
70	.997	.534	.288	.856
		20	20	24

and α_h, α_{h-1}, λ_{h-1} and λ_{h-2} should be estimated as outlined in section five.

In Model II one can expect some gain from the use of composite estimation the first year it is used (i.e. with sampling on two occasions). In general, using the leapfrog approach of section five, gives results that are similar to those of Model I but the trends are not so clear cut. It should be noted that Model II agrees with the variance assumptions as stated in section two while Model I, as a test for robustness, deviates from these assumptions.

These estimation techniques need to be further tested on "real" data and compared to other types of composite estimators that use difference adjustments rather than ratio adjustments. Conditions under which one type of adjustment gives smaller mean square error than the other needs further investigation.

I would like to thank Stuart Scott for his review of this paper and Alfreda Reeves for her labors in the typing of this paper.

BIBLIOGRAPHY

1. Kendall, M., Stuart, A. (1969), The Advanced Theory of Statistics, Vol. I, London: Charles W. Griffin & Co.
2. Wolter K.M., "Composite Estimation in Finite Populations", JASA, Sept. 1979.
3. Woodruff, R.S., "The Use of Rotating Samples in the Census Bureau's Monthly Surveys", JASA, June 1963.
4. Eckler, A.R., "Rotation Sampling", JRSS, 1955.
5. Patterson, H.D., "Sampling on Successive Occasions with Partial Replacement of Units", JRSS, 1950.
6. Cochran, G.W., "Sampling Techniques" (second edition), Wiley.
7. Raj, D., "Sampling Theory", McGraw-Hill.
8. Hansen, M.H., Hurwitz, W.N., Madow, W.J., (1953)"Sample Survey Methods and Theory." John Wiley & Sons, Inc. New York.