CHAPTER 2

Unified Sampling Theory: Design-Based Inference

2.1 INTRODUCTION

In this chapter we consider the inferential aspects of sampling from a finite population under a fixed population setup, where each of the unit is associated with a fixed unknown real number. A sample is selected from the population using some man-made randomization procedure called sampling design. The design-based inference is based on all possible samples that might be selected according to the sampling design. Expectation is the average of all possible samples. Different types of linear unbiased estimators have been proposed and conditions of unbiasedness of the estimators have been derived. The nonexistence theorems proposed by Godambe (1955), Hanurav (1966), and Basu (1971) have been discussed. Concepts of admissibility and sufficient statistics in finite population sampling have been introduced as well.

2.2 DEFINITIONS AND TERMINOLOGIES

2.2.1 Noninformative and Adaptive (Sequential) Sampling Designs

A sampling design p is said to be noninformative if the selection probability $p(s)$ of a sample s does not depend on the value of the study variable y. In adaptive or sequential sampling procedures, the selection probability $p(s)$ may depend on the values of the variable of interest y for the units selected in the sample s.

2.2.2 Estimator and Estimate

After selection of a sample s using a suitable sampling design p, information on the study variable y is collected from each of the units selected in the sample. Here we assume that all units in the sample have responded and there is no measurement error in measuring a response, i.e., the true value

Survey Sampling Theory and Applications
ISBN 978-0-12-811848-1
http://dx.doi.org/10.1016/B978-0-12-811848-1.00002-9

y_i, of the study variable y is obtained from each of the ith unit ($i \in s$). The information gathered from the selected units in the sample and their values y_i's is known as data, and it will be denoted by $d = ((i,y_i), i \in s)$. The collection of all possible values of d is known as the sample space, and it will be denoted by $\mathcal{X}$. A real valued function $T(s,\mathbf{y}) = T(d)$ of d on the sample space $\mathcal{X}$ is known as a statistic. When the statistic $T(s,\mathbf{y})$ is used as a guess value of a certain parametric function $\theta = \theta(\mathbf{y})$ of interest (such as the population mean, total, median etc.), we call $T(s,\mathbf{y})$ as an estimator of the parameter θ. Obviously, an estimator is a random variable whose value depends on the sample selected (i.e., data). The numerical value of an estimator for a given data is called an estimate.

2.2.3 Unbiased Estimator

An estimator $T = T(s,\mathbf{y})$ is said to be design unbiased (p-unbiased or unbiased) for estimating a population parameter θ if and only if

$$E_p(T) = \sum_{s \in \mathfrak{S}} T(s,\mathbf{y})p(s) = \theta \quad \forall \quad \mathbf{y} \in R^N \tag{2.2.1}$$

where, E_p denotes the expectation with respect to the sampling design p, $p(s)$ is the probability of the selection of the sample s according to design p, $\mathfrak{S}$ is the collection of all possible samples, and R^N is the N-dimensional Euclidean space. The class of all unbiased estimators of θ satisfying (2.2.1) will be denoted by C_θ.

An estimator, which is not unbiased, is called a biased (or design biased) estimator. The amount of bias of an estimator T is defined as

$$B(T) = E_p(T) - \theta = \sum_{s \in \mathfrak{S}} T(s,\mathbf{y})p(s) - \theta \tag{2.2.2}$$

2.2.4 Mean Square Error and Variance

The mean square error of an estimator T is denoted by

$$M(T) = E_p(T-\theta)^2 = \sum_{s \in \mathfrak{S}} [T(s,\mathbf{y}) - \theta]^2 \tag{2.2.3}$$

The mean square error measures the closeness of an estimator T around the true value θ.

The variance of an estimator T with respect to the sampling design p is denoted by

$$V_p(T) = E_p[T - E_p(T)]^2 = \sum_{s \in \mathfrak{S}} [T(s,\mathbf{y}) - E_p(T)]^2 p(s) \tag{2.2.4}$$

It can be easily checked that

$$M(T) = V_p(T) + [B(T)]^2 \tag{2.2.5}$$

For an unbiased estimator $B(T) = 0$ and hence the mean square is equal to its variance.

2.2.5 Uniformly Minimum Variance Unbiased Estimator

Let T_1 and $T_2(\neq T_1)$ be two unbiased estimators that belong to a certain class of unbiased estimators C_θ. The estimator T_1 is said to be better than T_2 if:

(i) $V_p(T_1) \leq V_p(T_2) \quad \forall \quad \mathbf{y} \in R^N$
and

(ii) the strict inequality $V_p(T_1) < V_p(T_2)$ holds for at least one $\mathbf{y} \in R^N$.

In case at least one of the estimators T_1 and T_2 is biased, T_1 is said to be better than T_2 if:

(i) $M(T_1) \leq M(T_2) \ \forall \ \mathbf{y} \in R^N$
and

(ii) the strict inequality $M(T_1) < M(T_2)$ holds for at least one $\mathbf{y} \in R^N$.

An estimator T_0 belonging to the class of unbiased estimators C_θ is called the uniformly minimum variance unbiased estimator (UMVUE) for estimating a parametric function θ if T_0 is better than any other unbiased estimators belonging to the class C_θ, i.e., any $\widetilde{T}(\neq T_0) \in C_\theta$ satisfies

$$\begin{aligned} &\text{(i)} \ V_p(T_0) \leq V_p\left(\widetilde{T}\right) \ \forall \ \mathbf{y} \in R^N \\ &\quad \text{and} \\ &\text{(ii)} \ V_p(T_0) < V_p\left(\widetilde{T}\right) \ \text{for at least one} \ \mathbf{y} \in R^N \end{aligned} \tag{2.2.6}$$

2.3 LINEAR UNBIASED ESTIMATORS

In case θ is a linear function of $\mathbf{y}$, such as population total Y or mean $\overline{Y}$, we very often use a linear estimator for Y as follows:

$$t^* = t^*(s, \mathbf{y}) = a_s + \sum_{i \in s} b_{si} y_i \tag{2.3.1}$$

where, a_s, a known constant, depends on the selected sample s but is independent of the units selected in the sample and their y-values. b_{si}'s are known constants free from y_i's, $i \in s$, but may be dependent on the selected sample s and units $i(\in s)$. $\sum_{i \in s}$ denotes the sum over distinct units in s.

In case a_s in (2.3.1) is equal to zero, then t^* reduces to a linear homogeneous unbiased estimator for Y and it is given by

$$t = t(s, \mathbf{y}) = \sum_{i \in s} b_{si} y_i \tag{2.3.2}$$

The different choices of the constants a_s and b_{si}'s yield different estimators. Our objective is to choose certain specific estimators, which must possess certain desirable properties.

2.3.1 Conditions of Unbiasedness

The estimator t^* in (2.3.1) will be unbiased for the population total Y if and only if

$$E_p(t^*) = Y \quad \forall \quad \mathbf{y} \in R^N \tag{2.3.3}$$

Now noting

$$\begin{aligned} E_p(t^*) &= \sum_{s \in \mathfrak{S}} t^*(s, \mathbf{y}) p(s) \\ &= \sum_{s \in \mathfrak{S}} a_s p(s) + \sum_{s \in \mathfrak{S}} \left(\sum_{i \in s} b_{si} y_i \right) p(s) \\ &= \sum_{s \in \mathfrak{S}} a_s p(s) + \sum_{s \in \mathfrak{S}} \left(\sum_{i=1}^{N} I_{si} b_{si} y_i \right) p(s) \\ &\quad \text{(writing } I_{si} = 1 \text{ if } i \in s \text{ and } I_{si} = 0 \text{ if } i \notin s) \\ &= \sum_{s \in \mathfrak{S}} a_s p(s) + \sum_{i=1}^{N} y_i \left(\sum_{s \in \mathfrak{S}} I_{si} b_{si} p(s) \right) \\ &= \alpha_0 + \sum_{i=1}^{N} \alpha_i y_i \end{aligned} \tag{2.3.4}$$

where, $\alpha_0 = \sum_{s \in \mathfrak{S}} a_s p(s)$ and $\alpha_i = \sum_{s \in \mathfrak{S}} I_{si} b_{si} p(s) = \sum_{s \supset i} b_{si} p(s)$.

From Eqs. (2.3.3) and (2.3.4), we note that t^* is unbiased for Y if and only if

$$\alpha_0 + \sum_{i=1}^{N} \alpha_i y_i = Y \quad \forall \quad \mathbf{y} \in R^N \tag{2.3.5}$$

Now, putting $\mathbf{y} = \mathbf{y}^{(0)} = (0,\ldots,0,\ldots,0)$, all coordinates of $\mathbf{y}$ are zero and $\mathbf{y} = \mathbf{y}^{(i)} = (0,\ldots,y_i,\ldots,0)$ whose ith coordinate y_i is nonzero and the remaining coordinates are zero, in (2.3.5) the unbiasedness condition (2.3.5) reduces to

$$\alpha_0 = \sum_{s\in\mathfrak{S}} a_s p(s) = 0 \quad \text{and} \quad \alpha_i = \sum_{s\supset i} b_{si} p(s) = 1 \quad \forall \quad i = 1,\ldots,N \tag{2.3.6}$$

Substituting $a_s = 0$ in (2.3.5), we find the condition of unbiasedness of a linear homogeneous estimator $t = \sum_{i\in s} b_{si} y_i$ for the total Y as

$$\alpha_i = \sum_{s\supset i} b_{si} p(s) = 1 \quad \forall \quad i = 1,\ldots,N \tag{2.3.7}$$

If $\pi_i > 0$, then $b_{si} = 1/\pi_i$ meets the unbiased condition (2.3.7). On the other hand, if $\pi_i = \sum_{s\supset i} p(s) = \sum_{s\in\mathfrak{S}} I_{si} p(s) = 0$, then $I_{si} = 0$ for s with $p(s) > 0$ and hence we cannot find b_{si}'s ($\neq 0$) that satisfies the unbiasedness condition (2.3.7). This leads to the following theorem attributed to Godambe (1955).

Theorem 2.3.1

The necessary and sufficient condition for existence of a linear unbiased estimator t of the population total Y is that the inclusion probability π_i should be positive for all $i = 1,\ldots,N$.

2.3.2 Horvitz–Thompson Estimator

Consider the linear homogeneous unbiased estimator $t = \sum_{i\in s} b_{si} y_i$ for the total Y. If we put $b_{si} = c_i$ in the expression of t, then the unbiasedness condition (2.3.7) yields $c_i = 1/\pi_i$. In this case the estimator t reduces to

$$\widehat{Y}_{ht} = \sum_{i\in s} y_i/\pi_i \tag{2.3.8}$$

The estimator $\widehat{Y}_{ht}$ is known as Horvitz–Thompson (1952) estimator for the population total Y.

For an simple random sampling without replacement (SRSWOR), sampling design of size n, $\pi_i = n/N$ and the Horvitz–Thompson estimator (HTE) reduces to

$$\widehat{Y} = N\,\overline{y}(s) \quad \text{with} \quad \overline{y}(s) = \sum_{i\in s} y_i/n \tag{2.3.9}$$

2.3.3 Hansen–Hurwitz Estimator

If we take $b_{si} = k\, n_i(s)$, with k as a constant and $n_i(s)$ = number of times ith unit is repeated in s, then the unbiasedness condition (2.3.7) reduces to

$$k\sum_{s\supset i} n_i(s)p(s) = k\sum_{s\in\mathfrak{S}} n_i(s)p(s) = k\, E_p\{n_i(s)\} = 1$$

and the estimator t reduces to

$$t = \sum_{i\in s} n_i(s)y_i/E_p(n_i(s)) \tag{2.3.10}$$

In particular, for a probability proportional to size with replacement (PPSWR) sampling design with normed size measure p_i for the ith unit, $E_p(n_i(s)) = np_i$ and (2.3.10) reduces to Hansen–Hurwitz (1943) estimator

$$\widehat{Y}_{hh} = \frac{1}{n}\sum_{i\in s} n_i(s)\frac{y_i}{p_i} \tag{2.3.11}$$

The PPSWR sampling reduces to simple random sampling with replacement (SRSWR), if $p_i = 1/N\ \forall\, i = 1,\ldots,N$ and in this case we get

$$t = \frac{N}{n}\overline{y}_n$$

where, $\overline{y}_n$ is the sample mean of all the n units including repetition.

2.3.4 Unbiased Ratio Estimator

Let us choose $b_{si} = c_i/p(s)$. In this case the unbiasedness condition (2.3.7) reduces to $c_i = 1/\beta_i$, where $\beta_i = \sum_{s\supset i} = \sum_{s\in\mathfrak{S}} I_{si}$ = total number of times ith unit appears in all possible samples with $p(s) > 0$ and the estimator (2.3.2) reduces to

$$t = \sum_{i\in s}\frac{1}{\beta_i}\frac{y_i}{p(s)}$$

In case $\mathfrak{S}$ consists of all possible $\binom{N}{n}$ samples each of n distinct units with positive probabilities, then $\beta_i = \binom{N-1}{n-1} = M_1$ (say) and the expression of t becomes

$$t = \frac{1}{M_1}\sum_{i\in s}\frac{y_i}{p(s)} \tag{2.3.12}$$

For the Lahiri–Midzuno–Sen (LMS) sampling scheme, $p(s) = x_s/(M_1X)$, where $x_s = \sum_{i\in s} x_i$, $X = \sum_{i\in U} x_i$ and $x_i(>0)$ is a known positive number (measure of size) for the ith unit, the estimator (2.3.12) reduces to the unbiased ratio estimator for population total Y proposed by LMS (1951, 1952, 1953) and it is given by

$$\widehat{Y}_{lms} = (y_s/x_s)X \tag{2.3.13}$$

where $y_s = \sum_{i\in s} y_i$

2.3.5 Difference and Generalized Difference Estimator

Let $t\ (s, y) = \sum_{i\in s} b_{si}y_i$ be a linear homogeneous unbiased estimator of the total Y, x_i the known value of a certain character x of the ith unit, and $X = \sum_{i=1}^{N} x_i$. Then the linear estimator

$$t_D = \sum_{i\in s} b_{si}y_i - \beta\left(\sum_{i\in s} b_{si}x_i - X\right) \tag{2.3.14}$$

is unbiased for the total Y for any known value β. The estimator (2.3.14) is known as a difference estimator. In particular when $b_{si} = 1/\pi_i$ and $\beta = 1$, (2.3.14) takes the following elegant form

$$t_D = \sum_{i\in s} y_i/\pi_i - \left(\sum_{i\in s} x_i/\pi_i - X\right) \tag{2.3.15}$$

The estimator t_D is known as generalized difference estimator.

2.4 PROPERTIES OF THE HORVITZ–THOMPSON ESTIMATOR

The HTE of the finite population total Y is

$$\widehat{Y}_{ht} = \sum_{i\in s} y_i/\pi_i$$

where, $\sum_{i\in s}$ denotes sum over distinct units in s and $\pi_i > 0\ \forall\ i = 1,\ldots,N$.

The HTE is widely used in survey sampling theory because it possesses some important properties. The properties are presented in this section.

Theorem 2.4.1

Let $I_{si} = 1$ for $i \in s$ and $I_{si} = 0$ for $i \notin s$. Then,
(i) $E_p(I_{si}) = \pi_i$, (ii) $V_p(I_{si}) = \pi_i(1 - \pi_i)$, and (iii) $C_p(I_{si},I_{sj}) = -\Delta_{ij}$ for $i \neq j$.
where C_p is the covariance operator with respect to the sampling design p and $\Delta_{ij} = \pi_i\pi_j - \pi_{ij}$.

Proof

(i) $E_p(I_{si}) = \sum_s I_{si} p(s)$

$= \pi_i$

(ii) $V_p(I_{si}) = E_p(I_{si})^2 - \pi_i^2$

$= E_p(I_{si}) - \pi_i^2$

$= \pi_i(1 - \pi_i)$

and

(iii) $C_p(I_{si}, I_{sj}) = E_p(I_{si}I_{sj}) - E_p(I_{si})E_p(I_{sj})$

$= \sum_s I_{si}I_{si}p(s) - \pi_i\pi_j$

$= \pi_{ij} - \pi_i\pi_j$

$= -\Delta_{ij}$

Theorem 2.4.2

(i) $\widehat{Y}_{ht}$ is unbiased for Y i.e., $E_p\left(\widehat{Y}_{ht}\right) = Y$

(ii) The variance of $\widehat{Y}_{ht}$ is

$$V_p\left(\widehat{Y}_{ht}\right) = \sum_{i=1}^{N} \gamma_i^2 \left(\frac{1}{\pi_i} - 1\right) + \sum_{i\neq}^{N} \sum_{j=1}^{N} \gamma_i\gamma_j \left(\frac{\pi_{ij}}{\pi_i\pi_j} - 1\right)$$

$$= V_{ht}$$

(iii) An unbiased estimator of $V_p\left(\widehat{Y}_{ht}\right)$ is

$$\widehat{V}_{ht} = \sum_{i\in s} \gamma_i^2 \frac{1-\pi_i}{\pi_i^2} + \sum_{i\neq} \sum_{j\in s} \gamma_i\gamma_j \frac{\pi_{ij} - \pi_i\pi_j}{\pi_{ij}}$$

for $\pi_{ij} > 0$ for $i \neq j = 1,\ldots,N$.

Proof

(i) $E_p\left(\widehat{Y}_{ht}\right) = E_p\left(\sum_{i\in s} \frac{\gamma_i}{\pi_i}\right)$

$= E_p\left(\sum_{i=1}^{N} \frac{\gamma_i}{\pi_i} I_{si}\right)$

$= \sum_{i=1}^{N} \frac{\gamma_i}{\pi_i} E_p(I_{si})$

$= Y$ (since $E_p(I_{si}) = \pi_i$)

(ii) $$V_p\left(\widehat{Y}_{ht}\right) = V_p\left(\sum_{i=1}^{N}\frac{\gamma_i}{\pi_i}I_{si}\right)$$

$$= \sum_{i=1}^{N}\frac{\gamma_i^2}{\pi_i^2}\{V_p(I_{si})\} + \sum_{i\neq}^{N}\sum_{j=1}^{N}\frac{\gamma_i}{\pi_i}\frac{\gamma_j}{\pi_j}\{C_p(I_{si}, I_{sj})\}$$

Now noting $V_p(I_{si}) = \pi_i(1-\pi_i)$ and $C_p(I_{si},I_{sj}) = \pi_{ij} - \pi_i\pi_j$, we find $V_p\left(\widehat{Y}_{ht}\right) = V_{ht}$.

(iii) Writing, $\widehat{V}_{ht} = \sum_{i=1}^{N}\frac{\gamma_i^2}{\pi_i^2}(1-\pi_i)I_{si} + \sum_{i\neq}^{N}\sum_{j=1}^{N}\frac{\gamma_i}{\pi_i}\frac{\gamma_j}{\pi_j}\frac{\pi_{ij}-\pi_i\pi_j}{\pi_{ij}}I_{si}I_{sj}$, we get

$$E_p\left(\widehat{V}_{ht}\right) = \sum_{i=1}^{N}\frac{\gamma_i^2}{\pi_i^2}(1-\pi_i)E_p(I_{si}) + \sum_{i\neq}^{N}\sum_{j=1}^{N}\frac{\gamma_i}{\pi_i}\frac{\gamma_j}{\pi_j}\frac{\pi_{ij}-\pi_i\pi_j}{\pi_{ij}}E_p(I_{si}I_{sj})$$

$$= V_{ht} \text{ (since } E_p(I_{si}) = \pi_i \text{ and } E_p(I_{si}I_{sj}) = \pi_{ij})$$

Theorem 2.4.3

For a fixed effective size (ν) sampling design,

(i) $V_p\left(\widehat{Y}_{ht}\right) = V_{ht} = \frac{1}{2}\sum_{i\neq}^{N}\sum_{j=1}^{N}\Delta_{ij}\left(\frac{\gamma_i}{\pi_i} - \frac{\gamma_j}{\pi_j}\right)^2 = V_{YG}$

(ii) An unbiased estimator of $V_p\left(\widehat{Y}_{ht}\right)$ is given by

$$\widehat{V}_{YG} = \frac{1}{2}\sum_{i\neq}\sum_{j\in s}\frac{\Delta_{ij}}{\pi_{ij}}\left(\frac{\gamma_i}{\pi_i} - \frac{\gamma_j}{\pi_j}\right) \text{ for } \pi_{ij} > 0 \text{ for } i \neq j = 1, \ldots, N$$

Proof

(i) $$V_{YG} = \frac{1}{2}\sum_{i\neq}^{N}\sum_{j=1}^{N}\Delta_{ij}\left(\frac{\gamma_i}{\pi_i} - \frac{\gamma_j}{\pi_j}\right)^2$$

$$= \frac{1}{2}\sum_{i\neq}^{N}\sum_{j=1}^{N}\Delta_{ij}\left(\frac{\gamma_i^2}{\pi_i^2} + \frac{\gamma_j^2}{\pi_j^2} - 2\frac{\gamma_i}{\pi_i}\frac{\gamma_j}{\pi_j}\right)$$

$$= \sum_{i=1}^{N}\frac{\gamma_i^2}{\pi_i^2}\sum_{j(\neq i)=1}^{N}\Delta_{ij} - \sum_{i\neq}^{N}\sum_{j=1}^{N}\Delta_{ij}\frac{\gamma_i}{\pi_i}\frac{\gamma_j}{\pi_j}$$

$$= V_{ht}$$

since

$$\sum_{j(\neq i)=1}^{N} \Delta_{ij} = \pi_i \sum_{j(\neq i)=1}^{N} \pi_j - \sum_{j(\neq i)=1}^{N} \pi_{ij}$$

$$= \pi_i(\nu - \pi_i) - (\nu - 1)\pi_i$$

$$= \pi_i(1 - \pi_i) \text{ (using corollary 1.3.2).}$$

(ii) $$E_p\left(\widehat{V}_{YG}\right) = \frac{1}{2}E_p\left[\sum_{i\neq}\sum_{j\in s}\frac{\Delta_{ij}}{\pi_{ij}}\left(\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j}\right)^2\right]$$

$$= \frac{1}{2}\sum_{i\neq}^{N}\sum_{j=1}^{N}\frac{\Delta_{ij}}{\pi_{ij}}\left(\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j}\right)^2 E_p(I_{si}I_{sj})$$

$$= V_{YG}$$

Remark 2.4.1

The variance estimator $\widehat{V}_{ht}$ was proposed by Horvitz and Thompson (1952) and is applicable for any sampling design with $\pi_{ij} > 0$ for $i \neq j = 1,\ldots,N$. The variance estimator $\widehat{V}_{YG}$ was proposed by Yates and Grundy (1953) and is known as the Yates–Grundy variance estimator. $\widehat{V}_{YG}$ is unbiased for only a fixed effective size sampling design. Both the estimators $\widehat{V}_{ht}$ and $\widehat{V}_{YG}$ suffer from the drawback that they can be negative. No simple sufficient condition of nonnegativity is available for $\widehat{V}_{ht}$. The Yates–Grundy variance estimator $\widehat{V}_{YG}$ is nonnegative for a sampling design, for which $\Delta_{ij} \geq 0$. Various sampling designs have been proposed, for which $\Delta_{ij} \geq 0$ for $i \neq j = 1,\ldots,N$. As far as efficiency is concerned, $\widehat{V}_{ht}$ is known to be admissible but $\widehat{V}_{YG}$ is inadmissible. Details are given in Section 2.6.

2.5 NONEXISTENCE THEOREMS

We will call the collection of the estimators $t = \sum_{i\in s} b_{si}y_i$, whose coefficients b_{si}'s satisfy the unbiasedness condition (2.3.7) as the class of linear homogeneous unbiased estimator C_{lh}. The class of the linear unbiased estimators C_l comprises estimators (2.3.1) and is subject to (2.3.6). The class of all possible unbiased estimators, which includes linear, linear homogeneous, and nonlinear estimators, will be denoted by C_u and clearly, $C_u \supset C_l \supset C_{lh}$.

In Section 2.2, it is shown that for a given sampling design p we can construct numerous linear unbiased estimators for a finite population total Y. Therefore, we need to select the best estimator in the sense of having uniformly minimum variance. Godambe (1955) proved that in the class of linear homogeneous unbiased estimators C_{llh}, the UMVUE does not exist, i.e., none of the estimators can be termed as the best. Hanurav (1966) modified Godambe's result and showed the existence of the UMVUE for a unicluster sampling design (defined in the section next). Basu (1971) generalized the result further and proved the nonexistence of the UMVUE in the class of all unbiased estimators C_u. Godambe (1955) showed that the HTE is the UMVUE in the reduced subspace R_0 of the parameter space R^N, where $R_0 = \bigcup_{i=1}^{N} \mathbf{y}^{(i)}$ and $\mathbf{y}^{(i)}$ = vector $\mathbf{y}$, whose ith coordinate y_i is nonzero and the remaining coordinates are zeros.

2.5.1 Unicluster Sampling Design

A design $\widetilde{p}$ is called as a unicluster design if any two samples $s, s^* \in \mathfrak{S}$ with $\widetilde{p}(s), \widetilde{p}(s^*) > 0$ imply either (i) $s \cap s^* = \phi$ (null set) or (ii) the samples s and s^* are equivalent ($s \sim s^*$) in the sense that they contain the same set of distinct units.

Example 2.5.1

Let $U = (1,2,3,4,5,6)$ be a finite population, and $\mathfrak{S}$ consists of the samples $s_1 = (1,2)$, $s_2 = (3,4)$, $s_3 = (5,6)$ with probabilities $p(s_1) = 0.4$ and $p(s_2) = p(s_3) = 0.3$. Then $\widetilde{p}$ is a unicluster sampling design.

2.5.2 Class of Linear Homogeneous Unbiased Estimators

Theorem 2.5.1

In the class of linear homogeneous unbiased estimators C_{llh}, the UMVUE of a finite population total Y based on a noncensus sampling design, p with $\pi_i > 0 \ \ \forall \ i = 1,\ldots,N$ does not exist if the sampling design p is non-unicluster. However, the UMVUE does exist if p is a unicluster design.

Proof

The class C_{llh} consists of the estimators $t(s, \mathbf{y}) = \sum_{i \in s} b_{si} y_i$, where the constants b_{si}'s satisfy the unbiasedness condition (2.3.7) viz.

$$\sum_{s \supset i} b_{si} p(s) = 1 \quad \text{for every} \ \ i = 1, \cdots, N$$

Here we want to find the constants, b_{si}'s, that minimize

$$V_p(t) = \sum_{s \in \mathfrak{S}} \left(\sum_{i \in s} b_{si} y_i \right)^2 p(s) - Y^2$$

subject to the unbiasedness condition (2.3.7).

For minimization, we consider

$$\Psi = \sum_{s \in \mathfrak{S}} \left(\sum_{i \in s} b_{si} y_i \right)^2 p(s) - Y^2 - \sum_{i=1}^{N} \lambda_i \left\{ \sum_{s \supset i} b_{si} p(s) - 1 \right\} \tag{2.5.1}$$

with λ_i's as undetermined Lagrange multipliers.

Differentiating Ψ with respect to b_{si} and equating to zero, we get

$$\frac{\partial \Psi}{\partial b_{si}} = 2 y_i \left(\sum_{i \in s} b_{si} y_i \right) p(s) - \lambda_i p(s) = 0 \tag{2.5.2}$$

Eq. (2.5.2) should be valid for $\forall\ \mathbf{y} \in R^N$. In particular for $\mathbf{y} = \mathbf{y}^{(i)} = (0,\ldots,0,y_i,0,\ldots,0)$ with $y_i \neq 0$, Eq. (2.5.2) yields the optimum value of b_{si} as

$$b_{si} = \lambda_i / \left(2 y_i^2\right) = b_{0i} \text{ (say)} \tag{2.5.3}$$

The unbiasedness condition (2.3.7) and (2.5.3) yield

$$b_{si} = b_{0i} = 1/\pi_i \tag{2.5.4}$$

Thus the UMVUE should necessarily be the HTE viz. $\sum_{i \in s} y_i/\pi_i$ and should satisfy Eq. (2.5.2), i.e.,

$$\sum_{i \in s} y_i/\pi_i = \frac{\lambda_i}{2 y_i} \quad \text{for} \quad \forall \quad i \in s, y_i \neq 0 \tag{2.5.5}$$

For a noncensus non-unicluster sampling design, we always find two different samples, s_1 and s_2, which have a common unit, i.e., $s_1 \cap s_2 \neq \phi$ for which $p(s_1) > 0$ and $p(s_2) > 0$. In this case Eq. (2.5.5) indicates that

$$\sum_{i \in s_1} y_i/\pi_i = \sum_{i \in s_2} y_i/\pi_i \quad \forall \quad \mathbf{y} \in R^N, \tag{2.5.6}$$

which is clearly impossible. Hence the UMVUE does not exist for a noncensus non-unicluster sampling design.

For a non-census unicluster sampling design any two samples s_1 and s_2 with $p(s_1), p(s_2) > 0$ implies $s_1 \cap s_2 = \phi$. In this case, the condition of unbiasedness (2.3.7) yields $b_{si} = 1/\pi_i = 1/p(s)$ for $i \in s$. Hence the

Horvitz–Thompson estimator is the unique linear homogeneous unbiased estimator for the total Y. Therefore, it is trivially the MVUE in C_{lh}.

2.5.3 Optimality of the Horvitz–Thompson Estimator

The following theorem shows the optimality the HTE in the restricted parametric space $R_0 = \bigcup_{i=1}^{n} \mathbf{y}^{(i)} (\subset R^N)$.

Theorem 2.5.2

For any sampling design p, with $\pi_i > 0 \quad \forall\ i = 1,\ldots,N$, the HTE $\widehat{Y}_{ht} = \sum_{i \in s} y_i/\pi_i$ is the UMVUE in C_{lh}, the class of linear homogeneous unbiased estimators of the population total Y when $\mathbf{y} \in R_0$.

Proof

$$
\begin{aligned}
V_p[t(s,y)] &= \sum_{s \in \mathfrak{F}} \left(\sum_{i \in s} b_{si} y_i \right)^2 p(s) - Y^2 \\
&= \sum_{s \in \mathfrak{F}} \left(\sum_{i \in s} b_{si}^2 y_i^2 + \sum_{i \neq} \sum_{j \in s} b_{si} b_{sj} y_i y_j \right) p(s) - Y^2 \\
&= \sum_{i=1}^{N} y_i^2 \left(\sum_{s \supset i} b_{si}^2 p(s) - 1 \right) + \sum_{i \neq}^{N} \sum_{j}^{N} y_i y_j \left(\sum_{s \supset i,j} b_{si} b_{sj} p(s) - 1 \right)
\end{aligned}
\tag{2.5.7}
$$

Substituting $\mathbf{y} = \mathbf{y}^{(j)} = (0,\ldots,0,y_j,0,\ldots,0)$ in (2.5.7), we get

$$
\begin{aligned}
V_p\left[t\left(s, \mathbf{y}^{(j)}\right)\right] &= y_j^2 \left(\sum_{s \supset j} b_{sj}^2 p(s) - 1 \right) \\
&\geq y_j^2 \left[\frac{\left(\sum_{s \supset j} b_{sj} p(s) \right)^2}{\sum_{s \supset j} p(s)} - 1 \right] \text{ (using Cauchy inequality)}
\end{aligned}
$$

The unbiasedness condition $\sum_{s \supset j} b_{si} p(s) = 1$ yields

$$
V_p\left[t\left(s, \mathbf{y}^{(j)}\right)\right] \geq y_j^2 \left(\frac{1}{\pi_j} - 1 \right) \tag{2.5.8}
$$

The equality in (2.5.8) holds if and only if $b_{sj} = 1/\pi_j$. Hence,

$$V_p[t(s,\gamma)] \geq V_p\left(\sum_{i\in s}\frac{\gamma_i}{\pi_i}\right) \text{ for } \forall \ \mathbf{y} \in R_0 \tag{2.5.9}$$

The inequality in (2.5.9) is strict if and only if $t(s,\mathbf{y}) \neq \widehat{Y}_{ht}$. Hence for the restricted parametric space R_0, $\widehat{Y}_{ht} = \sum_{i\in s}\frac{\gamma_i}{\pi_i}$ is the UMVUE in the class C_{lh}.

2.5.4 Class of All Unbiased Estimators

Let $T(s,\mathbf{y})$ be an unbiased estimator for an arbitrary parametric function $\theta = \theta(\mathbf{y})$. The value of $T(s,\mathbf{y})$ depends on the values of γ_i's belonging to the sample s but is independent of γ_i's, which do not belong to s. The value of $\theta = \theta(\mathbf{y})$ depends on all the values of $\gamma_i, i = 1,\ldots N$. Let C_θ be the class of all unbiased estimators of θ. Basu (1971) proved the nonexistence of a UMVUE of $\theta(\mathbf{y})$ in the class C_θ of all unbiased estimators. The theorem is described as follows:

Theorem 2.5.3

For a noncensus design, there does not exist the UMVUE of $\theta = \theta(\mathbf{y})$ in the class of all unbiased estimators C_θ.

Proof

If possible, let $T_0(s,\mathbf{y})(\in C_\theta)$ be the UMVUE of the population parameter $\theta = \theta(\mathbf{y})$. Since the design p is noncensus and the value of $T_0(s,\mathbf{y})$ depends on γ_i's for $i \in s$ but not on the values of γ_i's for $i \notin s$, we can find a known vector $\mathbf{y}^{(a)} = (a_1,\ldots,a_i,\ldots,a_N)$ for which $T_0(s,\mathbf{y}^{(a)}) \neq \theta(\mathbf{y}^{(a)})$ with $p(s) > 0$. Consider the following estimator

$$T^*(s,\mathbf{y}) = T_0(s,\mathbf{y}) - T_0\left(s,\mathbf{y}^{(a)}\right) + \theta\left(\mathbf{y}^{(a)}\right)$$

$T^*(s,\mathbf{y})$ is unbiased for $\theta(\mathbf{y})$ because

$$E_p[T^*(s,\mathbf{y})] = \theta(\mathbf{y}) - \theta\left(\mathbf{y}^{(a)}\right) + \theta\left(\mathbf{y}^{(a)}\right) = \theta(\mathbf{y}).$$

Since $T_0(s,\mathbf{y})$ is assumed to be the UMUVE for $\theta(\mathbf{y})$, we must have

$$V_p[T_0(s,\mathbf{y})] \leq V_p[T^*(s,\mathbf{y})] \quad \forall \ \mathbf{y} \in R^N \tag{2.5.10}$$

Now for $\mathbf{y} = \mathbf{y}^{(a)}$, $V_p[T^*(s,\mathbf{y})] = V_p\left[T^*\left(s,\mathbf{y}^{(a)}\right)\right] = V_p\left[\theta\left(\mathbf{y}^{(a)}\right)\right] = 0$ while $V_p[T_0(s,\mathbf{y}^{(a)})] > 0$ since we assumed $T_0\left(s,\mathbf{y}^{(a)}\right) \neq \theta(\mathbf{y}^{(a)})$ with $p(s) > 0$. Hence the inequality (2.5.10) is violated at $\mathbf{y} = \mathbf{y}^{(a)}$ and the nonexistence of the UMVUE for $\theta(\mathbf{y})$ is proved.

In particular, for $\theta(\mathbf{y}) = Y$, Theorem 2.5.3 reduces to following theorem.

Theorem 2.5.4

For a noncensus design, p, the UMVUE of Y in the class of unbiased estimators C_u does not exist.

2.5.5 Class of Linear Unbiased Estimators

Now writing $\theta(\mathbf{y}) = Y$, $T_0(s,\mathbf{y}) = t_0{}^*(s,\mathbf{y}) = a_{s0} + \sum_{i\in s} b_{si0}y_i$ and

$$T^*(s,\mathbf{y}) = t_0{}^*(s,\mathbf{y}) - t_0{}^*\left(s,\mathbf{y}^{(a)}\right) + \sum_{i=1}^{N} a_i = \sum_{i\in s} b_{si0}(y_i - a_i) + \sum_{i=1}^{N} a_i$$

in Theorem 2.5.3, we derive the following theorem.

Theorem 2.5.5

For a noncensus design, p, the UMVUE of the population total, Y, in the class of linear unbiased estimators C_l does not exist.

2.6 ADMISSIBLE ESTIMATORS

We have seen in Section 2.5 that in almost all practical situations, the UMVUE for a finite population total does not exist. The criterion of admissibility is used to guard against the selection of a bad estimator.

An estimator T is said to be admissible in the class C of estimators for a given sampling design p if there does not exist any other estimator in the class C better than T. In other words, there does not exist an alternative estimator $T^*(\neq T) \in C$, for which following inequalities hold.

$$\begin{aligned} &\text{(i)}\ \ V_p(T^*) \le V_p(T) \quad \forall\, T^*(\neq T)\in C \ \text{ and } \ \mathbf{y}\in R^N \\ &\text{and} \quad \text{(ii)}\ \ V_p(T^*) < V_p(T) \ \text{ for at least one } \ \mathbf{y}\in R^N \end{aligned} \tag{2.6.1}$$

Theorem 2.6.1

In the class of linear homogeneous unbiased estimators C_{lh}, the HTE $\widehat{Y}_{ht}$ based on a sampling design p with $\pi_i > 0\ \forall\ i = 1,\ldots,N$ is admissible for estimating the population total Y.

Proof

The proof is immediate from Theorem 2.5.2. Since $\widehat{Y}_{ht}$ is the UMVUE when $\mathbf{y} \in R_0$, we cannot find an estimator $\forall\, T^*\left(\neq \widehat{Y}_{ht}\right)\in C_{lh}$ for which (2.6.1) holds.

The Theorem 2.6.1 of admissibility of the HTE $\widehat{Y}_{ht}$ in the class C_{lh} was proved by Godambe (1960). Godambe and Joshi (1965) proved the admissibility of $\widehat{Y}_{ht}$ in the class of all unbiased estimators C_u, and it is given in Theorem 2.6.2.

Theorem 2.6.2

For a given sampling design p with $\pi_i > 0 \ \forall \ i = 1,\ldots,N$, the HTE $\widehat{Y}_{ht}$ is admissible in the class C_u of all unbiased estimator for estimating the total Y.

Proof

Suppose $\widehat{Y}_{ht}$ is not admissible in C_u, then there exists an estimator $t(s,\mathbf{y})(\neq \widehat{Y}_{ht}) \in C_u$, for which the following inequalities hold.

$$\begin{aligned} &\text{(i)} \ \ V_p[t(s,\mathbf{y})] \leq V_p\left(\widehat{Y}_{ht}\right) \ \ \forall \ \ \mathbf{y} \in R^N \\ &\qquad\qquad \text{and} \\ &\text{(ii)} \ \ V_p[t(s,\mathbf{y})] < V_p\left(\widehat{Y}_{ht}\right) \ \text{at least one} \ \ \mathbf{y} \in R^N \end{aligned} \tag{2.6.2}$$

The estimator $t(s,\mathbf{y})$ can be written as

$$t(s,\mathbf{y}) = \widehat{Y}_{ht} + h(s,\mathbf{y}) \tag{2.6.3}$$

$$\text{i.e.,} \ h(s,\mathbf{y}) = t(s,\mathbf{y}) - \widehat{Y}_{ht}$$

Since $t(s,\mathbf{y})$ and $\widehat{Y}_{ht}$ are unbiased for the total Y, we must have

$$E_p[h(s,\mathbf{y})] = \sum_{s \in \mathfrak{S}} h(s,\mathbf{y})p(s) = 0 \tag{2.6.4}$$

Furthermore, (2.6.2) yields

$$V_p[h(s,y)] + 2C_p\left[\widehat{Y}_{ht}, h(s,y)\right] \leq 0$$

i.e.,

$$\sum_{s \in \mathfrak{S}} \{h(s,y)\}^2 p(s) + 2\sum_{s \in \mathfrak{S}} h(s,y)\left(\sum_{i \in s} \frac{y_i}{\pi_i}\right)p(s) \leq 0 \quad \forall \ \ \mathbf{y} \in R^N \tag{2.6.5}$$

Let us define $\mathbf{y}_{(j)}$ = collection of all vectors $\mathbf{y} = (y_1,\ldots,y_k,\ldots,y_N)$ having exactly j $(=0,1,\ldots,N)$ nonzero coordinates and $\mathfrak{S}_{(j)}(\subset \mathfrak{S})$ collection of samples consisting of units with nonzero $\mathbf{y}$-values for exactly j units and $\mathfrak{S}$ is the collections of all possible samples with positive selection probabilities. Clearly, $\mathbf{y}_{(j)} \cap \mathbf{y}_{(k)} = \phi$ for $j \neq k = 0,1,\ldots,N$ and $\bigcup_{j=1}^{N} \mathbf{y}_{(j)} = R^N$.

To prove this the theorem, we prove the following lemma by Godambe and Joshi (1965).

Lemma 2.6.1

$h(s,\mathbf{y}) = 0$ for $\forall\ s \in \mathfrak{F}$ and $\forall\ \mathbf{y} \in \mathbf{y}_{(j)}$ implies $h(s,\mathbf{y}) = 0$ for $\forall s \in \mathfrak{F}$ and $\forall\ \mathbf{y} \in \mathbf{y}_{(j+1)}, j = 0, 1, .., N-1$.

Proof

Let $h(s,\mathbf{y}) = 0$ for $\forall s \in \mathfrak{F}$ and $\forall\ \mathbf{y} \in \mathbf{y}_{(j)}$, then for any $\forall\ \mathbf{y} \in \mathbf{y}_{(j+1)}$, Eqs. (2.6.4) and (2.6.5) yield

$$\sum_{i=0}^{j+1} \sum_{s \in \mathfrak{F}_{(i)}} h(s,y)p(s) = 0 \tag{2.6.6}$$

and

$$\sum_{i=0}^{j+1} \sum_{s \in \mathfrak{F}_{(i)}} \{h(s,y)\}^2 p(s) + 2\sum_{i=0}^{j+1} \sum_{s \in \mathfrak{F}_{(i)}} h(s,y)\left(\sum_{i \in s} \frac{y_i}{\pi_i}\right)p(s) \le 0 \tag{2.6.7}$$

Now $h(s,\mathbf{y}) = 0$ for $\forall s \in \mathfrak{F}$ and $\forall\ \mathbf{y} \in \mathbf{y}_{(j)}$ implies $h(s,\mathbf{y}) = 0$ for $\forall s \in \mathfrak{F}_{(i)}, i = 0, 1, \ldots, j$. Hence (2.6.6) and (2.6.7) implies

$$\sum_{s \in \mathfrak{F}_{(j+1)}} h(s,\mathbf{y})p(s) = 0 \tag{2.6.8}$$

and

$$\sum_{s \in \mathfrak{F}_{(j+1)}} \{h(s,\mathbf{y})\}^2 p(s) + 2\sum_{s \in \mathfrak{F}_{(j+1)}} h(s,\mathbf{y})\left(\sum_{i \in s} \frac{y_i}{\pi_i}\right)p(s) \le 0 \tag{2.6.9}$$

The magnitude of $\sum_{i \in s} \frac{y_i}{\pi_i}$ is the same as $\sum_{i=1}^{N} \frac{y_i}{\pi_i}$ for every $s \in \mathfrak{F}_{(j+1)}$ because, when $\mathbf{y} \in \mathbf{y}_{(j+1)}$, each of the samples belonging to $\mathfrak{F}_{(j+1)}$ contains all the $j+1$ nonzero y-values that belong to $\mathbf{y}_{(j+1)}$. Thus (2.6.9) yields

$$\sum_{s \in \mathfrak{F}_{(j+1)}} \{h(s,\mathbf{y})\}^2 p(s) + \left(\sum_{i=1}^{N} \frac{y_i}{\pi_i}\right) \sum_{s \in \mathfrak{F}_{(j+1)}} h(s,\mathbf{y})p(s) \le 0 \quad \forall\ \mathbf{y} \in \mathbf{y}_{(j+1)}$$

Finally using (2.6.8), we get

$$h(s,\mathbf{y}) = 0 \quad \forall s \in \mathfrak{F}_{(j+1)} \text{ and } \forall\ \mathbf{y} \in \mathbf{y}_{(j+1)} \tag{2.6.10}$$

Since $\mathfrak{F} = \bigcup_{i=1}^{j+1} \mathfrak{F}_{(i)}$, (2.6.10) implies

$$h(s,\mathbf{y}) = 0 \quad \forall\ s \in \mathfrak{F} \text{ and } \forall\ \mathbf{y} \in \mathbf{y}_{(j+1)} \tag{2.6.11}$$

Hence the lemma is proved.

For $\mathbf{y} \in \mathbf{y}_{(0)}$, $\sum_{i \in s} \frac{y_i}{\pi_i} = 0$, hence using (2.6.5), we find $h(s,\mathbf{y}) = 0 \ \forall \ s \in \mathfrak{S}$ with $p(s) > 0$. Now from the Lemma 2.6.1, we see that

$$h(s, \mathbf{y}) = 0 \ \forall \ s \in \mathfrak{S} \text{ and } \forall \ \mathbf{y} \in R^N \tag{2.6.12}$$

Finally Eqs. (2.6.3) and (2.6.12) yield $t(s, y) = \widehat{Y}_{ht}$. Hence there does not exist an estimator, $t(s,\mathbf{y})(\neq \widehat{Y}_{ht}) \in C_u$, for which (2.6.2) holds. Thus $\widehat{Y}_{ht}$ is an admissible in the class of unbiased estimators C_u.

Remark 2.6.1

The admissible estimator is not unique. In fact we can construct a number of admissible estimators in a given class of estimators. It follows from the following corollary derived by Cassel et al. (1977).

Corollary 2.6.1

For any sampling design p with $\pi_i > 0 \ \forall \ i = 1,\ldots,N$, the generalized difference estimator

$$t_D = \sum_{i \in s} \frac{y_i - a_i}{\pi_i} + A,$$

with a_i's are known constants and $A = \sum_{i=1}^{N} a_i$ is admissible in the class C_u.

Proof

From Theorem 2.6.2, we note that $\sum_{i \in s} \frac{y_i - a_i}{\pi_i}$ is admissible in the class of all unbiased estimators of $Y - A$. Hence $\sum_{i \in s} \frac{y_i - a_i}{\pi_i} + A$ is admissible in the class of all unbiased estimators C_u for estimating Y.

Remark 2.6.2

Godambe and Joshi (1965) showed that Theorem 2.6.2 is also valid for the subspace $-a < y_i < b$ for $\forall \ i = 1,\ldots,N$ with $a,b \geq 0$ of the parameter space R^N.

Remark 2.6.3

Godambe and Joshi (1965) also proved that the variance estimator $\widehat{V}_{ht}$ is admissible in the entire class of unbiased estimators of V_{ht}.

Remark 2.6.4

It is important to note that if an estimator is admissible in a class C, then it is also admissible in a subclass of $C^*(\subset C)$ containing the estimator. Hence, Theorem 2.6.1 can be proved as a corollary of Theorem 2.6.2.

For further results of admissibility, readers are referred to Joshi (1965a,b), Cassel et al. (1977), Patel and Dharmadikari (1978), and Sengupta (1980), among others.

2.7 SUFFICIENCY IN FINITE POPULATION

An estimator $e(s,\mathbf{y})$ is said to be inadmissible in a class of estimators C if there exists an estimator $e^*(s,y)(\in C)$ better than $e(s,y)$. Hence it is natural to question how an inadmissible estimator could be improved. The method of improvement of an inadmissible estimator with the aid of sufficient statistics is known as Rao–Blackwellization. The concept of sufficient statistics in survey sampling was introduced by Basu (1958), while the concepts of linear sufficiency, distribution-free sufficient statistics, and Bayesian sufficiency were also introduced by Godambe (1966, 1968). Details have been given by Cassel et al. (1977), Chaudhuri and Stenger (1992), and Thompson and Seber (1996).

Let $s = (i_1, \ldots, i_k, \ldots, i_{n_s})$ be an ordered sample of size n_s selected from a population U with probability $p(s)$ using a sampling design p, where the unit i_k is selected at the kth draw. After selection of sample s, the responses $y_{i_1}, \ldots, y_{i_{n_s}}$ were obtained from sampled units $i_1, \ldots, i_{n_s}$, respectively. The ordered data based on the ordered sample s are denoted by $d = \{(i_1, y_{i_1}), \ldots, (i_k, y_{i_{n_s}})\} = (i_k, y_{i_k}; i_k \in s)$.

Let $\tilde{s} = (j_1, j_2, \ldots, j_{\nu_s})$ with $j_1 < j_2 < \ldots < j_{\nu_s}$ be the unordered sample obtained from s by taking distinct units of s and arranging them in ascending order of their labels. Let us denote the operator r, which transforms the ordered sample s to the unordered sample $\tilde{s}$, i.e., $r(s) = \tilde{s}$. The unordered data are denoted by $\tilde{d} = \left\{(j_1, y_{j_1}), \ldots, \left(j_{\nu_s}, y_{j_{\nu_s}}\right)\right\} = \left(j, y_j; j \in \tilde{s}\right)$. We define the operator R to get unordered data $\tilde{d}$ from ordered data d, i.e., $R(d) = \tilde{d}$.

Example 2.7.1

Let $U = (1,2,3,4,5)$, $\mathbf{y} = (10,15,15,20,10)$, and $s = (5,2,5)$. Here $y_1 = 10$, $y_2 = 15$, $y_3 = 15$, $y_4 = 20$, and $y_5 = 10$; $r(s) = \tilde{s} = (2,5)$, $d = \{(5,10),(2,15),(5,10)\}$ and $R(d) = \tilde{d} = \{(2,15),(5,10)\}$.

2.7.1 Sufficiency and Likelihood

Suppose that prior to the survey the surveyor had no idea about the values of the parameter $\mathbf{y} = (y_1,\ldots,y_N)$ and hence $\Omega_y = R^N$, the N-dimensional Euclidean space was considered the parametric space. After surveying the

sample s, the surveyor collects the ordered data $d = \left(i_k, y_{i_k}; i_k \in s\right)$. The data d is said to be consistent with parameter vector $\mathbf{y}_0 = (y_{10},\ldots,y_{i0},\ldots,y_{N0})$ if $y_{i_k} = y_{i_k 0}$ for $i_k \in s$. After collection the data d, the surveyor knows that the values of y_i's belong to s and hence the parametric space Ω_y reduces to Ω_{yd} $(\subset \Omega_y)$, where Ω_{yd} consists of the vectors $\mathbf{y}$ with $y_j = y_{j0}$ for $j \in s$.

The unordered data $\widetilde{d} = (i_{j_k}, y_{j_k}; j_k \in \widetilde{s})$ obtained from d will be consistent with $\mathbf{y}_0$ if $y_{jk} = y_{jk0}$ for $j_k \in \widetilde{s}$, i.e., $y_{j_1} = y_{j_1 0}, \ldots, y_{j_{\nu_s}} = y_{j_{\nu_s} 0}$. Here also, given the unordered data $\widetilde{d}$, the parametric space Ω_y reduces to $\Omega_{\widetilde{yd}}$, which consists of the vectors $\mathbf{y}$ with $y_j = y_{j0}$ for $j \in \widetilde{s}$. Clearly, $\Omega_{yd} = \Omega_{\widetilde{yd}}$ as s and $\widetilde{s}$ consist of the same set of distinct units.

Example 2.7.2

Consider the population $U = (1,2,3,4)$ of 4 units from which an ordered sample $s = (1,2,2)$ is selected. Let the y-values of the units selected in the sample s be $y_1 = 50$ and $y_2 = 100$. In this case, $d = \{(1,50),(2,100),(2,100)\}$; $\Omega_y = (-\infty < y_1 < \infty, -\infty < y_2 < \infty, -\infty < y_3 < \infty, -\infty < y_4 < \infty) = R^4$, $\widetilde{d} = \{(1,50),(2,100)\}$, $\Omega_{\widetilde{yd}} = \Omega_{yd} = (50, 100, -\infty < y_3 < \infty, -\infty < y_4 < \infty)$. Here both d and $\widetilde{d}$ are consistent with parameter $\mathbf{y} = (50,100,500,600)$ but inconsistent with $\mathbf{y} = (100,100,500,600)$. Data (D), a random variable, depends on the selection of the sample and realization of the parametric vector $\mathbf{y}$. Given data $D = d$, the likelihood function of the parameter $\mathbf{y}$ was obtained by Godambe (1966) as

$$\begin{aligned} L(\mathbf{y}|d) &= P\{D = d\} \\ &= \begin{cases} p(s) & \text{for } \mathbf{y} \in \Omega_{yd} \\ 0 & \text{for } \mathbf{y} \notin \Omega_{yd} \end{cases} \\ &= p(s)\ I_d(\mathbf{y}) \quad \text{for } \forall\ \mathbf{y} \in \Omega_y \end{aligned} \tag{2.7.1}$$

where, $I_d(\mathbf{y})$ is an indicator variable defined as

$$I_d(\mathbf{y}) = \begin{cases} 1 & \text{for } \mathbf{y} \in \Omega_{yd} \\ 0 & \text{for } \mathbf{y} \notin \Omega_{yd} \end{cases} \tag{2.7.2}$$

The likelihood function (2.7.2) is flat (constant), equal to $p(s)$ for $\mathbf{y} \in \Omega_{yd}$, and zero outside Ω_{yd}. Hence no unique maximum likelihood of $\mathbf{y}$ exists, and the likelihood function is noninformative.

Theorem 2.7.1

For a noninformative sampling design p, the unordered data $\widetilde{d} = R(d)$ derived from an ordered data d is a sufficient statistic for $\mathbf{y}$.

Proof

Suppose an ordered sample s is selected with probability $p(s)$ using a sampling design p and $d = \left(i_k, y_{i_k}; i_k \in s\right)$ be the ordered data based on s. Here we denote D as a random variable, which takes the value d with probability

$$P\{D = d\} = p(s)\ \ I_d(\mathbf{y})\ \text{ for }\ \forall\ \ \mathbf{y} \in \Omega_y$$

Similarly, if we denote $\widetilde{D}$ as a random variable of getting an unordered data $\widetilde{d}$, then

$$\begin{aligned} P\left(\widetilde{D} = \widetilde{d}\right) &= p(\widetilde{s}) = \sum_{s|r(s)=\widetilde{s}} p(s)\ \text{ for }\ \mathbf{y} \in \Omega_{\underset{yd}{\sim}}\ \text{and} \\ P\left(\widetilde{D} = \widetilde{d}\right) &= 0\ \text{ when }\ \mathbf{y} \notin \Omega_{\underset{yd}{\sim}} \end{aligned} \tag{2.7.3}$$

where, $\sum_{s|r(s)=\widetilde{s}}$ denotes the sum over the ordered samples s, which produces the unordered sample $\widetilde{s}$.

Furthermore, Eq. (2.7.3) can be written as

$$P\left\{\widetilde{D} = \widetilde{d}\right\} = \left\{\sum_{s|r(s)=\widetilde{s}} p(s)\right\} I_{\underset{d}{\sim}}(y)\ \text{ for }\ \mathbf{y} \in \Omega_y$$

The conditional distribution of obtaining ordered data d_0 (based on an ordered sample s_0), given unordered data $\widetilde{d}_1$, is given by

$$P\left(D = d_0 \middle| \widetilde{D} = \widetilde{d}_1\right) = \frac{P\left(D = d_0 \cap \widetilde{D} = \widetilde{d}_1\right)}{P\left(\widetilde{D} = \widetilde{d}_1\right)} \tag{2.7.4}$$

Now there are two cases.

Case I: $P(\widetilde{D} = \widetilde{d}_1) > 0$ and $R(d_0) = \widetilde{d}_1$, here $I_{d_0}(\mathbf{y}) = I_{\underset{d_1}{\sim}}(\mathbf{y})$ and

$$P\left(D = d_0 \middle| \widetilde{D} = \widetilde{d}_1\right) = \frac{P(D = d_0)}{P\left(\widetilde{D} = \widetilde{d}_1\right)} = \frac{ps_0}{\left\{\sum_{s|r(s)=\widetilde{s}_1} p(s)\right\}} \tag{2.7.5}$$

Case II: $P\left(\widetilde{D}=\widetilde{d}_1\right)>0$ and $R(d_0)\neq\widetilde{d}_1$, here $P\left(D=d_0\cap\widetilde{D}=\widetilde{d}_1\right)=0$ and hence

$$P\left(D=d_0\middle|\widetilde{D}=\widetilde{d}_1\right)=0 \tag{2.7.6}$$

From (2.7.5) and (2.7.6), we note that $P\left(D=d_0\middle|\widetilde{D}=\widetilde{d}_1\right)$ is independent of parameter $\mathbf{y}$ as $p(s)$ is noninformative (does not involve $\mathbf{y}$). Hence, the unordered data $\widetilde{d}=R(d)$ is a sufficient statistic for $\mathbf{y}$.

2.7.2 Minimal Sufficient Statistic

A statistic $f(d)$, a function of data d, partitions the sample space of d. Let $\mathcal{P}_f$ be the partition associated with f. The statistic $f^*(d)$ is called a minimal sufficient statistic if and only if for any statistic $f(\neq f^*)$, each partition set of $\mathcal{P}_f$ is a subset of a partition set $\mathcal{P}_{f^*}$ induced by f^*. In other words, every set of $\mathcal{P}_{f^*}$ can be expressed as the union of the sets of $\mathcal{P}_f$.

Theorem 2.7.2

For a noninformative sampling design p, the unordered data $\widetilde{d}=R(d)$ derived from an ordered data d is a minimal sufficient statistic for $\mathbf{y}$.

Proof

Let s_1 and s_2 be two samples with $p(s_1)>0$ and $p(s_2)>0$ yielding data d_1 and d_2, respectively. Following Thompson and Seber (1996), we note that $\widetilde{d}$ is a minimal sufficient for $\mathbf{y}$ if for any two data points d_1 and d_2, the following holds:

$$\widetilde{d}_1=R(d_1)=R(d_2)=\widetilde{d}_2\Leftrightarrow P(D=d_1)=kP(D=d_2)\forall\ \ \mathbf{y}\in\Omega_y \tag{2.7.7}$$

where k is a constant independent of $\mathbf{y}$.

Let $\widetilde{d}_1=\widetilde{d}_2$, then $I_{\underset{\sim}{d}_1}(\mathbf{y})=I_{d_1}(\mathbf{y})=I_{\underset{\sim}{d}_2}(\mathbf{y})=I_{d_2}(\mathbf{y})$ and

$$P(D=d_1)=p(s_1)I_{d_1}(\mathbf{y})=\frac{p(s_1)}{p(s_2)}p(s_2)I_{d_2}(\mathbf{y})=kP(D=d_2)$$

where $k=\dfrac{p(s_1)}{p(s_2)}$ is independent of $\mathbf{y}$.

Similarly, $P(D=d_1)=k\ P(D=d_2)\forall\ \mathbf{y}\in\Omega_y$ implies

$$p(s_1)I_{d_1}(\mathbf{y})=k\ \ p(s_2)I_{d_2}(\mathbf{y})\ \ \forall\ \ \mathbf{y}\in\Omega_y \tag{2.7.8}$$

Since, $p(s_1), p(s_2)>0$ and $I_{d_1}(\mathbf{y})$ and $I_{d_2}(\mathbf{y})$ can take only two values 0 or 1, Eq. (2.7.8) implies $I_{d_1}(\mathbf{y})=I_{d_2}(\mathbf{y})$, i.e., $\widetilde{d}_1=R(d_1)=R(d_2)=\widetilde{d}_2$. Hence $\widetilde{d}$ is a minimal sufficient statistic for $\mathbf{y}$.

2.7.3 Rao–Blackwellization

Let an ordered sample s be selected from a population with a probability $p(s)$ and d be the corresponding ordered data. Let $t(d)$ be an estimator (not necessarily unbiased) for a parametric function $\theta(\mathbf{y}) = \theta$ and $\tilde{t}\left(\tilde{d}\right) = E_p\left[t(d)\middle|\tilde{d}\right]$, where $\tilde{d}$ be the unordered data obtained from d. Then we have the following theorem.

Theorem 2.7.3

(i) $E_p\{t(d)\} = E_p\left\{\tilde{t}\left(\tilde{d}\right)\right\}$

and

(ii) $MSE\{t(d)\} = E_p[t(d) - \theta]^2 \geq MSE\left\{\tilde{t}\left(\tilde{d}\right)\right\} = E_p\left[\tilde{t}\left(\tilde{d}\right) - \theta\right]^2$

Proof

(i) $E_p[t(d)] = E_p\left[E_p\left\{t(d)\middle|\tilde{d}\right\}\right] = E_p\left[\tilde{t}\left(\tilde{d}\right)\right]$

(ii)
$$\begin{aligned} MSE[t(d)] &= E_p[t(d) - \theta]^2 \\ &= E_p\left[\left\{\tilde{t}\left(\tilde{d}\right) - \theta\right\} + \left\{t(d) - \tilde{t}\left(\tilde{d}\right)\right\}\right]^2 \\ &= E_p\left\{\tilde{t}\left(\tilde{d}\right) - \theta\right\}^2 + E_p\left\{t(d) - \tilde{t}\left(\tilde{d}\right)\right\}^2 \end{aligned}$$

Now noting $E_p\left[\left\{\tilde{t}\left(\tilde{d}\right) - \theta\right\}\left\{t(d) - \tilde{t}\left(\tilde{d}\right)\right\}\right] = E_p\left[\left\{\tilde{t}\left(\tilde{d}\right) - \theta\right\} E_p\left(\left\{t(d) - \tilde{t}\left(\tilde{d}\right)\right\}\middle|\tilde{d}\right)\right] = 0$, we find

$$\begin{aligned} MSE\{t(d)\} &= MSE\left\{\tilde{t}\left(\tilde{d}\right)\right\} + E_p\left\{t(d) - \tilde{t}\left(\tilde{d}\right)\right\}^2 \\ &\geq MSE\left\{\tilde{t}\left(\tilde{d}\right)\right\} \end{aligned}$$

Corollary 2.7.1

In case $t(d)$ is unbiased for θ, we find that,

(i) $\tilde{t}\left(\tilde{d}\right)$ is also unbiased for θ.

and

(ii) $V_p\left\{\tilde{t}\left(\tilde{d}\right)\right\} \leq V_p\{t(d)\}$.

Corollary 2.7.2

An estimator based on an ordered data can be improved by applying Rao–Blackwellization technique. Hence we conclude that an estimator based on ordered data is inadmissible.

Remark 2.7.1

Theorem 2.7.3 is greatly used for improvement of the conventional ordered estimators of various sampling designs such as SRSWR, PPSWR, and probability proportional to size without replacement (PPSWOR) sampling schemes.

Remark 2.7.2

It is natural to question as to whether two different unbiased estimators based on the same unordered data (sufficient statistic) for a particular population parameter, θ exists. The answer is "yes" because, although the unordered data $\widetilde{d}$ is a sufficient statistic, it is not a complete sufficient statistic (vide Theorem 2.7.4 below). Hence one can construct an infinite number of unbiased estimators as a function of the sufficient statistic for estimating a population parameter θ. But none of them attains the minimum variance for all possible values of θ, as we have shown earlier that the UMVUE does not exist.

Theorem 2.7.4 (Cassel et al., 1977)

The statistic $\widetilde{d}$ is not a complete sufficient statistic.

Proof

By definition, the statistic $\widetilde{d}$ will be a complete sufficient statistic for the parameter $\mathbf{y}$ if for any function $H\left(\widetilde{d}\right)$ the following condition holds:

$$E\left\{H\left(\widetilde{d}\right)\right\} = 0 \quad \forall \quad \mathbf{y} \in \Omega_y \Rightarrow P\left\{H\left(\widetilde{d}\right) = 0\right\} = 1 \quad \forall \quad \mathbf{y} \in \Omega_y \qquad (2.7.9)$$

To prove the theorem, consider a finite population $U = \{1,2,3,4\}$ of known size $N = 4$. Let $s_1 = (1,1,2)$, $s_2 = (1,2,2)$, $s_3 = (3,3,4)$ and $s_4 = (4)$ be all the possible samples with respective selection probabilities $p(s_1) = 0.10$, $p(s_2) = 0.20$, $p(s_3) = 0.30$ and $p(s_4) = 0.40$. Here the order samples s_1, s_2, s_3, and s_4 generates the unordered samples $\widetilde{s}_1 = (1,2)$, $\widetilde{s}_2 = (3,4)$, and $\widetilde{s}_3 = (4)$ with probabilities $p(\widetilde{s}_1) = 0.30$, $p(\widetilde{s}_2) = 0.30$, and $p(\widetilde{s}_3) = 0.40$. The inclusion probabilities of the four units are $\pi_1 = \pi_2 = 0.30$, $\pi_3 = 0.30$, and $\pi_4 = 0.70$. Consider the function

$$H\left(\widetilde{d}\right) = \sum_{i \in \widetilde{s}} \frac{1}{\pi_i} - N = \sum_{i \in \widetilde{s}} \frac{1}{\pi_i} - 4$$

Here $E\left\{H\left(\widetilde{d}\right)\right\} = 0 \quad \forall \quad \mathbf{y} \in \Omega_y$. But, $H\left(\widetilde{d}_1\right) = \frac{1}{\pi_1} + \frac{1}{\pi_2} - 4 = 2.667$, $H\left(\widetilde{d}_2\right) = \frac{1}{\pi_3} + \frac{1}{\pi_4} - 4 = 0.762$, and $H\left(\widetilde{d}_3\right) = \frac{1}{0.70} - 4 = -2.571$ are all nonzero. Hence $\widetilde{d}$ is not a complete sufficient statistic.

2.8 SAMPLING STRATEGIES

So far, we have compared performances of several estimators for a given sampling design. It is found that UMVUE exists only for a unicluster sampling design. Among the sampling designs that are used in practice, only the linear systematic sampling scheme (described in Chapter 4) is a unicluster sampling design. Other sampling designs do not belong to this category. So, it is a natural to search for a combination of a sampling design and an estimator, which may yield better estimators than an alternative combination of a sampling design and an estimator. A combination of an estimator and a sampling design is known as a sampling strategy. We will denote a sampling strategy by the pair $h = (p,t)$.

2.8.1 Unbiased Strategy

The strategy $h = (p,t)$ is to be p-unbiased for the parameter θ if the associated estimator t is p-unbiased, i.e., $E_p(t) = \theta \ \forall \ \mathbf{y} \in R^N$.

2.8.2 Uniformly Minimum Variance Unbiased Strategy

An unbiased strategy $h_1 = (p_1,t_1)$ is said to be better than another unbiased strategy $h_2 = (p_2,t_2)$ for estimating θ if

$$V_{p_1}(t_1) \leq V_{p_2}(t_2) \ \text{ for } \ \forall \ \ \mathbf{y} \in R^N$$

and

$$V_{p_1}(t_1) < V_{p_2}(t_2) \ \text{ for at least one } \ \mathbf{y} \in R^N.$$

A strategy $h_0 = (p_0,t_0)$ belonging to the class of unbiased strategies $\mathcal{H}$ is said to be the uniformly minimum variance unbiased strategy in $\mathcal{H}$ if

$$V_{p_0}(t_0) \leq V_p(t) \ \text{ for } \ h = (p,t)(\neq h_0) \in \mathcal{H} \ \text{ and } \ \forall \ \ \mathbf{y} \in R^N$$

Cassel et al. (1977) showed that there does not exist a UMV strategy in the class $\mathcal{H}_u(n)$ of unbiased fixed effective size n sampling strategies. The class $\mathcal{H}_u(n)$ consists of strategies $h = (p,t)$, where $p \in P_n$ and $t \in C_u$, P_n is the class of fixed effective size n sampling design, and C_u is the class of unbiased estimators.

2.8.3 Admissible Strategies

A strategy h_0 belonging to the class of strategies $\mathcal{H}$ is said to be admissible in $\mathcal{H}$ if there does not exist a strategy in $\mathcal{H}$, which is better than h_0.

Admissibility of strategies was considered by Joshi (1966), Godambe (1969), Ramakrishnan (1975), and Ericson (1970) among others. Good

details have been given by Cassel et al. (1977). The concept of hyperadmissibility was introduced by Hanurav (1966). It was found that HTE is the unique hyperadmissible estimator in a certain class of estimators. The results of the comparisons of strategies and properties of admissibility have not been presented here. However, the comparisons of some strategies under super population models have been discussed in Chapter 6.

2.8.4 Minimax Strategy

A strategy $h_0 = (p_0, t_0)$ belonging to the class $\mathcal{H}$ of strategies is said to be a minimax strategy in $\mathcal{H}$ for estimating parameter θ if and only if

$$\inf_{h \in \mathcal{H}} \left[\sup_{\mathbf{y} \in \Omega_\gamma} E_p (t - \theta)^2 \right] = \sup_{\mathbf{y} \in \Omega_\gamma} E_{p_0} (t_0 - \theta)^2$$

The minimax strategies for estimating the finite population mean were studied by several authors including Aggarwal (1959, 1966), Godambe (1960), Godambe and Joshi (1965), and Chaudhuri (1969) and they derived minimax property of the sample mean in the restricted parameter space. Details are given by Cassel et al. (1977).

2.9 DISCUSSIONS

In a fixed population setup, the population vector $\mathbf{y}$ is considered as fixed, i.e., each unit is associated with a real number. The vector $\mathbf{y}$ is considered unknown and is called the parameter. Our object is to estimate certain parametric function $\theta = \theta(\mathbf{y})$ by selecting a sample by using a suitable sampling design. After collecting the sample, information of the study variable y is collected from the selected sample. The sample and the observed y-values constitute the data d. An estimator is a function of data d. The estimator is a random variable since its value varies from one data to another. The expectation is the average value of the estimator computed over all possible samples that might be realized by selection through a sampling design. The design-based inference comprises all possible samples that can be selected from a sampling design with positive probabilities. Given a sampling design, an unbiased estimator can be chosen in various ways. Godambe (1955) proved that for a given sampling design, UMVUE doest not exist in the class of linear homogeneous unbiased estimators. Hanurav (1966) showed that the UMVUE exists only in unicluster sampling design, which includes systematic sampling design. Basu (1971)

proved the nonexistence of UMVUE in the class of all unbiased estimators. Because of the nonexistence of the UMVUE, alternative criteria for admissibility were proposed to guard against selection of unsuitable estimators. It was found that the HTE is admissible for both the classes of linear unbiased and all unbiased estimators. Numerous admissible estimators for estimating a parametric function exist. The concept of sufficiency in finite population sampling was introduced by Basu (1958). Basu showed that unordered data constitute a sufficient statistic. An estimator based on ordered data is inadmissible. An inadmissible estimator can be improved by using a sufficient statistic. This technique is known as Rao–Blackwellization. The unordered data are a sufficient statistic but not a complete sufficient statistic. So, we can construct numerous unbiased estimators for a particular parameter θ, which are function of a sufficient statistic. But, none of them is the UMVUE. The combination of a sampling design and an estimator is known as a sampling strategy. The best sampling strategy does not exist in a finite population setup. However, several admissible strategies for estimating a parameter θ exist.

2.10 EXERCISES

2.10.1 Consider the marks of four students

Students	1	2	3	4
Marks	80	52	30	70

Select a sample using the following design:

Sample	(1, 2, 1)	(1, 3, 4)	(1, 3)	(2, 4)	(1, 2, 3)	(2, 3, 4)
Probability	0.2	0.1	0.4	0.1	0.1	0.1

(i) Compute the HTE of the total marks of the four students.
(ii) Obtain the variance of the HTE.
(iii) Give an unbiased estimate of the variance of the HTE.

2.10.2 The following data relate consumption of food in six households in a certain locality.

Households	1	2	3	4	5	6
Expenditure on food ($)	1500	2000	3000	2500	1500	3000

Select a sample of three households using the following design.

Sample	(1,2,3)	(1,2,4)	(1,2,5)	(1,2,6)	(1,3,4)	(1,3,5)	(1,3,6)	(1,4,5)	(1,4,6)	(1,5,6)
Probability	0.05	0.05	0.05	0.04	0.04	0.04	0.06	0.06	0.06	0.04
Sample	(2,3,4)	(2,3,5)	(2,3,6)	(2,4,5)	(2,4,6)	(2,5,6)	(3,4,5)	(3,4,6)	(3,5,6)	(4,5,6)
Probability	0.05	0.05	0.05	0.03	0.03	0.07	0.07	0.05	0.05	0.06

(i) Estimate the average consumption of food by using the HTE.

(ii) Obtain the standard error of the HTE using both the HTE and Yates–Grundy's variance estimators.

2.10.3 Show that for an SRSWOR sampling design of size n

(i) HTE for the population total is $\widehat{Y}_{ht} = N\,\overline{y}(s)$, where $\overline{y}(s) = \sum_{i\in s} y_i/n$.

(ii) $V_{ht} = V_{YG} = N^2(1/n - 1/N)S_y^2$,

where $S_y^2 = \sum_{i=1}^{N} \left(y_i - \overline{Y}\right)^2/(N-1)$

(iii) $\widehat{V}_{ht} = \widehat{V}_{YG} = N^2(1/n - 1/N)s_y^2$,

where $s_y^2 = \sum_{i\in s} \left(y_i - \overline{y}(s)\right)^2/(n-1)$

2.10.4

(a) Prove that an unbiased estimator of the population variance $S_y^2 = \sum_{i=1}^{N} \left(y_i - \overline{Y}\right)^2/(N-1)$ is available if and only if $\pi_{ij} > 0\ \forall\ i \neq j = 1,\ldots,N$.

(b) Show that $\widehat{S}_y^2 = a\sum_{i\in s} y_i^2/\pi_i + b\sum_{i\neq} \sum_{j\in s} y_i y_j/\pi_{ij}$ is an unbiased estimator for S_y^2, where $\pi_{ij} > 0\ \forall\ i \neq j$, $a = 1/N$, and $b = -1/N(N-1)$.

(c) Show that $\widehat{S}_y^2$ is admissible in the class of quadratic unbiased estimators of S_y^2.

2.10.5 Show that the estimator $e(s,y) = \sum_{i\in s} c_i y_i$ with $c_i \geq 1, \sum_{i\in s} 1/c_i = 1$ based on a fixed effective sampling design is admissible in the class of linear estimators (non necessarily unbiased) of Y (Godambe and Joshi, 1965).

2.10.6 Show that for a given sampling design (i) the sample mean based of distinct units and (ii) ratio estimator $\widehat{Y}_R = \dfrac{\sum_{i\in s} x_i}{\sum_{i\in s} y_i}\overline{X}$ (where $\overline{X}$ is known population mean of the charter x) are admissible in the class of all estimators for estimating the mean $\overline{Y}$ (Joshi, 1965a,b).