*CHAPTER 8*

**SAMPLING WITH UNEQUAL PROBABILITIES WITHOUT REPLACEMENT; USING HORVITZ – THOMPSON ESTIMATOR**

* 1. **INTRODUCTION**

In chapter 7 we have given a detailed description of unequal probability sampling with replacement. In chapters 8 and 9 we will explain unequal probability sampling without replacement. In Chapter 8 matter relating to Horvitz and Thompson estimator will be treated whereas in chapter 9 special estimators will be discussed. Sampling with unequal probabilities without replacement is not so simple as compared to with replacement procedure. For that let a sample of two units is selected from a population of N units. Let the probability of the selection of the ith unit is Pi = Zi/Z. Suppose that the ith unit is not selected at the first draw but the jth unit is selected (j ≠ i) then the probability of selecting the jth unit at the first draw is Pj = Zj/Z; and the conditional probability of selecting the ith unit at the second draw is Zi/(Z - Zj) = Pi/(1 - Pj). The probability of inclusion of ith unit at the second draw in the sample is the sum of the product that the jth unit is selected at the first draw and the ith unit is selected at the second draw, given the that jth unit is selected at the first draw i.e.

. (8.1.1)

The probability of inclusion of the ith population unit to be in the sample denoted by  will be; either the ith unit is included in the sample at the first draw or ith unit is included at the second draw i.e.

 . (8.1.2)

The probability that both ith and jth units are in the sample, denoted by πij ,will be

|  |  |  |
| --- | --- | --- |
| πij | = | Probability of the selection of the ith unit × probability of selection of jth unit given the ith unit has already been selected at the first draw + probability of the selection of jth unit × probability of the ith unit given jth unit has already been selected at the first draw |



. (8.1.3)

For any sample size similar expressions may be obtained but as the sample size increases the expressions become more complicated.

.

* 1. **NOTATIONS**

In this section some notations of unequal probability sampling without replacement has been explained.

 Probability of selection of the i-th population unit in the sample

 Probability of selection of the j-th unit given that i-th unit has already been selected

 Probability of inclusion of i-th population unit in the sample

 Joint probability of inclusion of the i-th and j-th population units in the sample

For a fixed sample size design properties of quantities  and  are stated in following theorem.

**Theorem 8.1:** The quantities  and  satisfy following relations:

i)  (8.2.1)

ii)  (8.2.2)

iii)  (8.2.3)

iv)  (8.2.4)

v)  (8.2.5)

**Proof**: To prove these relations we define an indicator variable ai that assumes a value of 1 if i-th unit is selected with probability  and  value of 0 with probability 1–. Clearly ai follows the binomial distribution with sample size 1

(There will be n values of ai and N-n values are zero)

Clearly; E(a­i) = πi; E(ai,aj) = πij; (8.2.6)

Var(ai) = πi (1-πi) ; Cov (ai,aj) = πij - πiπj. (8.2.7)

(i) 

We consider, .

Taking expectation of the above relation

Now 

(ii) 

From (8.2.6) we can write

(iii) 



Using this relation  we get



(iv) 



Using (8.2.2) we have,

◊

(v) 



* 1. **HORVITZ AND THOMPSON ESTIMATOR**

Horvitz and Thompson (1952) were the first to provide the general theory for unequal probability sampling without replacement; though Madow (1949), Narain (1951) and Goodman and Kish (1949) had already provided some concepts for unequal probability sampling without replacement but no mathematical foundation was developed.. The estimator of population total proposed by Horvitz and Thompson is:

 **(8.3.1)**

The Horvitz and Thompson (1952) estimator is unbiased over repeated sampling. If the inclusion probabilities are proportional to benchmark variable, it possesses the property described by Hajek (1981) as representative ness, which might better be termed as ratio estimator property (described in Chapter-6). An estimator that possesses this property is free from sampling error when the current (estimated) variable is exactly proportional to the auxiliary (benchmark) variable.

* + 1. **EXPECTATION AND VARIANCE OF HORVITZ AND THOMPSON ESTIMATOR**

In this section expectation, variance and variance estimator are be proved.

**Theorem 8.2:** The estimator proposed by Horvitz and Thompson (1952) is an unbiased estimator of population total *Y*.

*Proof:*

Using the indicator variable *ai* (as explained in chapter 2), (8.3.1) can be written as:

 (8.3.2)

In order to prove unbiasedness, we take expectation of (8.3.2)



Since E(ai) = πI; so

**◊**

**Theorem 8.3:** The variance of Horvitz and Thompson estimator is:

 (8.3.3)

or  (8.3.4)

**Proof:**

We have 

Substituting the value of  from (8.3.2):







= (8.3.5)

Substituting the values of  and  from (8.2.7) in (8.3.5) and on simplification we get

 ◊ (8.3.3)

Or alternatively let



 (8.3.5)

Substituting the values of  and  from (8.2.7) in (8.3.5) we have (8.3.3):

Substituting the value of  from (8.2.4) in first part of (8.3.3) we have:





On simplification we get:

 ◊ (8.3.4)

Expression (8.3.4) was put forwarded by Sen (1953) and independently by Yates and Grundy (1953) and is valid if the sample size is fixed. This is known as Sen-Yates-Grundy variance expression.

Expression (8.3.3) and (8.3.4) are valid provided πij are not equal to zero for j ≠ i, i.e. if all the possible pairs of distinct population units have non- zero probability of inclusion in the sample.

Robinson (1982) investigated the degree to which the Horvitz and Thompson estimator approximates the population mean in survey sampling. Specifically, conditions (for mean square), consistency were given, along with rates of convergence. The condition involved only first and second order inclusion probabilities, and for various sampling designs they are easy to check.

An unbiased variance estimator of (8.3.3) suggested by Horvitz and Thompson (1952) is:

. (8.3.7)

This estimator suffers from the disadvantage that it is not always zero when the variance is zero. The following alternative *conditionally* unbiased variance estimator was suggested by Sen (1953) and by Yates and Grundy (1953) for use when the number of sample units is fixed.

. (8.3.8)

This is known as Sen-Yates-Grundy variance estimator which is also convenient for use in an equal probability sampling context where the πij are known but unequal, example of this can be seen in Agrawal, Singh and Singh (1984). The unbiasedness of (8.3.7) and (8.3.8) may be proved easily. Both these estimators can assume negative values but (8.3.7) rarely seems to do so in practice. The estimator given in (8.3.8) has performed much better than (8.3.7) in a number of empirical comparisons, commencing with that in Yates and Grundy’s (1953) paper. Sen (1953) also compared the efficiency of (8.3.7) and (8.3.8) taking a population of five units and selecting all possible samples of two units under the following two schemes:

* The first unit is selected with pps and the second unit with pps without replacement.
* The first unit is drawn with pps and the second unit with equal probabilities without replacement.

Sen (1953) demonstrated that the expression ( 8.3.8) took positive values for all these samples, but that (8.3.7) was negative for some of the pairs. He further showed that j>j; for all j ≠ i for n=2.

Rao (1961, 1963) obtained same result under two well-known selection procedures for unequal probabilities without replacement. Rao and Sing (1973) compared (8.3.7) and (8.3.8) for 34 populations using Brewer’s selection procedure for n = 2 and came to the conclusion that (8.3.8) is always non-negative, whereas (8.3.7) is negative for some of the pairs. Moreover (8.3.8) has performed much better than (8.3.7). For n = 2 it is the only possible non-negative variance estimator [Vijayan, 1975]. Similar results were obtained by Lanke (1974) using Hajek’s Method -1(1964), and by Cumberland and Royall (1981) in a theoretical result. Liu(1984) proved that a fixed sample is essential for the *existence* of Sen-Yates-Grundy form of variance estimator and it is design unbiased estimator in the problem of estimating the mean, variance and covariance of a finite population. Rao (1979) has shown that for n = 2, (8.3.8) is the *unique unbiased* *variance estimator*. Recently Shahbaz (2003) has compared (8.3.7) and (8.3.8) for his method and found that (8.3.8) is always non–negative but (8.3.7) assume negative value for some of the samples. Joshi (1970) proved the *admissibility* of the Sen-Yates-Grundy variance estimator given in (8.3.7) and attempted to extend this proof to Murthy estimator (to be discussed in Chapter 9). Patel and Dharmadhikari (1978) disputed this extension and provided alternative proof. They also provided a proof of the admissibility of several estimators used with selection procedures such as Lahiri’s(1951) where the probability of selection of a sample is proportional to its aggregate measure of size, namely the (unbiased) classical (conventional) ratio estimator, an estimator given by Murthy (1963) and *linear invariant* estimator.

Das and Tripathi (1977) showed that Horvitz and Thompson variance estimator (8.3.7) was admissible in the class of unbiased estimators of variance (8.3.7). Biyani (1980) gave an example to show that there exist sample designs for any sample size greater than two for which the Sen-Yates-and Grundy estimator is *inadmissible* in the class of non-negative unbiased quadratic estimators of variance of the Horvitz and Thompson estimator. He also pointed out that a *posterior* lower bound can be obtained for any non-negative definite quadratic function of a finite population and gave an example to show that the Sen-Yates-Grundy variance estimator; even when is non-negative; can take values smaller than this bound.

A number of authors cited by Rao and Singh (1973) have written about necessary bestness and *hyper admissibility*, concepts which Rao and Singh showed to be meaningful only for samples ( of size >1) for which only single observation was non-zero. The earlier authors have argued or implied that because the Horvitz and Thompson estimator was necessary best and / or uniquely hyper-admissible in certain classes it should be used for any sample design, irrespective of whether there was any positive correlation between Yi and πi or not. This can be shown to lead to nonsensical results in some instances[ Rao (1966a), Basu(1971)] Rao and Singh(173) showed, however, that Horvitz and Thompson Variance estimator(8.3.7) was uniquely hyper-admissible in the class of all unbiased estimators of the variances(8.3.7) and further provided extensive empirical evidence that (8.3.7) was inferior to the Sen-Yates-Grundy variance estimator(8.3.8).

Lanke(1972), Patel(1974) Yogi and Gupta(1975) and Prabhu-Ajgaonkar (1984) have each presented proofs that necessary best estimator of order two does not exist for the most general class of linear homogeneous unbiased estimators.

Chaudhuri and Mukhopadhyay (1978) pointed out that the incrimination of sampling without replacement did not necessarily reduce the variance of the Horvitz and Thompson estimator. This result can also occur in equal probability sampling if πij are not all equal. Consider case where N=4, n=2,πi =1/2 for all i πij =1/2, if i=1 or 2 and j = 3 or 4 but zero otherwise. Then Y1 = Y2 =1 and Y3 = Y4 = 0 the variance is zero, but increasing the sample size to three by putting πi= ¾ for all i results in a non –zero variance regardless of the choice of the πij .

Singh and Singh (1979) considered the effect of random non-response both on sampling without replacement and *multinomial sampling* in each case treating the respondents as a simple random sub-sample of the original sample. Arnab (1979) in a addendum to their paper considered the same problems in terms of Poisson sampling ( Chapter 9)

The desirable properties of Horvitz and Thompson estimator are as;

1. It is the only unbiased estimator of the class in which same weight is attached to a particular population unit whenever it is selected (Horvitz and Thompson, 1952).
2. It is admissible in the class of all homogeneous linear unbiased estimators of population total Y, that is, there does not exist any member of that class which has smaller variance than  (Roy and Chakravarti, 1960, Godambe, 1960).
3. If the Yi are exactly proportional to πi and the number of units in the sample is fixed the variance of  is zero, this is a property usually associated with ratio estimator and will be referred to as the ratio estimator property. [Brewer (1963b)].
4. Under the model (6.8.2) the expected variance of the Horvitz-Thompson estimator achieves the lower bound of the expected variance for any design-unbiased estimator (Godambe-Joshi, 1965). i.e.



It follows from ratio estimator property if the values of the measure of size are known for all units in the population and Yi are approximately proportional to the measure of size Zi the variance of  can be made small by setting the proportional to. **This is a principal reason why selection with probability proportional to size has assumed importance in unequal probability sampling.**

* + 1. **Unequal Probability Sampling as a Unifying Thread For Sample Survey Theory**

Stuart (1963) demonstrated that sampling without replacement for unequal probability sampling generally, and the Horvitz–Thompson estimator in particular, could provide a framework for the understanding of number of sampling techniques commonly used in equal probability sampling, such as stratification, clustering and multi-stage sampling. His argument hinged on the interpretation of the Sen–Yates–Grundy form of the variance (8.3.7) as a weighted sum of squares  with weight.

Consider first simple random sampling without replacement. Here all the are  and all  are . Suppose we wish to hold  constant, so that the sampling is still with equal probabilities, but vary the  is such a fashion as to reduce the variance (8.3.4). The constraint on the  are that their sum is to be fixed at , that none be less than zero and that none exceed min. Expression (8.3.4) will be reduced if the corresponding at the largest values of  are increased. Suppose they are all increased to the value . Then the weights on these large squares are reduced to zero. Since the sum of the is fixed, the  corresponding to the remainder of the  must be reduced, with a consequent increase in the weights, but the net result is still a reduction in (8.3.4). But the setting of ,means that the ith and jth units must be selected independently. A way of achieving this is to arrange the population in groups such that if the ith and jth units are in the same group,  is likely to be small, while if they are in different groups,  is likely to be large. Selection of the sample within each group is then carried out independently of the selection within other groups. Thus , if the ith and jth units are from different groups and  if they are form the same group. These groups are familiar to the survey statistician as ‘strata’ and the selection procedure as ‘stratification with proportionate allocation’.

Stuart (1963) next suggests that the weights for pairs  corresponding to large values of  may be reduced still further by allowing the  for these pairs to take their maximum value of .

This implies that if the ith unit is selected, the jth is certain to be selected also. Such groups of units selected together are familiar to the survey statistician as ‘clusters’, and the selection procedure as ‘cluster sampling’. The clusters differ from the strata, naturally enough, in that the  for units within the same stratum should be as small as possible. The clustering principle is thus antithetic to the sampling principle: clusters should be heterogeneous internally whereas strata should be homogenous internally.

Since clustering changes the values of the  much more radically than stratification does, there is scope for a much greater reduction in variance if the clusters are chosen well. Conversely there is the danger of a great increase in variance if they are chosen poorly, or for reasons of convenience. If clusters are defined on the basis of geographical contiguity, in order to reduce survey travel costs, this danger is enhanced. Stuart highlights this difference by describing stratification as an investment and clustering as a speculation.

Fellegi (1963), in his discussion of Stuart’s paper, pointed out that while the setting of , i and j being in different groups makes these groups strata and the setting of , i and j being in the same group makes these groups clusters, the requirement where i belongs to group “s” and i belongs to group “t” may be only formally equivalent to two stage sampling, (chapter-5) and not specify a two–stage sampling design in fact.

**8.4 Comparison of with and without Replacement**

**for unequal Probability Sampling**

In this section we have given a comparison of unequal probability sampling with and without replacement. The comparison is given as under:

We know that

 (7.4.3)

Substituting πi = nPi in (7.4.3)

 (8.4.1)

Comparing (8.3.4) and (8.4.1) , if:

 (8.4.2)

Further comparison can also be made as:

 (7.4.1)

 (7.4.4)

and  (8.3.3)

 (8.4.3)

If ; then (8.4.3) can be written as:

 (8.4.4)

 (8.4.5)

From (7.4.4) and (8.4.5) we have

 (8.4.6)

This shows that  if

 (8.4.7)

* 1. **SELECTION PROCEDURES FOR USE WITH HORVITZ–THOMPSON (1952) ESTIMATOR**

We can see that the expressions of variance and variance estimator of Horvitz–Thompson (1952) estimator requires the knowledge of quantities  and  known as probability of inclusion of i-th unit and joint probability of inclusion of i-th and j-th units. Number of selection procedures is available for use with the Horvitz and Thompson (1952) estimator in literature to obtain expression for these quantities. The main purpose of designing a selection procedure is that  and  should be such that the variance should be as small as possible. The choice of a selection procedure to be used with Horvitz and Thompson (1952) estimator depends upon the characteristics that a selection procedure have. Some of the key features that a selection procedure must possess are listed below:

1. The number of units in the sample should be fixed.
2. The  should be precisely proportional to the measure of size.
3. Selection should be strictly without replacement.
4. There should be no difficulty in selecting a sample of more than two units i.e. should be applicable for any n.
5. The joint probabilities of selection should be simple to calculate as they are required for estimation of variance.
6. Each pair of distinct population units should have a non-zero probability of selection i.e. ≠ 0 for all j ≠ i.
7. The value of the should be such as to minimize the variance of the variance estimator.
8. The selection procedure should be simple.
9. For any value of n, the should take any value upto the theoretical limit n-1.
10. It should be simple to rotate. (Rotation means to drop some of the selected units and add new units in place of those).

Further, the available list of selection procedures differ from each other with respect to manner of selection of a sample. These selection procedures are divided in following groups by Carroll and Hartley (1964) depending upon their working.

1. **Draw by Draw Procedure**

At each draw one unit is selected from among those units not already selected. Probabilities of selection defined from each draw (since the selection is without replacement) are almost always depend on the units already selected. If the probability of selection at a given draw are (apart from a normalizing factor) independent of which units are selected at previous draw, these probabilities are sometimes called ***working probabilities.***

1. **Systematic Procedure**

Systematic procedure involves an ordering of population units or cumulating of size measures. The order of units may or may not be random. A random real number r (0 < r < 1) is chosen and n units selected are those whose cumulated values of πi (the desired probabilities of inclusion) are equal or to next greater than r, r + 1, r + 2, …… + r + n – 1.

1. **Rejective Procedure**

The term “rejective” has been employed by Hajek (1964) and is somewhat wider in its connotation than the term mass draw used by Carroll and Hartley (1964). Rejective procedure resembles draw by draw procedures in that only a single unit is selected at each draw of n successive draws. They differ from ordinary draw by draw procedures in that the selection at a given draw may give rise to the selection of an already selected unit, in such cases the selected sample is abandoned and the selection recommenced.

1. **Whole Sample Procedure**

In these selection procedures the units are not individually drawn: a probability is specified for each possible sample of n distinct units and one selection, using these probabilities, selects the whole sample.

1. **Sequential Procedures**

In these procedures each population unit is considered in turn and a decision is made whether or not it may be included in the sample.

**8.6 SELECTION PROCEDURES FOR SAMPLING WITH UNEQUAL**

**PROBABILITIES WITHOUT REPLACEMENT**

The Horvitz and Thompson (1952) estimator, its variance and variance estimators require the calculation of quantities  and . Since the emergence of this estimator a lot of selection procedures have been developed that can be used with this estimator. A comprehensive bibliography of these procedures can be found in Brewer and Hanif (1983), Chaudhuri and Vos (1988), and Hanif et al (1992). Some of these selection procedures are given below:

1. Midzuno’s Selection Procedure
2. Yates and Grundy (1953) Draw by Draw Procedure
3. Brewer’s (1963) Draw by Draw Procedure
4. Durbin’s (1967) Draw by Draw Procedure
5. Shahbaz–Hanif–Samiuddin (2003) Draw by Draw Procedure
6. Shahbaz–Hanif (2003) Draw by Draw Procedure
7. Yates and Grundy (1953) Rejective Procedure
8. Durbin’s (1953) Rejective Procedure
9. Rao’s (1965b) and Sampford’s (1967) Rejective Procedure
10. Prabhu and Ajgonkar (1982) Procedure
11. Hanif–Samiuddin–Shahbaz (2004) General Draw by Draw Procedure
12. Shahbaz–Khawaja–Hanif (2005) General class of draw by draw procedures.
13. Samiuddin and Asad (1981) Procedure
14. Random Systematic Procedure
15. **Midzuno’s Procedure (1950)**

In this selection procedure, selection is made as:

* Select first unit with probability proportional to measure of size, Zi
* Select a simple random sample of size n – 1 from remaining population.

The probability that ith unit is included in the sample is:

**(8.6.1)**

The joint probability of inclusion of both ith and jth units is:

 = P[i-th unit is selected at first draw and j-th unit is selected at second draw]

+ P[j-th unit is selected at first draw and i-th unit is selected at second draw]

+ P[Both units are selected after first two draws]



=   (8.6.2)

Similarly πijk : probability of inclusion of ith, jth and kth units is



 (8.6.3)

Under this scheme of sampling; Yates and Grundy variance estimator is always non-negative. The main advantage of this selection procedure is that it is simple to compute.

It is also interesting to note that the probability of getting a particular unordered sample is the sum of the probabilities of the units in the sample i.e.

 (8.6.4)

and with this property the classical (conventional) ratio estimator, defined in Chapter 6, is unbiased. This procedure breaks down unless  which is a very stringent requirement, consequently this procedure is not frequently applicable. Rao (1963) has shown that for n = 2, the variance of  with this procedure is always smaller than the variance of  provided  which is also condition for non-negativity of the working probabilities.

**EXAMPLE 8.1:** Consider following hypothetical population of size 4.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Unit** | 1 | 2 | 3 | 4 |
| **Yi** | 0.5 | 1.2 | 2.1 | 3.2 |
| **Zi** | 1 | 2 | 3 | 4 |
| **pi** | 0.1 | 0.2 | 0.3 | 0.4 |

Draw all possible samples of size 2 from this population using sampling without replacement. Show that the Horvitz and Thompson estimator is unbiased for population total under Midzuno’s selection procedure. Calculate the variance of this estimator.

# SOLUTION:

We first calculate the probabilities of selection and inclusion for given population under the Midzuno selection procedure. The expression for  under this selection procedure is given in (8.6.1). These values are given below:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Unit No.** | 1 | 2 | 3 | 4 | Total |
| **Yi** | 0.5 | 1.2 | 2.1 | 3.2 | 7.00 |
| **Zi** | 1 | 2 | 3 | 4 | 10.00 |
| **Pi** | 0.1 | 0.2 | 0.3 | 0.4 | 1.00 |
|  | 0.4000 | 0.4667 | 0.5333 | 0.6000 | 2.00 |

Now the Horvitz and Thompson estimator for a sample of size 2 is:



The expected value and variance of this estimator is obtained by using

 (8.6.5)

, (8.6.6)

where  are calculated by using (8.6.2). The necessary calculations to obtain these values are given in Table 8.1 below:

**Table 8.1**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Unit** | **yi** | **yj** | **pi** | **pj** | **πij** |  |  |  |
| 1 , 2 | 0.5 | 1.2 | 0.1 | 0.2 | 0.1000 | 3.8214 | 0.3821 | 1.4603 |
| 1 , 3 | 0.5 | 2.1 | 0.1 | 0.3 | 0.1333 | 5.1875 | 0.6917 | 3.5880 |
| 1 , 4 | 0.5 | 3.2 | 0.1 | 0.4 | 0.1667 | 6.5833 | 1.0972 | 7.2234 |
| 2 , 3 | 1.2 | 2.1 | 0.2 | 0.3 | 0.1667 | 6.5089 | 1.0848 | 7.0610 |
| 2 , 4 | 1.2 | 3.2 | 0.2 | 0.4 | 0.2000 | 7.9048 | 1.5810 | 12.4971 |
| 3 , 4 | 2.1 | 3.2 | 0.3 | 0.4 | 0.2333 | 9.2708 | 2.1632 | 20.0546 |
|  |  |  |  |  | 1.0000 |  | 7.0000 | 51.8844 |

[note that and  (sample size)]

Using (8.6.5) we have .

So Horvitz and Thompson estimator is unbiased under Midzuno selection procedure. The variance is obtained by using (8.6.6) as:



1. **Yates and Grundy (1953) Draw by Draw Procedure:**

This selection procedure is developed by Yates and Grundy (1953) for use with the Horvitz and Thompson (1952) estimator. The selection procedure is strictly without replacement and units are drawn one by one from the population. This selection procedure; for a sample of size 2; is described as under:

* + Select first unit with probability proportional to size.
  + Select second unit with probability proportional to size of the remaining units.

The probability of inclusion for ith unit in the sample is:



 (8.6.7)

The joint probability of inclusion for ith and jth unit in the sample is:

(8.6.8)

This selection procedure can be easily extended to larger sample sizes but calculation of  and  became very tedious. Sen (1953) has proved that under this procedure the Sen–Yates–Grundy (1953) variance estimator is always positive. For this consider the equation

, (8.6.9)

where .

For given values of *Pi* and *Pj*, the minimum value of *B* = *Bm* occurs when each  and in that case , where . The numerator of the expression on right side of (8.6.9) is than:



for , since .

**EXAMPLE 8.2:**

Consider the data given in Example 8.1.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Unit No.** | 1 | 2 | 3 | 4 |
| **Yi** | 0.5 | 1.2 | 2.1 | 3.2 |
| **Zi** | 1 | 2 | 3 | 4 |

Draw all possible samples of size 2 from this population. Show that the Horvitz and Thompson estimator is unbiased for population total under Yates and Grundy (1953) draw-by-draw selection procedure. Calculate the variance of this estimator.

# SOLUTION:

We first calculate the probabilities of inclusion for given population units under the Yates and Grundy selection procedure. The expression for  under this selection procedure is given in (8.6.7). These values are given below:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Unit No.** | 1 | 2 | 3 | 4 | **Total** |
| **Yi** | 0.5 | 1.2 | 2.1 | 3.2 | 7 |
| **Zi** | 1 | 2 | 3 | 4 | 10 |
| **Pi** | 0.1 | 0.2 | 0.3 | 0.4 | 1 |
|  | 0.2345 | 0.4413 | 0.6083 | 0.7159 | 2 |

Now the Horvitz and Thompson estimator for a sample of size 2 is:



The necessary calculations to obtain these values are given in the Table 8.2 below:

**Table 8.2**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Unit** | **yi** | **yj** | **pi** | **pj** | **πij** |  |  |  |
| 1 , 2 | 0.5 | 1.2 | 0.1 | 0.2 | 0.0472 | 4.8514 | 0.2291 | 1.1114 |
| 1 , 3 | 0.5 | 2.1 | 0.1 | 0.3 | 0.0762 | 5.5840 | 0.4255 | 2.3757 |
| 1 , 4 | 0.5 | 3.2 | 0.1 | 0.4 | 0.1111 | 6.6020 | 0.7336 | 4.8430 |
| 2 , 3 | 1.2 | 2.1 | 0.2 | 0.3 | 0.1607 | 6.1715 | 0.9918 | 6.1212 |
| 2 , 4 | 1.2 | 3.2 | 0.2 | 0.4 | 0.2333 | 7.1895 | 1.6775 | 12.0607 |
| 3 , 4 | 2.1 | 3.2 | 0.3 | 0.4 | 0.3714 | 7.9221 | 2.9425 | 23.3109 |
|  |  |  |  |  | 1.0000 |  | 7.0000 | 49.8229 |

.

So Horvitz and Thompson estimator is unbiased estimator of population total Y under Yates–Grundy selection procedure. The variance is obtained by using (8.6.6) as:



1. **Brewer’s Procedure (1963a)**

Brewer (1963a) suggested a selection procedure. The procedure for a sample of size 2 is given as:

* Select the first unit with probability proportional to 
* Select the second unit with probability proportional to size of remaining units.

The quantity  for this procedure is obtained as:

 

 (8.6.9)

Now the denominator may be simplified as

 (8.6.10)

where  . (8.6.11)

Substituting (8.6.10) in (8.6.9) we get 

The joint probabilities of inclusion of two units ;; may be obtained as



Since subscripts are dummy and *i* and *j* are symmetrical, therefore  is:

 (8.6.12)

It can be easily seen that (πiπj - πij) > 0 for all j ≠ i so that Yates and Grundy variance estimator is always positive. With this selection procedure, variance of  is always less than  [Brewer, 1963a]. Moreover, (πiπj - πij) is always positive [Rao, 1965], hence Yates-Grundy variance estimators is always positive. This method was extended for any n by the same author in 1975 through induction process.

**EXAMPLE 8.3**

For Brewer’s method with n = 2, we have



1. Show that if every pi < , 0 < πij < 4pipj for j ≠ i.
2. Show that this makes the Sen – Yates – Grundy variance estimator always positive for this selection procedure.

# SOLUTION

Consider

(a) (1 – pi – pj) = 1 – pi – p­j + 2pipj  - 2pipj.

= 1 – 2pi – p­j + 2pipj + pi - 2pipj.

= (1 – 2pi) – p­j (1 – 2pi) + pi  (1 – 2pj).

= (1 – 2pi) (1 – 2j) + pi (1 – 2pj).

= (1 – 2pi) (1 – 2pj)  = (1 – 2pi) (1 – 2pj) 

= (1 – 2pi) (1 – 2pj) 

Now



Substituting the value of (1 – pi – pj) from the above relation we have







since



so

⇒ 

⇒ 

(b)  (8.3.8)



since therefore



⇒ 

This shows that Yen-Yates-Grundy variance estimator is always positive if for all values Hence .

**(d) Durbin’s Procedure (1967)**

Durbin (1967) procedure is stated as:

* Select first unit with probability proportional to size
* Select second unit with probability proportional to 

The quantity is obtained as



 (8.6.14)

Now 

 (8.6.15)

Also  (8.6.16)

Substituting these values from (8.6.15) and (8.6.16) in (8.6.14) we have .

The quantity  turns out to be





Substituting value from (8.6.15) we have:

 **(8.6.17)**

which is the same as (8.6.12).

It can be easily shown that (πiπj - πij) > 0 hence Yates Grundy variance estimator is always positive. This procedure can be extended for n > 2 in principle.

**EXAMPLE 8.4:** Consider Example 8.1. Draw all possible samples and obtain the variance of Horvitz and Thompson estimator under Brewer selection procedure.

**SOLUTION:**

The probabilities of selection and inclusion for Brewer selection procedure are given below:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Unit No.** | 1 | 2 | 3 | 4 |
| **Yi** | 0.5 | 1.2 | 2.1 | 3.2 |
| **Zi** | 1 | 2 | 3 | 4 |
| **Pi** | 0.1 | 0.2 | 0.3 | 0.4 |
|  | 0.2 | 0.4 | 0.6 | 0.8 |

The necessary calculations to obtain the variance of Horvitz-Thompson (1952) estimators are given in Table 8.2 below:

**Table 8.3**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Unit** | **yi** | **yj** | **pi** | **pj** | **πij** |  |  |  |
| 1 , 2 | 0.5 | 1.2 | 0.1 | 0.2 | 0.0277 | 5.5000 | 0.1525 | 0.8386 |
| 1 , 3 | 0.5 | 2.1 | 0.1 | 0.3 | 0.0535 | 6.0000 | 0.3208 | 1.9248 |
| 1 , 4 | 0.5 | 3.2 | 0.1 | 0.4 | 0.1188 | 6.5000 | 0.7723 | 5.0198 |
| 2 , 3 | 1.2 | 2.1 | 0.2 | 0.3 | 0.1188 | 6.5000 | 0.7723 | 5.0198 |
| 2 , 4 | 1.2 | 3.2 | 0.2 | 0.4 | 0.2535 | 7.0000 | 1.7742 | 12.4198 |
| 3 , 4 | 2.1 | 3.2 | 0.3 | 0.4 | 0.4277 | 7.5000 | 3.2079 | 24.0594 |
|  |  |  |  |  | 1.0000 |  | 7.0000 | 49.2822 |

Using (8.6.5) we have 

So Horvitz and Thompson estimator is unbiased under Brewer selection procedure. The variance is obtained by using (8.6.6) as:



Since Durbin (1967) selection procedure is in same equivalence class as of Brewer (1963a) selection procedure therefore the variance of Horvitz and Thompson estimator will be same for Brewer and Durbin selection procedures.

**(e) Shahbaz–Hanif–Samiuddin Procedure (2003)**

This procedure is given by Shahbaz et al (2003) and is is stated as:

* Select the first unit with probability proportional to 
* Select the second unit with probability proportional to 

The probability of inclusion of ith unit in the sample is obtained as:



Substituting the values of  and  from (8.6.10) and (8.6.16) in above equation we have:



= 

=  (8.6.18)

The quantity  is worked out as under:



= 

=  (8.6.19)

**EXAMPLE 8.5:** Consider population of Example 8.1. Calculate variance of Horvitz and Thompson estimator under Shahbaz et al (2003) selection procedure.

**SOLUTION**:

The expression for  and  under Shahbaz–Hanif–Samiuddin procedure are given in (8.6.18) and (8.6.19). The values of  for population of Example 8.1 under this procedure are given below:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Unit No.** | 1 | 2 | 3 | 4 |
| **Yi** | 0.5 | 1.2 | 2.1 | 3.2 |
| **Zi** | 1 | 2 | 3 | 4 |
| **Pi** | 0.1 | 0.2 | 0.3 | 0.4 |
|  | 0.1692 | 0.3642 | 0.6064 | 0.8602 |

Necessary calculations to prove unbiasedness and to calculate variance of Horvitz and Thompson estimator are given in Table 8.3.

**Table 8.3**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Unit** | **yi** | **yj** | **pi** | **pj** | **πij** |  |  |  |
| 1 , 2 | 0.5 | 1.2 | 0.1 | 0.2 | 0.0162 | 6.2500 | 0.1012 | 0.6326 |
| 1 , 3 | 0.5 | 2.1 | 0.1 | 0.3 | 0.0365 | 6.4188 | 0.2345 | 1.5049 |
| 1 , 4 | 0.5 | 3.2 | 0.1 | 0.4 | 0.1165 | 6.6754 | 0.7774 | 5.1895 |
| 2 , 3 | 1.2 | 2.1 | 0.2 | 0.3 | 0.0871 | 6.7579 | 0.5883 | 3.9755 |
| 2 , 4 | 1.2 | 3.2 | 0.2 | 0.4 | 0.2610 | 7.0145 | 1.8307 | 12.8417 |
| 3 , 4 | 2.1 | 3.2 | 0.3 | 0.4 | 0.4828 | 7.1833 | 3.4679 | 24.9110 |
|  |  |  |  |  | 1.0000 |  | 7.0000 | 49.0552 |

Using results from last two columns in (8.4.7) and (8.4.8) we get:



So Horvitz and Thompson estimator is unbiased under Shahbaz–Hanif–Samiuddin selection procedure. The variance is obtained below:



**(f) Shahbaz–Hanif Procedure (2003)**

This procedure also uses revised probabilities and is described as:

* Select first unit with probability proportional to 
* Select second unit with probability proportional to size of the remaining units.

The probability of inclusion for ith unit in the sample is:

i = 

= 



=  (8.6.20)

The probability of inclusion of both ith and jth unit in the sample (ij) for this selection procedure is:

ij = 

= 

=  (8.6.21)

The Sen–Yates–Grundy variance expression is always positive under this selection procedure.

**EXAMPLE 8.6:** Consider population of Example 8.1. Obtain variance of Horvitz and Thompson estimator under Shahbaz–Hanif selection procedure.

**SOLUTION**:

The expression for  and  under Shahbaz–Hanif–Samiuddin procedure are given in (8.6.20) and (8.6.21). The values of  for population of Example 8.1 under this procedure are given below:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Unit No.** | 1 | 2 | 3 | 4 |
| **Yi** | 0.5 | 1.2 | 2.1 | 3.2 |
| **Zi** | 1 | 2 | 3 | 4 |
| **Pi** | 0.1 | 0.2 | 0.3 | 0.4 |
|  | 0.1892 | 0.3871 | 0.5974 | 0.8262 |

Necessary calculations to obtain the variance of Horvitz and Thompson estimator are given in Table 8.4 below:

**Table 8.4**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Unit** | **yi** | **yj** | **pi** | **pj** | **πij** |  |  |  |
| 1 , 2 | 0.5 | 1.2 | 0.1 | 0.2 | 0.0216 | 5.7418 | 0.1243 | 0.7136 |
| 1 , 3 | 0.5 | 2.1 | 0.1 | 0.3 | 0.0464 | 6.1574 | 0.2856 | 1.7585 |
| 1 , 4 | 0.5 | 3.2 | 0.1 | 0.4 | 0.1212 | 6.5152 | 0.7897 | 5.1452 |
| 2 , 3 | 1.2 | 2.1 | 0.2 | 0.3 | 0.1058 | 6.6149 | 0.6995 | 4.6274 |
| 2 , 4 | 1.2 | 3.2 | 0.2 | 0.4 | 0.2597 | 6.9727 | 1.8111 | 12.6283 |
| 3 , 4 | 2.1 | 3.2 | 0.3 | 0.4 | 0.4453 | 7.3883 | 3.2898 | 24.3057 |
|  |  |  |  |  | 1.0000 |  | 7.0000 | 49.1787 |

Using (8.4.8) we have:



**(g) Yates and Grundy Rejective Procedure (1953)**

This procedure was given by Yates and Grundy (1953) and is not a strictly without replacement procedure. For a sample of size 2 the selection is made as:

* + Select first unit with probability proportional to size and with replacement.
  + Select second unit with probability proportional to size.
  + Repeat first two steps if same unit is selected twice.

In order to obtain the probability of inclusion for ith unit in the sample one can readily see that the probability of cases when same unit can be selected twice is and the probability that the same unit cannot be selected twice is . Now the probability of inclusion of ith unit to be in the sample is:

 (8.6.22)

The joint probability of inclusion for ith and jth units is:

 (8.6.23)

**EXAMPLE 8.7:** Consider population of Example 8.1. Obtain the variance of Horvitz and Thompson estimator under Yates–Grundy rejective procedure.

**SOLUTION**:

The expression for  and  under Yates-Grundy procedure are given in (8.6.22) and (8.6.23). The values of  for Example 8.1 under this procedure are given below:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Unit No.** | 1 | 2 | 3 | 4 |
| **Yi** | 0.5 | 1.2 | 2.1 | 3.2 |
| **Zi** | 1 | 2 | 3 | 4 |
| **Pi** | 0.1 | 0.2 | 0.3 | 0.4 |
|  | 0.2571 | 0.4571 | 0.6000 | 0.6857 |

Necessary calculations to obtain the variance of Horvitz and Thompson estimator are given in Table 8.5.

**Table 8.5**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Unit** | **yi** | **yj** | **pi** | **pj** | **πij** |  |  |  |
| 1 , 2 | 0.5 | 1.2 | 0.1 | 0.2 | 0.0571 | 4.5694 | 0.2611 | 1.1931 |
| 1 , 3 | 0.5 | 2.1 | 0.1 | 0.3 | 0.0857 | 5.4444 | 0.4667 | 2.5407 |
| 1 , 4 | 0.5 | 3.2 | 0.1 | 0.4 | 0.1143 | 6.6111 | 0.7556 | 4.9951 |
| 2 , 3 | 1.2 | 2.1 | 0.2 | 0.3 | 0.1714 | 6.1250 | 1.0500 | 6.4313 |
| 2 , 4 | 1.2 | 3.2 | 0.2 | 0.4 | 0.2286 | 7.2917 | 1.6667 | 12.1528 |
| 3 , 4 | 2.1 | 3.2 | 0.3 | 0.4 | 0.3429 | 8.1667 | 2.8000 | 22.8667 |
|  |  |  |  |  | 1.0000 |  | 7.0000 | 50.1796 |

Using (8.4.8) we have:



**(h)Durbin Rejective Procedure (1953)**

The Durbin’s Rejective procedure is parallel to the Yates–Grundy (1953) procedure as it uses same methodology as used by the procedure of Yates and Grundy. The quantities  and  for this procedure are given in (8.6.22) and (8.6.23).

The difference in Yates–Grundy and Durbin procedure is that the former uses Horvitz–Thompson estimator and latter uses the estimator given as:

 (8.6.24)

Durbin (1953) has reported that estimator given in (8.4.23) is biased and the amount of bias is given as

 (8.6.25)

The bias is reported to be negligible.

**EXAMPLE 8.8:** Consider population given in Example 8.1. Obtain the variance of Horvitz and Thompson estimator under Durbin rejective procedure.

**SOLUTION**:

The expression for  and  under Durbin procedure are given in (8.6.22) and (8.6.23). The values of  for population of Example 8.1 under this procedure are given below:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Unit No.** | 1 | 2 | 3 | 4 |
| **Yi** | 0.5 | 1.2 | 2.1 | 3.2 |
| **Zi** | 1 | 2 | 3 | 4 |
| **Pi** | 0.1 | 0.2 | 0.3 | 0.4 |
|  | 0.2571 | 0.4571 | 0.6000 | 0.6857 |

Necessary calculations to obtain the variance of Horvitz and Thompson estimator are given in Table 8.6.

**Table 8.6**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Unit** | **yi** | **yj** | **pi** | **pj** | **πij** |  |  |  |
| 1 , 2 | 0.5 | 1.2 | 0.1 | 0.2 | 0.0571 | 5.5000 | 0.3143 | 1.7286 |
| 1 , 3 | 0.5 | 2.1 | 0.1 | 0.3 | 0.0857 | 6.0000 | 0.5143 | 3.0857 |
| 1 , 4 | 0.5 | 3.2 | 0.1 | 0.4 | 0.1143 | 6.5000 | 0.7429 | 4.8286 |
| 2 , 3 | 1.2 | 2.1 | 0.2 | 0.3 | 0.1714 | 6.5000 | 1.1143 | 7.2429 |
| 2 , 4 | 1.2 | 3.2 | 0.2 | 0.4 | 0.2286 | 7.0000 | 1.6000 | 11.2000 |
| 3 , 4 | 2.1 | 3.2 | 0.3 | 0.4 | 0.3429 | 7.5000 | 2.5714 | 19.2857 |
|  |  |  |  |  | 1.0000 |  | 6.8571 | 47.3714 |

Now 

Using (8.4.8) we have:



Finally 

**(i) Rao and Sampford Rejective Procedure (Rao 1965b, Sampford 1967)**

These selection procedure are rejective procedures and realize same set of  and . Both of these procedures use revised probabilities of selection. The procedure of Rao (1965b) for n = 2 is given as:

* Select first unit with probability proportional to  and with replacement.
* Select second unit with probability proportional to size.
* Repeat first two steps if same unit is selected twice.

The procedure of Sampford (1967) for n=2 is stated as:

* Select first unit with probability proportional to size and with replacement.
* Select second unit with probability proportional to .
* Repeat first two steps if same unit is selected twice.

The quantities  and  for Sampford (1967) procedure are obtained as under:

The probability of drawing two different units is:



The probability of inclusion for ith unit is obtained as:



=

 = .

Substituting the values from (8.6.10) and (8.6.11) in above equation we have .

The expression for can be worked as:



= 

Substituting the value of  from (8.4.10) in above equation we have:

 (8.6.26)

The quantity given in (8.4.26) is same as that for Brewer (1963a) and Durbin (1967) selection procedure.

For Brewer’s selection procedure, Durbin’s selection procedure and Sampford’s selection procedure, the quantities and  is same, therefore one can say that for n = 2, they belong to same equivalence class and also know as **BDS** selection procedure.

**(j) Prabu and Ajgonkar Procedure (1982)**

This procedure is like the rejective procedure but differ in that in this procedure we do not reject the whole sample but only reject the unit that is selected twice. This procedure is:

* Select first unit with probability proportional to size and with replacement.
* Select second element with probability  where A is a normalizing constant.
* If same unit is selected twice, then select one more unit, from remainder of the population, with probability proportional to size.

The probability of inclusion of ith unit in the sample for this procedure is:



=  (8.6.27)

The joint probability of inclusion of both ith and jth unit in the sample for this procedure is:

 (8.6.28)

1. **Hanif–Samiuddin–Shahbaz General Procedure (2004)**

This procedure is a generalization of Brewer (1963) procedure for a sample of size 2. This procedure also uses the revised probabilities of selection at the first draw. The procedure is:

* Select first unit with probability proportional to  where “*a*” is a constant.
* Select second unit with probability proportional to size of the remaining units.

The probability of inclusion for the i-th unit; i; in the sample is:



= 

=  (8.6.29)

The joint probability of inclusion for i-th and j-th unit in the sample for this selection procedure is given as:



= 

=  (8.6.30)

Some special cases of this general selection procedure can be readily obtained by using selected values of constant “*a*”. Using *a* = 0.0 in (8.4.29) the probability of inclusion  becomes:

 (8.6.22)

Also for *a* = 0.0 the joint probability of inclusion  given in (8.4.30) transforms to:

 (8.4.23)

So for *a* = 0.0 this procedure is equivalent to the Yates–Grundy (1953) rejective procedure in that this procedure realizes same set of  and .

Again using *a* = 0.5 in (8.6.30) the probability of inclusion  becomes:

 (8.6.5)

For *a* = 0.5 the joint probability of inclusion given in (8.6.30) becomes:

 (8.6.6)

Again for *a* = 1.0 the expressions given in (8.6.29) and (8.6.30) transform to:

 and 

which are probabilities of inclusion and joint probability of inclusion for Brewer (1963) selection procedure

This shows that the general selection procedure transforms to the Yates–Grundy (1953) draw-by-draw procedure for *a* = 0.5. It can also be seen that for *a* = 1.0 the general selection procedure transforms to Brewer (1963) selection procedure as for this value of *a* the quantities  and  given in (8.6.29) and (8.6.30) transforms to the expression of Brewer (1963) selection procedure.

1. **Random Systematic Procedure (Goodman and Kish 1949)**

This procedure is stated as:

Arrange the population units in random order. Cumulate the measure of size, divide the total measure of size Z by the required number of units in sample, n, to obtain the skip interval Z/n. Choose a random start that is a random number greater than or equal to zero and less than the skip interval. The first unit selected is that for which the cumulate size measure is the smallest greater than or equal to the random start the second unit is that for which the cumulated size measure is the smallest greater than or equal to the random start plus the skip interval.

For this type of selection procedure Hartley and Rao (1962) have given a formula for π*ii*, which is asymptotically correct as *N →*  ∞ under certain conditions.

The main drawbacks of the systematic procedures are the difficulty of calculating the joint probabilities of inclusion for the purpose of estimating the variance, and the fact that one or more of these joint probabilities is sometimes zero. A simple example of a situation in which one of the π*IJ* is zero is given by *n* = 2 ; *N* = 5 ; *Zj* = 1, 2, 4, 5, 6.

An illustration of this procedure is given below:

Suppose we have a population of 8 units arranged in random order with the sizes. Let a sample of 3 units is to be selected

|  |  |  |
| --- | --- | --- |
| Unit | Size | Cumulative Total |
| 1 | 15 | 15 |
| 2 | 81 | 96 |
| 3 | 26 | 112 |
| 4 | 42 | 164 |
| 5 | 20 | 184 |
| 6 | 16 | 200 |
| 7 | 45 | 245 |
| 8 | 55 | 300 |

In this example Z/n = 300/3 = 100. Let the random start is 36, so the second unit is in the sample. For the selection of second unit Z/n = 100 will be added in 36 and comes out to be 36 + 100 = 135; this falls against 164 so ‘th unit is selected. Similarly for the selection of third unit 2Z/n will be added in 36 and so on till the required sample is obtained. The selection procedure is simple. The only disadvantage with this is that no exact formula for variance and variance estimator is available. Hartley and Rao (1962) have derived an expression for the joint probabilities πij which is asymptotically correct under certain conditions as . Using these values; Hartley and Rao (1962) obtained following asymptotic formula for variance of :

 **(8.6.31)**

 **(8.6.32)**

It was shown that for n = 2

 **(8.6.33)**

and the variance estimator of the same order is

 **(8.6.34)**

Note that .

Both these expression reduces to the variance expression of simple random sampling without replacement for .

If in (8.6.31) the term  is deleted it transforms to (7.4.2) which is variance estimator. The factor mentioned here can be treated as correction factors. Hence it is obvious that  for Hartley-Rao scheme is less than . Connor (1966) has derived an exact formula for πij for any n and J.N.K. Rao (1965) derived asymptotic expression for πij for various selection procedures. Expression (8.6.30) is valid for all the selection procedures using Horvitz and Thompson estimator provided  (Hanif 1974).

# EXAMPLE 8.9: From the population given in Example 3 and for n = 2; find the variance of under random systematic selection procedure.

# SOLUTION:

Since πi = nPi

 (8.6.35**)**

Substituting the value of Pi and Yi we get:



**(m) Samiuddin – Asad (1981) Procedure:**

In this selection procedure the population of size “N” is divided in (n+1) groups or blocks, let us call these blocks as R and T. To select a random sample of size “n” by using this selection procedure we obtain quantities  where . The blocks in this method are formed so that  and  for all R and T. The procedure works well when all  are nearly equal to . After forming the (n+1) blocks a random sample of “n” blocks is selected, so that probability of not selecting block R is , and then one unit is randomly selected from each block. Hence the probability of selecting a unit from block R is  and the probability of inclusion of I-th unit in the sample is . The joint probability of inclusion of I-th and J-th units in the sample can be easily obtained as under:

The probability of including block I and J in the sample is  and hence the probability of including both units I and J in the sample is  Also if  then  Also the variance of Horvitz–Thompson estimator, under this selection procedure, takes the form:

 (8.6.36)

where  and  This selection procedure enjoys the property that 

1. **Khawaja–Shahbaz–Hanif (2005) general class of draw by draw procedures:**

In this selection procedure:

• Select first unit with probability proportional to 

* Select second unit with probability proportional to size of the remaining units.

The probability of inclusion of i-th unit in this selection procedure is given as

i = 

which after slight algebraic manipulation becomes

i =  (8.3.37)

where  (8.3.38)

The joint probability of inclusion of both i-th and j-th unit in the sample for this selection procedure is given by



which after slight algebraic manipulation becomes

 (8.3.39)

This selection procedure gives number of other selection procedures as its members for selected values of the constants “a” and “b”.

**8.7 EMPIRICAL STUDY OF SELECTION PROCEDURES**

Some selected selection procedures for use with the Horvitz–Thompson (1952) estimator were studied in previous section along with a general selection procedure. This empirical study is two fold in that first the empirical study of selected selection procedures is given then empirical study for the general selection procedure is also given. The empirical study is based on fifty natural populations, selected from standard text on the subject. To carry out the empirical study the variance of Horvitz–Thompson estimator is calculated for various selection procedures along with variance of Hansen–Hurwitz (1943) estimator with equal probabilities. After calculating the variance for various procedures for each population they have been ranked in ascending order of magnitude. The procedure that yields smallest variance in a population is assigned a rank 1, the second smallest is assigned a rank 2 and so on. The results of these rankings are given in Table 8.4.

**Table 8.7: Ranks of Various Estimators with Ranks of Coefficient of Variation of X and Correlation Coefficient between X and Y**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Pop.** | **CV (X)** | **Rho** | **SRS** | **HH** | **MS** | **YG(dbd)** | **Brewer** | **SHS** | **Dur (Rej)** | **YG (Rej)** |
| 1 | 20 | 7 | 3 | 8 | 1 | 4 | 5 | 7 | 6 | 2 |
| 2 | 46 | 49 | 8 | 5 | 1 | 3 | 4 | 6 | 7 | 2 |
| 3 | 38 | 35 | 8 | 5 | 1 | 3 | 4 | 7 | 6 | 2 |
| 4 | 10 | 41 | 8 | 4 | 7 | 5 | 2 | 1 | 3 | 6 |
| 5 | 9 | 10 | 8 | 7 | 2 | 3 | 4 | 6 | 5 | 1 |
| 6 | 45 | 50 | 8 | 2 | 7 | 3 | 1 | 5 | 4 | 6 |
| 7 | 44 | 44 | 8 | 4 | 7 | 1 | 2 | 5 | 3 | 6 |
| 8 | 4 | 9 | 8 | 6 | 7 | 5 | 4 | 2 | 3 | 1 |
| 9 | 12 | 21 | 8 | 6 | 7 | 5 | 4 | 1 | 2 | 3 |
| 10 | 33 | 23 | 2 | 7 | 1 | 4 | 5 | 8 | 6 | 3 |
| 11 | 39 | 25 | 8 | 5 | 7 | 2 | 3 | 6 | 4 | 1 |
| 12 | 15 | 28 | 8 | 6 | 7 | 5 | 3 | 1 | 2 | 4 |
| 13 | 48 | 29 | 8 | 6 | 7 | 2 | 3 | 5 | 4 | 1 |
| 14 | 34 | 45 | 8 | 5 | 7 | 2 | 3 | 6 | 4 | 1 |
| 15 | 37 | 40 | 8 | 6 | 7 | 2 | 3 | 5 | 4 | 1 |
| 16 | 19 | 27 | 8 | 6 | 7 | 5 | 2 | 1 | 3 | 4 |
| 17 | 25 | 12 | 8 | 7 | 2 | 3 | 4 | 6 | 5 | 1 |
| 18 | 40 | 32 | 2 | 6 | 1 | 4 | 5 | 8 | 7 | 3 |
| 19 | 43 | 13 | 2 | 7 | 1 | 4 | 5 | 8 | 6 | 3 |
| 20 | 22 | 16 | 8 | 6 | 7 | 5 | 3 | 1 | 2 | 4 |
| 21 | 29 | 33 | 8 | 6 | 7 | 3 | 4 | 5 | 2 | 1 |
| 22 | 23 | 11 | 8 | 7 | 1 | 3 | 4 | 6 | 5 | 2 |
| 23 | 18 | 14 | 8 | 7 | 1 | 3 | 4 | 6 | 5 | 2 |
| 24 | 50 | 4 | 1 | 7 | 2 | 4 | 5 | 6 | 8 | 3 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Pop.** | **CV (X)** | **Rho** | **SRS** | **HH** | **MS** | **YG(dbd)** | **Brewer** | **SHS** | **Dur (Rej)** | **YG (Rej)** |
| 25 | 31 | 30 | 8 | 6 | 7 | 2 | 3 | 5 | 4 | 1 |
| 26 | 36 | 26 | 8 | 6 | 7 | 2 | 3 | 4 | 5 | 1 |
| 27 | 21 | 8 | 8 | 7 | 6 | 2 | 3 | 5 | 4 | 1 |
| 28 | 35 | 2 | 2 | 6 | 1 | 4 | 5 | 7 | 8 | 3 |
| 29 | 16 | 20 | 8 | 7 | 2 | 3 | 4 | 6 | 5 | 1 |
| 30 | 17 | 22 | 8 | 6 | 7 | 5 | 3 | 2 | 4 | 1 |
| 31 | 13 | 1 | 1 | 8 | 2 | 4 | 5 | 7 | 6 | 3 |
| 32 | 5 | 3 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 |
| 33 | 11 | 24 | 8 | 6 | 7 | 4 | 3 | 2 | 5 | 1 |
| 34 | 14 | 37 | 2 | 7 | 1 | 4 | 5 | 8 | 6 | 3 |
| 35 | 27 | 18 | 8 | 7 | 3 | 2 | 4 | 6 | 5 | 1 |
| 36 | 30 | 42 | 8 | 6 | 7 | 2 | 1 | 4 | 3 | 5 |
| 37 | 41 | 38 | 8 | 5 | 7 | 2 | 3 | 6 | 4 | 1 |
| 38 | 1 | 19 | 7 | 6 | 1 | 2 | 4 | 5 | 3 | 8 |
| 39 | 24 | 15 | 8 | 6 | 7 | 4 | 3 | 2 | 5 | 1 |
| 40 | 28 | 17 | 8 | 7 | 1 | 3 | 4 | 6 | 5 | 2 |
| 41 | 3 | 5 | 8 | 7 | 6 | 2 | 3 | 5 | 4 | 1 |
| 42 | 47 | 48 | 8 | 4 | 7 | 5 | 2 | 1 | 3 | 6 |
| 43 | 49 | 47 | 8 | 3 | 7 | 1 | 2 | 4 | 5 | 6 |
| 44 | 7 | 39 | 8 | 5 | 7 | 2 | 4 | 6 | 3 | 1 |
| 45 | 26 | 34 | 8 | 4 | 7 | 2 | 3 | 6 | 5 | 1 |
| 46 | 8 | 6 | 8 | 7 | 5 | 2 | 4 | 6 | 3 | 1 |
| 47 | 6 | 46 | 8 | 5 | 7 | 1 | 2 | 6 | 3 | 4 |
| 48 | 32 | 43 | 8 | 5 | 7 | 4 | 2 | 1 | 3 | 6 |
| 49 | 42 | 31 | 8 | 6 | 7 | 2 | 3 | 5 | 4 | 1 |
| 50 | 2 | 36 | 8 | 5 | 6 | 4 | 2 | 1 | 3 | 7 |

Table 8.5 below contains the frequency of ranks of various selection procedures

**Table 8.8: Frequency of Ranks of Various Procedures**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Rank** | **SRS** | **HH** | **MS** | **YG(dbd)** | **Brewer** | **SHS** | **Dur (Rej)** | **YG (Rej)** |
| 1 | 2 | 0 | 12 | 3 | 2 | 8 | 0 | 23 |
| 2 | 5 | 1 | 5 | 16 | 8 | 4 | 5 | 6 |
| 3 | 1 | 1 | 1 | 10 | 16 | 1 | 12 | 8 |
| 4 | 0 | 4 | 0 | 12 | 16 | 3 | 11 | 4 |
| 5 | 0 | 9 | 1 | 9 | 8 | 10 | 12 | 1 |
| 6 | 0 | 18 | 4 | 0 | 0 | 16 | 6 | 6 |
| 7 | 1 | 15 | 27 | 0 | 0 | 4 | 2 | 1 |
| 8 | 41 | 2 | 0 | 0 | 0 | 4 | 2 | 1 |
| **Average Rank** | 7.00 | 5.90 | 4.86 | 3.16 | 3.40 | 4.74 | 4.32 | 2.62 |

From Table 8.5 it can be seen that the Yates–Grundy (1953) rejective procedure dominates other selection procedures included in the study as this procedure has smallest average rank. This procedure is closely followed by the Yates–Grundy (1953) draw-by-draw procedure. Table 8.5 also shows that the Yates–Grundy (1953) rejective procedure perform better than all other procedures in 23 out of 50 populations. The Brewer (1963) procedure performs better in only 2 populations. Shahbaz–Hanif–Samiuddin (2003) procedure is better than all other procedures in 8 out of fifty populations.

The performance of a selection procedure also depends upon coefficient of variation of measure of size variable and correlation coefficient between measure of size variable and variable under study (Y). To see the effect of these measures on performance of a selection procedure the ranking can be categorized with respect to various ranges of coefficient of variation and correlation coefficient as given in Table 8.6 and Table 8.7 below:

**Table 8.9: Average Ranks of Various Procedures for Different Ranges**

**of Coefficient of Variation**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Rank**  **CV (Z)** | **SRS** | **HH** | **MS** | **YG(dbd)** | **Brewer** | **SHS** | **Dur (Rej)** | **YG (Rej)** |
| 1 – 10 | 7.9 | 5.9 | 5.4 | 3.1 | 3.3 | 4.1 | 3.2 | 3.1 |
| 11 – 20 | 6.2 | 6.7 | 4.2 | 4.2 | 3.8 | 4.1 | 4.4 | 2.4 |
| 21 – 30 | 8.0 | 6.3 | 4.8 | 2.9 | 3.3 | 4.7 | 4.1 | 1.9 |
| 31 – 40 | 6.2 | 5.7 | 4.6 | 2.9 | 3.6 | 5.7 | 5.1 | 2.2 |
| 41 – 50 | 6.7 | 4.9 | 5.3 | 2.7 | 3.0 | 5.1 | 4.8 | 3.5 |

From results of Table 8.6 it can be seen that for small and moderate ranges of coefficient of variation Yates–Grundy (1953) rejective procedure outperform other procedures. For large values of coefficient of variation Yates–Grundy (1953) draw-by-draw procedure outperform all other procedures. The performance of Brewer (1963a) draw-by-draw procedure is also reasonable well. Table 8.6 clearly indicates that this procedure shows least variation in its performance as its minimum average rank is 3.0 and maximum is 3.8.

**Table 8.10: Average Ranks of Various Procedures for Different Ranges**

**of Correlation Coefficient**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Rank**  **ρYZ** | **SRS** | **HH** | **MS** | **YG(dbd)** | **Brewer** | **SHS** | **Dur (Rej)** | **YG (Rej)** |
| 1 – 10 | 5.5 | 7.0 | 3.8 | 3.5 | 4.2 | 5.4 | 4.9 | 1.7 |
| 11 – 20 | 7.3 | 6.7 | 2.6 | 3.2 | 3.9 | 5.2 | 4.6 | 2.5 |
| 21 – 30 | 7.4 | 6.0 | 6.4 | 3.6 | 3.2 | 3.5 | 3.9 | 2.0 |
| 31 – 40 | 6.8 | 5.5 | 5.1 | 2.8 | 3.6 | 5.7 | 4.4 | 2.1 |
| 41 – 50 | 8.0 | 4.3 | 6.4 | 2.7 | 2.1 | 3.9 | 3.8 | 4.8 |

The performance of various selection procedures for various ranges of correlation coefficient is given in Table 8.7. This table clearly shows that for small and moderate correlation coefficient Yates–Grundy (1953) rejective procedure outperforms other procedures whereas for large correlation coefficient Brewer (1963a) procedure is best. The performance of Hansen–Hurwitz estimator is not good as it is a with replacement procedure. A general condition in which sampling with replacement performs better than sampling without replacement in unequal probability sampling is given in section 8.6.

The effect of coefficient of variation and correlation coefficient on performance of a selection procedure can also be judged by conducting formal regression analysis. Using the rankings given in Table 8.4 the regression model

Rank (Estimator) = β0 + β1 (Rank CV(X)) + β2(Rank ρXY) + ε

can be easily fitted. Summary results for above model are given in Table 8.8 below:

**Table 8.8: Regression Summaries for Ranks of Various Estimators**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **SRS** | **HH** | **MS** | **YG dbd** | **Brewer** | **SHS** | **Dur rej** | **YG rej** |
| **β0** | 6.491 | 7.665 | 3.482 | 3.975 | 4.344 | 4.455 | 3.999 | 1.532 |
| **p-value** | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.016 |
| **β1** | -0.052 | -0.005 | -0.037 | -0.013 | 0.014 | 0.056 | 0.060 | -0.019 |
| **p-value** | 0.021 | 0.513 | 0.153 | 0.293 | 0.106 | 0.013 | 0.000 | 0.336 |
| **β2** | 0.072 | -0.064 | 0.091 | -0.018 | -0.052 | -0.045 | -0.048 | 0.062 |
| **p-value** | 0.002 | 0.000 | 0.001 | 0.152 | 0.000 | 0.045 | 0.001 | 0.003 |
| **F** | 6.157 | 39.628 | 6.303 | 2.595 | 18.343 | 3.983 | 11.817 | 5.007 |
| **p-value** | 0.004 | 0.000 | 0.004 | 0.085 | 0.000 | 0.025 | 0.000 | 0.011 |

Table 8.8 clearly indicates effect of coefficient of variation and correlation coefficient on rank of a selection procedure. The **β1** coefficients indicate effect of coefficient of variation on rank of a procedure. These coefficients indicate that the rank of Brewer (1963a), Shahbaz–Hanif–Samiuddin (2003) and Durbin (1953) rejective procedures will increase with increase in coefficient of variation of measure of size. So these procedures are useful for populations with small coefficient of variation. Effect of coefficient of variation on ranks of other procedures involved in the study is inverse; that is these procedures will perform better for population having larger coefficient of variation. The **β2** coefficients show effect of correlation coefficient on rank of a selection procedure. This coefficient is positive for three procedures namely Simple Random Sampling, Midzuno–Sen (1951) and Yates–Grundy (1953) rejective procedures indicating that the performance of these procedures will be better for population having smaller correlation coefficient between measure of size and variable under study. For rest of the procedures this coefficient is negative indicating that these procedures will perform better for population having large correlation coefficient. The F–values for testing significance of regression indicates that the performance of a procedure can be judged by using the coefficient of variation of measure of size and correlation coefficient between measure of size and actual variable under study.

Hanif–Samiuddin–Shahbaz (2004) selection procedure is also given in section 8.4 (k). This procedure uses a constant “*a*” in working probabilities. The choice of a suitable value of “*a*” in this procedure depends upon the performance of this procedure for various values of “*a*”. To decide about this the empirical study of this selection procedure is also given below. For this empirical study same fifty natural populations have been used. The empirical study is also based upon ranks. The frequency of rank for various values of *“a*” is given in Table 8.9 below:

**Table 8.9: Frequency of Ranks of Various Values of “*a*”**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Ranks** | **Values of “*a*”** | | | | | | | | | | |
| **0.0** | **0.1** | **0.2** | **0.3** | **0.4** | **0.5** | **0.6** | **0.7** | **0.8** | **0.9** | **1.0** |
| 1 | 28 | 0 | 0 | 0 | 4 | 0 | 0 | 2 | 0 | 2 | 14 |
| 2 | 0 | 28 | 0 | 2 | 0 | 2 | 1 | 0 | 1 | 14 | 2 |
| 3 | 0 | 0 | 28 | 2 | 0 | 2 | 1 | 0 | 17 | 0 | 0 |
| 4 | 0 | 0 | 2 | 28 | 0 | 1 | 2 | 16 | 0 | 1 | 0 |
| 5 | 0 | 0 | 2 | 0 | 28 | 1 | 18 | 0 | 0 | 1 | 0 |
| 6 | 0 | 2 | 0 | 0 | 1 | 44 | 0 | 2 | 0 | 0 | 1 |
| 7 | 0 | 2 | 0 | 1 | 17 | 0 | 28 | 2 | 0 | 0 | 0 |
| 8 | 3 | 0 | 0 | 17 | 0 | 0 | 0 | 28 | 1 | 0 | 1 |
| 9 | 1 | 0 | 18 | 0 | 0 | 0 | 0 | 0 | 31 | 0 | 0 |
| 10 | 0 | 18 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 32 | 0 |
| 11 | 18 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 32 |
| **Average Ranks** | 5.18 | 5.24 | 5.28 | 5.30 | 5.38 | 5.66 | 5.98 | 6.32 | 6.8 | 7.18 | 7.68 |

It can be readily seen that the performance of general selection procedure is best at *a*=0.0; which yield the Yates–Grundy (1953) rejective procedure. The performance of general selection procedure for various ranges of coefficient of variation and correlation coefficient is given in following tables.

**Table 8.10: Average Ranks of Various Values of “a” with ranks of Coefficient of Variation.**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **CV** | **Values of Constant “a”** | | | | | | | | | | |
| **0.0** | **0.1** | **0.2** | **0.3** | **0.4** | **0.5** | **0.6** | **0.7** | **0.8** | **0.9** | **1.0** |
| 1 – 10 | 5.0 | 5.2 | 5.4 | 5.6 | 5.8 | 6.0 | 6.2 | 6.4 | 6.6 | 6.8 | 7.0 |
| 11 – 20 | 6.0 | 6.0 | 6.0 | 6.0 | 6.0 | 6.0 | 6.0 | 6.0 | 6.0 | 6.0 | 6.0 |
| 21 – 30 | 4.1 | 4.2 | 4.2 | 4.5 | 4.9 | 5.6 | 6.3 | 7.0 | 7.8 | 8.4 | 9.0 |
| 31 – 40 | 2.9 | 3.4 | 3.9 | 4.5 | 5.0 | 5.5 | 6.3 | 7.2 | 8.2 | 9.1 | 10.0 |
| 41 – 50 | 3.1 | 3.2 | 3.5 | 4.2 | 4.8 | 5.6 | 6.5 | 7.4 | 8.3 | 9.5 | 9.9 |

**Table 8.11: Average Ranks of Values of “a” with ranks of Correlation Coefficient.**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **XY** | **Values of Constant “a”** | | | | | | | | | | |
| **0.0** | **0.1** | **0.2** | **0.3** | **0.4** | **0.5** | **0.6** | **0.7** | **0.8** | **0.9** | **1.0** |
| 1 – 10 | 3.0 | 3.6 | 4.2 | 4.8 | 5.4 | 6.0 | 6.6 | 7.2 | 7.8 | 8.4 | 9.0 |
| 11 – 20 | 3.0 | 3.6 | 4.2 | 4.8 | 5.4 | 6.0 | 6.6 | 7.2 | 7.8 | 8.4 | 9.0 |
| 21 – 30 | 6.0 | 6.0 | 6.0 | 6.0 | 6.0 | 6.0 | 6.0 | 6.0 | 6.0 | 6.0 | 6.0 |
| 31 – 40 | 3.3 | 3.5 | 3.7 | 4.3 | 4.9 | 5.5 | 6.3 | 7.2 | 8.2 | 9.1 | 10.0 |
| 41 – 50 | 5.8 | 5.3 | 4.9 | 4.9 | 4.8 | 5.2 | 5.8 | 6.4 | 7.1 | 7.9 | 7.9 |

Tables 8.10 and 8.11 clearly indicate that the general selection procedure performs well at *a*=0.0 which yield the Yates–Grundy (1953) rejective procedure.

Regression summary for the general selection procedure is given in following table:

**Table 8.12: Regression Summary for Ranks of Various Values of “a”**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Values of Constant “a”** | | | | | | | | | | |
| **0.0** | **0.1** | **0.2** | **0.3** | **0.4** | **0.5** | **0.6** | **0.7** | **0.8** | **0.9** | **1.0** |
| **Β0** | 4.684 | 4.781 | 4.908 | 5.080 | 5.322 | 5.864 | 6.377 | 6.844 | 7.147 | 7.471 | 7.522 |
| **p-Value** | 0.003 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| **Β1** | -0.119 | -0.090 | -0.060 | -0.032 | -0.007 | 0.006 | 0.021 | 0.037 | 0.057 | 0.080 | 0.107 |
| **p-Value** | 0.014 | 0.020 | 0.040 | 0.143 | 0.697 | 0.549 | 0.105 | 0.093 | 0.062 | 0.044 | 0.026 |
| **Β2** | 0.139 | 0.108 | 0.075 | 0.040 | 0.009 | -0.014 | -0.037 | -0.057 | -0.070 | -0.092 | -0.101 |
| **p-Value** | 0.005 | 0.006 | 0.012 | 0.064 | 0.602 | 0.182 | 0.006 | 0.010 | 0.022 | 0.022 | 0.036 |
| **F** | 5.567 | 5.151 | 4.156 | 2.153 | 0.159 | 0.920 | 4.346 | 3.854 | 3.400 | 3.600 | 3.587 |
| **p-Value** | 0.007 | 0.009 | 0.022 | 0.130 | 0.853 | 0.405 | 0.019 | 0.028 | 0.042 | 0.035 | 0.035 |

The coefficients **β0** in **a**bove table clearly indicates that the average rank of general selection procedure will increase with increase in the value of *a*. Also the coefficients **β1** is negative for *a*=0.0 to *a*=0.4 indicating that for large coefficient of variation small value of *a* will give more precise results. Also the coefficient **β2** for these values of *a* is positive indicating that for small correlation coefficient small values of *a* should be a suitable choice. The large value of *a* is appropriate for populations having small coefficient of variation and large correlation coefficient.

* 1. **APPROXIMATE VARIANCE FORMULAS OF HORVITZ–THOMPSON ESTIMATOR USING FIRST ORDER INCLUSION PROBABILITIES:**

The variance of Horvitz and Thompson (1952) estimator of population total involve the quantities. Considerable difficulties are involved in the determination of the quantities for most of the selection procedures. As the sample size increases these quantities are hard to calculate.. Attempts have been made to obtain a simplified expression for variance of Horvitz–Thompson (1952) estimator that contains only first order inclusion probabilities.

Hartley and Rao (1962) derived following expression under the random systematic method:

  (8.8.1)

Expression (8.8.1) is correct to order N0. Rao (1963a) further showed that the asymptotic variance formula to order N0 for a sample of size 2 is given as:



 (8.8.2)

The value of in (8.8.2) is 3/32 for Narain’s (1951) procedure, 1/8 for Carroll–Hartley (1964) rejective procedure and 1/4 for the Random Systematic procedure (1950). Rao (1965) further showed that = 0 for the Brewer (1963a) selection procedure. Since Rao–Sampford, Rao (1965) and Sampford (1967), procedure, Durbin (1967) draw-by-draw procedure and the Brewer (1963a) procedure are in same equivalence class therefore = 0 for all these procedures. Rao (1963a) further showed that the approximate formula to the order N1 for a sample of size n is:

 (8.6.31)

Equation (8.8.4) was also shown by Rao (1963b) to be asymptotically valid for Narain (1951) procedure and Carroll – Hartley (1964) rejective procedure.

A simple approximation to ij in terms of I’s and j’s for selection procedures that ensure i = 2pi is given by Brewer (1963a), Durbin (1967), Rao (1965) and Sampford (1967) as:

 (8.8.3)

Brewer and Hanif (1983) gave two approximations to ij’s in terms of  and j’s. The first approximation is:

 (8.8.4)

with  ,  and 

The second one is much simpler than the first approximation, and is given as:

 (8.8.5)

Approximation (8.8.5) performs reasonably well even when one or two values of i are close to unity, each term being less than half the preceding one. Brewer and Hanif (1983)further showed that the approximation (8.8.5) may not result in a feasible set of ij when two of i are close to unity. Herzel (1986) suggested another approximation for ij. This approximation is given as:

 (8.8.6)

Approximation (8.8.8) may produce negative values of ij’s. An example of this is N = 4, n = 2, i = 0.2, 0.25, 0.75 and 0.8, Hanif (1994).

* + 1. **Alternative Expression for variance of Horvitz–Thompson estimator:**

The Sen–Yates–Grundy expression for variance of Horvitz–Thompson (1952) estimator is given in (8.3.4) as:

 (8.3.4)

The alternative expression of (8.3.4) is given below:

 = 

=

= 



 

= 

 (8.8.7)

Now (8.8.7) contains joint probabilities of inclusion,  in last term only. This last term can be manipulated to obtain the approximate expression for variance of Horvitz–Thompson estimator that contains only first order inclusion probabilities. These approximations to the variance are obtained in section 8.8.2 below.

* + 1. **Approximations to variance of Horvitz–Thompson Estimator:**

Shahbaz and Hanif (2003) have obtained following approximations by manipulating last term of (8.8.7).

1. **First Approximation:**

For this approximation consider last term (8.8.7) as:

 (8.8.8)

Using  in (8.8.8) we get:

= 

=  =  (8.8.9)

Using (8.8.9) in (8.8.7) we get:

 



 (8.8.10)

Using  in (8.8.10) the approximate variance formula for Horvitz–Thompson (1952) estimator is:

 (8.8.11)

1. **Second Approximation:**

The second approximation of Shahbaz and Hanif (2003) is obtained by using  in (8.8.8). Using this approximation the variance of Horvitz and Thompson estimator becomes:

 (8.8.12)

This is a special case of the one given by Hartley and Rao (1962) for *a* = . [See 8.6.31.]

1. **Third Approximation:**

The third approximation of Shahbaz and Hanif (2003) is obtained by using  in (8.8.8). Using this expression, approximate variance of Horvitz-Thompson estimator is:

 (8.8.13)

**(e) Fourth Approximation:**

For this approximation we use  in (8.8.8) to get:

 (8.8.14)

This approximation require suitable values of *“a”* and *“b”* for a close approximation of true variance of Horvitz–Thompson estimator. It is also interesting to note that expression given in (8.8.19) transforms to one given by Hartley and Rao (1962) for.

**(f) Fifth Approximation:**

For this approximation Shahbaz and Hanif has used in (8.8.8). The approximate formula for variance of Horvitz-Thompson estimator is:

 (8.8.15)

* + 1. **Empirical Study of Approximate Formulas:**

The choice of approximate variance formulae of Horvitz–Thompson estimator depends upon the closeness of approximate result with the actual one. To decide about this we have selected fifty natural populations and calculated the percentage absolute relative error in variance by using (8.8.11) and (8.8.15); where percentage absolute relative error is defined as:



The relative absolute percentage error has been obtained under Yates – Grundy (1953) draw-by-draw procedure and Brewer (1963) draw-by-draw procedure. The results are given below:

**Table 8.13: Percentage Absolute Relative Error for Various Approximations**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Pop.** | **Yates–Grundy** | | **Brewer** | | **Pop.** | **Yates–Grundy** | | **Brewer** | |
| **A–1** | **A–2** | **A–1** | **A–2** | **A–1** | **A–2** | **A–1** | **A–2** |
| 1. | 6.513 | 0.099 | 6.551 | 0.121 | 11. | 9.429 | 0.360 | 10.118 | 0.633 |
| 2. | 0.493 | 0.012 | 0.346 | 0.025 | 12. | 6.363 | 0.051 | 6.367 | 0.054 |
| 3. | 3.311 | 0.156 | 3.377 | 0.124 | 13. | 12.381 | 1.138 | 14.862 | 3.118 |
| 4. | 11.922 | 0.085 | 12.202 | 0.046 | 14. | 7.047 | 0.178 | 6.993 | 0.181 |
| 5. | 11.901 | 0.073 | 11.821 | 0.083 | 15. | 7.458 | 0.165 | 7.270 | 0.134 |
| 6. | 7.296 | 0.064 | 3.908 | 0.318 | 16. | 11.794 | 0.120 | 11.890 | 0.112 |
| 7. | 17.600 | 1.369 | 22.208 | 2.330 | 17. | 8.152 | 0.220 | 7.974 | 0.262 |
| 8. | 7.583 | 0.031 | 7.600 | 0.032 | 18. | 8.560 | 0.122 | 8.508 | 0.680 |
| 9. | 7.688 | 0.042 | 7.697 | 0.046 | 19. | 5.588 | 0.063 | 5.493 | 0.117 |
| 10 | 5.085 | 0.086 | 5.169 | 0.179 | 20. | 7.776 | 0.292 | 7.742 | 0.303 |

**Table 8.13: Percentage Absolute Relative Error for Various Approximations**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Pop.** | **Yates–Grundy** | | **Brewer** | | **Pop.** | **Yates–Grundy** | | **Brewer** | |
| **A–1** | **A–2** | **A–1** | **A–2** | **A–1** | **A–2** | **A–1** | **A–2** |
| 21. | 11.833 | 0.435 | 11.902 | 0.425 | 36. | 7.906 | 0.087 | 7.296 | 0.163 |
| 22. | 6.223 | 0.017 | 6.090 | 0.026 | 37. | 6.967 | 0.024 | 6.289 | 0.131 |
| 23. | 6.373 | 0.090 | 6.223 | 0.100 | 38. | 5.530 | 0.002 | 5.529 | 0.001 |
| 24. | 6.640 | 1.341 | 6.844 | 3.370 | 39. | 5.394 | 0.008 | 5.319 | 0.001 |
| 25. | 7.638 | 0.135 | 7.554 | 0.143 | 40. | 5.589 | 0.018 | 5.586 | 0.034 |
| 26. | 8.554 | 0.019 | 8.947 | 0.101 | 41. | 12.442 | 0.009 | 12.455 | 0.008 |
| 27. | 8.695 | 0.042 | 8.713 | 0.046 | 42. | 21.574 | 1.780 | 16.035 | 2.298 |
| 28. | 1.816 | 0.037 | 1.741 | 0.033 | 43. | 11.768 | 0.903 | 11.426 | 0.976 |
| 29. | 7.303 | 0.026 | 7.296 | 0.017 | 44. | 5.707 | 0.005 | 5.686 | 0.005 |
| 30. | 5.727 | 0.009 | 5.746 | 0.007 | 45. | 4.404 | 0.146 | 4.299 | 0.149 |
| 31. | 5.543 | 0.003 | 5.545 | 0.001 | 46. | 10.747 | 0.099 | 10.812 | 0.025 |
| 32. | 8.121 | 0.033 | 8.137 | 0.034 | 47. | 7.142 | 0.039 | 7.000 | 0.050 |
| 33. | 4.560 | 0.049 | 4.548 | 0.049 | 48. | 11.252 | 0.425 | 11.038 | 0.384 |
| 34. | 5.186 | 0.013 | 5.172 | 0.001 | 49. | 18.873 | 2.979 | 20.339 | 3.765 |
| 35. | 4.925 | 0.029 | 4.883 | 0.031 | 50. | 7.117 | 0.002 | 7.128 | 0.001 |

From entries of table 8.13 it can be seen that Approximation 2, that is (8.8.15), has less PARE. It can also be seen that PARE of Approximation 1, that is (8.8.11), varies from 1% to 22% for both Yates–Grundy and Brewer procedures. The PARE for Approximation 2, that is (8.8.15), is less than 3% for Yates – Grundy procedure and is less than 4% for Brewer procedures.

* 1. **GODAMBE-JOSHI LOWER BOUND OF VARIANCE**

In this section we have derived the Godambe–Joshi (1965) lower bound to the variance for an unequal probability sampling for Horvitz-Thompson estimator

**THEOREM 8.4**

The lower bound to the variance of Horvitz–Thompson estimator; under model (6.8.2); is gives as:

 (8.9.1)

**PROOF**

Consider

. (8.3.1)

The model in (6.8.2) is:

 (6.8.2)

The expected variance or anticipated variance [Isaki and Fuller (1984)] of the Horvitz–Thompson estimator is:



Substituting the value of  and *Y*, under model (6.8.2), we have from above equation:



 (8.9.2)

Now under the exact selection procedures we have. Substituting this value in (8.9.2) we have:







 (8.9.3)

Opening the square in (8.9.3) and applying the model expectation we have:



=

Applying design expectation we get

 





. (8.9.4)

Now consider the right hand side of (8.7.1)







 (8.9.3)

Now the second term of the expression (8.9.3) if of order N while the leading term contains only expressions of order N2 and N1. Hence asymptotically only the leading term is left, which is same as (8.9.1). Hence under model (6.8.2) the asymptotic variance formula (8.8.3) is valid for al unequal probability sampling which are strictly proportional to size.

**8.10 MODEL BASED STUDY OF APPROXIMATIONS:**

In this section model based treatment of the approximate variance formulae has been given.

**Approximation 1:**

For this consider first approximation as:



Under model (6.8.2) the approximate variance is:







Applying model based expectation:









 (8.10.3)

Approximation (2) given in (8.8.12), under model (6.8.2) is:

 (8.10.4)

Third approximation given in (8.8.13) is

 (8.8.13)

Under the model (6.8.2) the approximation is:

 (8.10.5)

Under model (6.8.2) fourth approximation given in (8.8.14) is:

 (8.10.6)

**8.11 A Generalization of the Horvitz–Thompson Estimator**

In unequal probability sampling some selection procedures cannot be categorized as either with replacement or without replacement procedures in the usual sense. The most important there are intermediate cases where for example, one or more of the population units may appear more than once in sample but the remaining units appear at most once. Hanif and Brewer (1980) developed a general theory of sampling with unequal probabilities which allows population units to appear more than once in a sample. The only condition imposed on the selection procedure is that the total number of appearances of a unit in the sample are fixed. Selection with replacement (multinomial sampling) using Hansen and Hurwitz (1943) and selection without replacement using the Horvitz and Thompson (1952) Estimator are special cases of this general theory. Examples of these are (i) ordinary systemic selection where one or more of the population units are large enough to be certain of selection at least once (ii) Deming’s (1960) procedure which selects several systematic samples with different random starts. (iii) constrained methods of selection, such as dumbbell selection, where one or more units are subject to selection.

This theory is also applicable in principle to a very wide range of sample designs. In particular, the theory is possible in multistage design, to evaluate the probability of selection of each possible final stage sample and then to treat the sampling procedure as though it were single stage selection. In practice, however, stratified and multistage samples will probably continue to be treated best as special cases. The idea is given below:

Let  be the number of times the ith population unit appear in sample and  be the number of times the ordered pair (i, j) appears in the set of  ordered pairs of sample units. Then

 (8.11.1)

The expected values of and  will be written as  and  respectively. Generalized Horvitz-Thompson (GHT) estimator may be defined as

 (8.11.2)

which is clearly unbiased. However, many optimal properties possessed by the Horvitz – Thompson estimator are not carried over to the GHT. The Hansen – Hurwitz estimator, for example, though convenient and widely used, is well known to be inadmissible, and this will generally be true of any estimator for which the  can take values other than 0 and 1.

The variance of the GHT estimator is





 (8.11.3)

 (8.11.4)

Expressions (8.11.3) and (8.11.4) are similar in form to (8.3.6) and (8.3.7) respectively, but more general in its meaning. When selection is strictly without replacement then . In this case (8.11.3) and (8.11.4) reduces to (8.3.6) and (8.3.7) respectively. Writing for convenience , the expressions (8.11.3) and (8.11.4) may be written as

 (8.11.5)  (8.11.6)

For sampling with replacement (multinomial sampling) and (8.11.5) and (8.11.6) reduces to expression (7.4.2) and (7.4.3) respectively. The generalization of the Sen–Yate–Grundy variance estimator is

 (8.11.7)

If then this estimator is unbiased for (8.3.7). If however,  for some , then the bias is non-zero and

 (8.11.8)

Expression (8.11.5) may also be written as

, (8.11.9)

where . (8.11.10)

The  are not functions of n. When sampling is without replacement then  can be chosen independently of n. More generally the  can be chosen such that they approach a finite limit as n tends to infinity. In consequence  may also be considered as approaching a finite limit as n approaches infinity. It then functions in the same way as the finite population correction term does in simple random sampling without replacement.

Since the  are involved in (8.9.7) that expression is not usually easy to calculate. A simpler but biased estimator is

 (8.11.11)

The expectation of (8.11.11)

(8.11.12)

Hence

 (8.11.13)

The bias of the simple variance estimator (8.11.11) is thus seen to be independent of n and to be directly proportional to the difference in variance between the estimator actually employed and the corresponding multinomial sampling estimator. Paradoxically, the lower the variance of the estimator employed, the higher the expectation of its variance estimator. Further, whenever this estimator is more efficient than the corresponding multinomial sampling estimator, it will always tend to appear less efficient and vice versa. This result was obtained for the special case of sampling without replacement by Raj (1954b)

A practical application of the above result is that the efficiencies of  under various sampling procedures may be compared using this biased estimator, the actual efficiencies bearing an inverse relation to the apparent efficiencies.

A factor to correct for the bias may be obtained by using the model: (6.8.2). Under the model

 (8.11.14)

 (8.11.15)

 (8.11.16)

From (8.11.14) to (8.11.16) we obtain

 (8.11.17)

The correction factor term in braces in equation (8.11.17) corresponds to the finite population correction factor in simple random sampling without replacement; when for all I, it actually takes that value.

A numerical example of the use of for the Randomized Systematic Procedure where one of the units has  is given in Hanif and Brewer (1980).

**EXERCISE**

8.1 A sample of 10 villages is drawn using Hansen and Hurwitz selection procedure from a population; size being the 1971 census of population. The relevant information is given as; the total population of the area is 456357. Estimate the total population and find .

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Village | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Z | 865 | 2535 | 2345 | 8356 | 5783 | 8322 | 756 | 1132 | 1156 | 5131 |
| y | 924 | 1945 | 1832 | 2540 | 3345 | 5506 | 4060 | 1013 | 983 | 765 |

8.2 In a sampling with probability proportional to size two units are drawn as:

1. The first unit is selected with probability proportional to size and the second unit is selected from the remaining units with equal probability without replacement.
2. The first unit is selected with probability proportional to size and the second unit is selected with probability proportional to size from the remaining units.

Prove under the above two schemes that Yates and Grundy variance estimator is positive.

8.3 Select all possible samples of size 2 under the above two schemes and prove that Yates and Grundy variance estimator is positive. The population is an under with the common probabilities.

|  |  |  |
| --- | --- | --- |
| **Pi** | **Population 1** | **Population 2** |
| 0.1 | 1.1 | 0.8 |
| 0.2 | 2.1 | 1.3 |
| 0.3 | 1.5 | 2.0 |
| 0.4 | 1.5 | 2.5 |



and t = (t1 + t2)/2 then E(t1) = E(t2) = E(t) = Y and Var(t2) < Var(t1).

8.4 Given a population of N units with measure of size Pi (i = 1, 2, ….N) normed to sum to unity. A sample of two distinct units is drawn with probability proportional to size as. The first unit is selected with probability Pi and second with replacement with probability proportional to Pi(1 – 2Pi) if the two units are selected are the same the sample is rejected and the same procedure is repeated again until two distinct units are selected.

1. Show that the probability of inclusion of the ith unit in the sample is 2Pi.
2. Obtain a formula for the joint probability of inclusion of the ith and jth units in the sample (i ≠ j).

8.5 Two units are selected draw by draw; the first with probability Pi (so that if the ith unit is in fact selected (P1 = Pi) and the second without replacement with probability proportional to is

Pj [(1-2P1)-1 + (1 – 2Pj)-1]

Find the joint probability of inclusion of the ith and jth unit in the sample.