**CHAPTER 9**

**SAMPLING WITH UNEQUAL PROBABILITIES WITHOUT REPLACEMENT (SPECIAL ESTIMATORS)**

**9.1 INTRODUCTION**

In chapter 8 unequal probability sampling without replacement using Horvitz–Thompson estimator has been discussed for estimation of population total. Some other estimators that can be used to estimate the population total with probability proportional to size are given in this chapter. These are:

* 1. Raj Estimator with Yates and Grundy draw-by-draw procedure
  2. Murthy Estimator with Yates and Grundy draw-by-draw procedure
  3. Modified Murthy Estimator with Brewer (1963) draw-by-draw selection procedure
  4. Modified Murthy Estimator with Durbin (1967) draw-by-draw selection procedure
  5. Modified Murthy Estimator with Shahbaz-Hanif-Samiuddin (2003) draw-by-draw selection procedure
  6. General Modified Murthy Estimator with Shahbaz and Hanif (2004) general draw-by-draw procedure
  7. Rao–Hartley–Cochran estimator with Rao–Hartley–Cochran selection procedure
  8. Lahiri’s Estimator with Lahiri’s selection procedure
  9. Poisson Sampling

These estimators are discussed as:

* 1. **RAJ’S ESTIMATOR, [Raj, (1956)]**

Raj (1956a) developed a series of estimators for use with unequal probability sampling without replacement. In this selection procedure the first unit ith (say) is selected with probability proportional to size p­i and the second unit jth (say) is selected from amongst the remaining units with probability proportional to pj(j ≠ i) and so on (Yates and Grundy draw by draw procedure). The probabilities of the successive units being drawn are  and so on. If  are the units selected in the sample of size n in the **same order** then series of unbiased estimators of population total Y proposed by Raj are

 (9.2.1)

Since order of selection is taken into account while selection of units, hence these estimators are called ***ORDERED*** estimators. All these estimators are unbiased of population total Y and are uncorrelated. Any linear combination ΣCiti where ΣCi = 1 is also an unbiased estimator of Y. These ordered estimates were proposed first time by Das (1951).

**THEOREM (9.1)**

Every estimator proposed by Raj is unbiased for population total, that is, . Also these estimators are uncorrelated.

PROOF:

To prove the unbiasedness we consider the estimators given in (9.2.1) in turn.

(i)

(ii)







Similarly it can be proved that.

(iii) Estimators t1 and t2 are uncorrelated i.e. E(t1 t2 ) = Y2

The estimators are uncorrelated if  This may be prove as:

 = .

**THEOREM 9.2:**

The variances of Raj’s estimators are in decreasing order, i.e.

, for r = 2, 3, …, n.

**PROOF**:

The variance of t1 will be



Putting the values of t1 we get



 (9.2.2)

Again 



= 

=

= 

= 

= 

= 



= 

=  **(9.2.3)**

It is obvious from (9.2.2) and (9.2.3) that . Similarly it can be proved .

Raj (1956) further suggested that any liner combination of the estimators given in (9.2.1) is also unbiased for population total. Specifically, Raj used the estimator  for estimation of population total. The estimator  for a sample of size 2 is given as:

 (9.2.4)

**THEOREM (9.3)**

The Raj estimator, , is unbiased estimator of population total with variance

 (9.2.5)

PROOF:

The unbiased ness can be proved easily. We know that t1 and t2 are independent and unbiased estimator of population total therefore their mean will also be unbiased of population total Y.

We know that the estimator  for a sample of size of size 2 is . Since estimators *t*1 and *t*2 are uncorrelated we can write

 (9.2.6)

Using expressions from (9.2.2) and (9.2.3) in (9.2.6) we get:

 = 

=

On simplification we get:

 = 

=  (9.2.5)

that is obviously less than (t1) and  (t2).

After some algebraic manipulation and on simplification we get

 (9.2.7)

This may be put as

(9.2.8)

The estimator  for any “n” proposed by Raj (1956a) is:

, (9.2.8)

Where  and  (9.2.9)

The variance for the general case is complicated and was derived by Pathak (1967a) for any n and is reproduced as:

 **(9.2.10)**

where  denotes the probability of non-inclusion of one or both of the units i and j in the first (r – 1) sample.

An unbiased variance estimator suggested by Raj (1956a) for any n is:

, **(9.2.11)**

which is non negative for 

Also an unbiased estimator of a sample of size 2 is given as:

 (9.2.12)

**EXAMPLE 9.1:**

From the population given in Example 8.1 draw all possible ordered sample of size 2 (Raj’s scheme) and show that the estimator is unbiased for the population total Y. Also obtain the variance of the estimator. Show further that the variance of t2 is less than variance of t1, that is 

# SOLUTION:

The population given in example 8.1 is:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **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 |

The expected value and variance of estimator t2 can be obtained by using:

 (9.2.13)

 (9.2.14)

where the quantity  for Raj scheme is given as .

The necessary calculations to obtain the expectation and variance are given in the table 9.1 below:

**Table 9.1**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Unit** | **yi** | **yj** | **pi** | **pj** |  |  |  |  |
| 1 , 2 | 0.5 | 1.2 | 0.1 | 0.2 | 0.0222 | 5.9 | 0.1311 | 0.7736 |
| 1 , 3 | 0.5 | 2.1 | 0.1 | 0.3 | 0.0333 | 6.8 | 0.2267 | 1.5413 |
| 1 , 4 | 0.5 | 3.2 | 0.1 | 0.4 | 0.0444 | 7.7 | 0.3422 | 2.6351 |
| 2 , 1 | 1.2 | 0.5 | 0.2 | 0.1 | 0.0250 | 5.2 | 0.1300 | 0.6760 |
| 2 , 3 | 1.2 | 2.1 | 0.2 | 0.3 | 0.0750 | 6.8 | 0.5100 | 3.4680 |
| 2 , 4 | 1.2 | 3.2 | 0.2 | 0.4 | 0.1000 | 7.6 | 0.7600 | 5.7760 |
| 3 , 1 | 2.1 | 0.5 | 0.3 | 0.1 | 0.0429 | 5.6 | 0.2400 | 1.3440 |
| 3 , 2 | 2.1 | 1.2 | 0.3 | 0.2 | 0.0857 | 6.3 | 0.5400 | 3.4020 |
| 3 , 4 | 2.1 | 3.2 | 0.3 | 0.4 | 0.1714 | 7.7 | 1.3200 | 10.1640 |
| 4 , 1 | 3.2 | 0.5 | 0.4 | 0.1 | 0.0667 | 6.2 | 0.4133 | 2.5627 |
| 4 , 2 | 3.2 | 1.2 | 0.4 | 0.2 | 0.1333 | 6.8 | 0.9067 | 6.1653 |
| 4 , 3 | 3.2 | 2.1 | 0.4 | 0.3 | 0.2000 | 7.4 | 1.4800 | 10.9520 |
|  |  |  |  |  | 1.0000 |  | 7.0000 | 49.4600 |

Using (9.3.9) we get, , so estimator t2 is unbiased for population total. Again using (9.3.10) we get



Also the sampling variance of t1 is given as:



Now , so 

This shows that 

**EXAMPLE 9.2:**

From the population given in Example 8.1 draw all possible ordered sample of size 2 (Raj’s scheme) and show that the estimator  is unbiased for the population total Y. Also obtain the variance of the estimator.

# SOLUTION:

The data of Example–8.1 is reproduced as:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **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 |

The expected value and variance of estimator  can be obtained by using:

 (9.2.15)

 (9.2.16)

Necessary calculations to obtain the expectation and variance of this estimator are given in table 9.2 below:

**Table 9.2**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Unit** | **yi** | **yj** | **pi** | **pj** |  |  |  |  |
| 1 , 2 | 0.5 | 1.2 | 0.1 | 0.2 | 0.0222 | 5.45 | 0.1211 | 0.6601 |
| 1 , 3 | 0.5 | 2.1 | 0.1 | 0.3 | 0.0333 | 5.90 | 0.1967 | 1.1603 |
| 1 , 4 | 0.5 | 3.2 | 0.1 | 0.4 | 0.0444 | 6.35 | 0.2822 | 1.7921 |
| 2 , 1 | 1.2 | 0.5 | 0.2 | 0.1 | 0.0250 | 5.60 | 0.1400 | 0.7840 |
| 2 , 3 | 1.2 | 2.1 | 0.2 | 0.3 | 0.0750 | 6.40 | 0.4800 | 3.0720 |
| 2 , 4 | 1.2 | 3.2 | 0.2 | 0.4 | 0.1000 | 6.80 | 0.6800 | 4.6240 |
| 3 , 1 | 2.1 | 0.5 | 0.3 | 0.1 | 0.0429 | 6.30 | 0.2700 | 1.7010 |
| 3 , 2 | 2.1 | 1.2 | 0.3 | 0.2 | 0.0857 | 6.65 | 0.5700 | 3.7905 |
| 3 , 4 | 2.1 | 3.2 | 0.3 | 0.4 | 0.1714 | 7.35 | 1.2600 | 9.2610 |
| 4 , 1 | 3.2 | 0.5 | 0.4 | 0.1 | 0.0667 | 7.10 | 0.4733 | 3.3607 |
| 4 , 2 | 3.2 | 1.2 | 0.4 | 0.2 | 0.1333 | 7.40 | 0.9867 | 7.3013 |
| 4 , 3 | 3.2 | 2.1 | 0.4 | 0.3 | 0.2000 | 7.70 | 1.5400 | 11.8580 |
|  |  |  |  |  | 1.0000 |  | 7.0000 | 49.3650 |

Using (9.2.11) we get , so estimator t2 is unbiased for population total. Again using (9.3.10) we get



It can be seen, .

The Raj’s estimator, given in section 9.3 uses the ordered method of selection of samples from a population. The Rao–Blackwell theorem states that the order of selection of units in a sample is not important and this does not play any significant role in improvement of an estimator. Further from the same theorem it is obvious that a sample written as a set is sufficient for estimation of population characteristics; therefore the ordered estimator can be easily avoided.

**9.3 MURTHY’S ESTIMATOR, [MURTHY (1957)]**

By using the idea of sufficiency, Murthy (1957) symmetries Raj’s (1956) estimator to obtain an unordered estimator for estimation of population total. Murthy has shown that the estimator  could be improved by the process of unordering. Murthy (1957) modified Raj’s ordered estimator of population total by a process of Rao – Blackwellization. In main stream statistics the process of Rao–Blackwellization produces the best unbiased estimate even if one starts with a very bad initial estimator such as an estimator based on a single observation. Murthy’s (1957) estimator can likewise be produced by starting with the estimator of total only on the first sample unit selected. The estimator proposed by Murthy (1957) is:

, **(9.3.1)**

where  is the probability of getting unordered sample and  is the conditional probability that the sth sample was selected given that the ith unit has been selected at the first drawn.

To prove the unbiasedness of tsymm, we first prove that  i.e. for any unit i in the population, taken over all samples having unit i drawn first.

(i) For n = 2, the probability of jth as second unit when ith unit is drawn in the first draw is





(ii)For n = 3, the probability of jth and kth as second and third unit when ith unit is drawn in the first draw is



 

Similarly for n = 4



The unbiased ness of Murthy’s estimator along with variance is given in the following theorem

**THEOREM ( 9.3)**

The Murthy Estimator is unbiased for population total Y.

**PROOF**:

The Murthy estimator of population total is:

 (9.3.1)

Taking the expectation and using that idea that  we have:



=  (9.3.2)

When we use  over all samples of size n, the coefficient yi in the sum is 1, hence (9.3.2) becomes

 (9.3.3)

In the following we will consider the Murthy (1957) estimator for a sample of size 2 only under the Yates–Grundy (1953) draw–by–draw selection procedure. For this we consider (9.3.2) as under:



The estimator for a sample of size 2 is:

 (9.3.4)

Now under the Yates–Grundy (1953) procedure we have:



Also the probability of sample under this procedure is:



=  (9.3.5)

Substituting these values in (9.3.2), the Murthy estimator for a sample of size 2 is.

 **(9.3.6)**

We may derive the Murthy estimator for a sample of size 2 alternatively by symmetrizing the Raj (1956) estimator as under;

For this derivation we can see that a sample of size 2 can be selected with probability proportional to size in the following two ways:

(i) ith unit first and jth second with 

(ii) jth unit first and ith second with 

Following this way there will be two ordered estimators i.e.  and .

The corresponding unordered estimator is, then, given by

 **(9.3.7)**

On simplification it comes to be (9.3.6).

Murthy (1957) symmetries the Raj(1956) estimator by using the idea of sufficiency . This definitely reduces the variance of resulting estimator. Murthy (1957) showed that variance of is less than variance of .

**THEOREM (9.3)**

The variance of Murthy’s estimator for a sample size n=2 is derived as

 (9.3.8)

**PROOF**



Putting  and  from (9.3.5) and (9.3.6) we have:

= 

=  = 



=  

 (9.3.9) Unbiased estimator of (7.12.15) is for n = 2.

 (9.3.10)

Murthy (1957) further showed that an unordered and therefore more efficient unbiased variance estimator for  for n = 2 is

 **(9.3.11)**

Pathak (1967a) derived the variance expression for Murthy (1957) estimator for a sample of size n. This expression is:

 (9.3.12)

An unbiased variance estimator of (9.4.4) given by Pathak (1967b) is given as:

 (9.3.13)

where  denotes the conditional probability of selecting a sample “s”, given that units *i* and *j* were selected in that order at the first two draws.

**EXAMPLE 9.3:**

Using the population given in Example–9.2 and drawing all possible samples of size 2, show that the Murthy estimator is unbiased for population total. Also find the variance of this estimator.

**SOLUTION:**

The population given in example 9.2 is:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 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 |

The expected value and variance of estimator  can be obtained by using:

 (9.3.14)

 (9.3.15)

Necessary calculations to obtain the expectation and variance of this estimator are given in table 9.3 below:

**Table 9.3**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Unit** | **yi** | **yj** | **pi** | **pj** |  |  |  |  |
| 1 , 2 | 0.5 | 1.2 | 0.1 | 0.2 | 0.0472 | 5.5294 | 0.2611 | 1.4438 |
| 1 , 3 | 0.5 | 2.1 | 0.1 | 0.3 | 0.0762 | 6.1250 | 0.4667 | 2.8583 |
| 1 , 4 | 0.5 | 3.2 | 0.1 | 0.4 | 0.1111 | 6.8000 | 0.7556 | 5.1378 |
| 2 , 3 | 1.2 | 2.1 | 0.2 | 0.3 | 0.1607 | 6.5333 | 1.0500 | 6.8600 |
| 2 , 4 | 1.2 | 3.2 | 0.2 | 0.4 | 0.2333 | 7.1429 | 1.6667 | 11.9048 |
| 3 , 4 | 2.1 | 3.2 | 0.3 | 0.4 | 0.3714 | 7.5385 | 2.8000 | 21.1077 |
|  |  |  |  |  | 1.0000 |  | 7.0000 | 49.3124 |

Using (9.4.12) we get , so estimator t2 is unbiased for population total. Again using (9.4.13) we get



It can be seen that .

* 1. **Modified Murthy Estimator with Brewer (1963a) Procedure**

**[Shahbaz and Hanif (2003)]**

Shahbaz and Hanif (2003) obtained a modification of Murthy (1957) estimator by using the Brewer (1963a) selection procedure in (9.3.1). The modified estimator given by Shahbaz and Hanif (2003) is derived in the following:

For the modification we consider the Murthy (1957) estimator given in (9.4.1) as:

 (9.3.1)

The estimator for a sample of size 2, given in (9.3.5), is:

 (9.3.5)

Now, for Brewer (1963a) selection procedure we have:

 , 

Now using these values in (9.3.5) we have following modified Murthy estimator:



= (9.4.1)

Where 

This estimator is unbiased for population total.

**THEOREM (9.4)** The variance of modified Murthy estimator under Brewer’s procedure is:

=  (9.4.2)

where  and  are defined appropriately.

**Proof:** The variance of modified Murthy (1957) estimator is given as:



Now, using the expression of from (9.4.1) and by using the fact that this estimator is unbiased for population total, we may write:

= 



= 





 =  (9.4.2)

where  (9.4.3)

 (9.4.4)

 (9.4.5)

**EXAMPLE 9.4:**

Considering data of Example–9.2, obtain modified Murthy estimator for all possible samples of size 2 and find its variance.

**SOLUTION:**

The population given in example 9.2 is:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **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 |

The expected value and variance of estimator  can be obtained by using:

 (9.4.6)



 (9.4.7)

where  for Brewer procedure is given in (8.3.7).

Necessary calculations to obtain the expectation and variance of this estimator are given in table 9.4 below:

**Table 9.4**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Unit | yi | yj | pi | pj |  |  |  |  |
| 1 , 2 | 0.5 | 1.2 | 0.1 | 0.2 | 0.0277 | 9.4187 | 0.2611 | 2.4593 |
| 1 , 3 | 0.5 | 2.1 | 0.1 | 0.3 | 0.0535 | 8.7284 | 0.4667 | 4.0733 |
| 1 , 4 | 0.5 | 3.2 | 0.1 | 0.4 | 0.1188 | 6.3593 | 0.7556 | 4.8048 |
| 2 , 3 | 1.2 | 2.1 | 0.2 | 0.3 | 0.1188 | 8.8375 | 1.0500 | 9.2794 |
| 2 , 4 | 1.2 | 3.2 | 0.2 | 0.4 | 0.2535 | 6.5755 | 1.6667 | 10.9592 |
| 3 , 4 | 2.1 | 3.2 | 0.3 | 0.4 | 0.4277 | 6.5463 | 2.8000 | 18.3296 |
|  |  |  |  |  | 1.0000 |  | 7.0000 | 49.9055 |

Using (9.5.6) we get , so estimator  is unbiased for population total. Again using (9.5.7) we get



**9.5 Modified Murthy Estimator with Durbin (1967) Procedure**

**[Shahbaz (2004)]**

Shahbaz (2004) has obtained this modification of Murthy (1957) estimator by using the Durbin (1967) selection procedure in (9.3.1). For this modification we consider the Murthy (1957) estimator for a sample of size 2 given in (9.3.5). Also, for the Durbin (1967) selection procedure we have:

 (9.5.1)

 (9.5.2)

Also for this selection procedure the quantity  is given in (8.3.7). Substituting these values in (9.3.5) we have the following modified Murthy estimator:



= 

=  (9.5.3)

The estimator given in (9.5.3) is same given by Durbin (1953) for his rejective selection procedure. The unbiasedness for a sample of size 2 is proved in the following theorem.

**THEOREM ( 9.6**) The modified Murthy estimator under Durbin’s (1967) procedure is unbiased for population total Y.

**Proof:** Consider the modified Murthy estimator given in (9.5.3) as:



The expected value of the estimator is obtained below:



= 

=

= 

= 

So the modified Murthy estimator under Durbin’s (1967) estimator is unbiased.

It is interesting to not that the same estimator when used by Durbin (1953) with his rejective procedure was biased.

We now obtain the sampling variance of the modified Murthy estimator (9.5.3) in the following theorem.

**THEOREM 9.7:**

The sampling variance of modified Murthy estimator under Durbin (1967) procedure is:

= 

**PROOF:** The sampling variance of  is obtained as under:

 = 

= 

= 

= 

= 



= 

=  (9.5.4)

Now equating the variance expression given in (9.6.4) with that of the variance of Murthy (1957) estimator given in (9.3.9) we can see that the only change is in the term just after the summation sign.

The unbiased estimator for variance given in (9.5.4) is:

 (9.5.5)

**EXAMPLE 9.5:**

Considering data of Example–9.2, obtain modified Murthy estimator under Durbin procedure for all possible samples of size 2 and find its variance.

**Solution:**

The population given in example 9.2 is:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **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 |

The expected value and variance of estimator  can be obtained by using:

 (9.5.6)



=  (9.5.7)

where  for Durbin procedure is given in (8.3.7).

Necessary calculations to obtain the expectation and variance of this estimator are given in table 9.5 below:

**Table 9.5**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Unit | yi | yj | pi | pj |  |  |  |  |
| 1 , 2 | 0.5 | 1.2 | 0.1 | 0.2 | 0.0277 | 5.5 | 0.1524 | 0.8379 |
| 1 , 3 | 0.5 | 2.1 | 0.1 | 0.3 | 0.0535 | 6.0 | 0.3210 | 1.9260 |
| 1 , 4 | 0.5 | 3.2 | 0.1 | 0.4 | 0.1188 | 6.5 | 0.7722 | 5.0193 |
| 2 , 3 | 1.2 | 2.1 | 0.2 | 0.3 | 0.1188 | 6.5 | 0.7722 | 5.0193 |
| 2 , 4 | 1.2 | 3.2 | 0.2 | 0.4 | 0.2535 | 7.0 | 1.7745 | 12.4215 |
| 3 , 4 | 2.1 | 3.2 | 0.3 | 0.4 | 0.4277 | 7.5 | 3.2078 | 24.0581 |
|  |  |  |  |  | 1.0000 |  | 7.0000 | 49.2822 |

Using (9.5.6) we get , so estimator  is unbiased for population total. Again using (9.5.7) we get



It can be seen that the modified Murthy estimator under Durbin’s (1967) procedure has smaller sampling variance as compared to the actual Murthy (1957) estimator.

**EXAMPLE 9.6:**

Considering data of Example–9.2. Show that the variance estimator given in (9.5.5) is unbiased for .

**SOLUTION:**

The variance estimator given in (9.6.5) is:



The expected value of above estimator is calculated as:

 (9.5.8)

Necessary calculations to obtain above expectation are given in table 9.6 below:

**Table 9.6**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Unit | yi | yj | pi | pj |  |  |  |
| 1 , 2 | 0.5 | 1.2 | 0.1 | 0.2 | 0.0277 | 0.5 | 0.0131 |
| 1 , 3 | 0.5 | 2.1 | 0.1 | 0.3 | 0.0535 | 1.2 | 0.0666 |
| 1 , 4 | 0.5 | 3.2 | 0.1 | 0.4 | 0.1188 | 0.8 | 0.0927 |
| 2 , 3 | 1.2 | 2.1 | 0.2 | 0.3 | 0.1188 | 0.3 | 0.0303 |
| 2 , 4 | 1.2 | 3.2 | 0.2 | 0.4 | 0.2535 | 0.3 | 0.0665 |
| 3 , 4 | 2.1 | 3.2 | 0.3 | 0.4 | 0.4277 | 0.0 | 0.0131 |
|  |  |  |  |  | 1.0000 |  | 0.2822 |

Using (9.6.8), and from last column of table 9.6, we can see that



* 1. **Modified Murthy Estimator with Shahbaz–Hanif–Samiuddin (2003)**

**Procedure [Shahbaz and Hanif (2003)]**

Shahbaz and Hanif (2003) have obtained another modification of Murthy (1957) estimator by using Shahbaz–Hanif–Samiuddin (2003) selection procedure in (9.3.1). For this modification we consider the Murthy (1957) estimator for a sample of size 2 given in (9.3.5). Also, for Shahbaz–Hanif–Samiuddin (2003) the quantities  and  are given before. The quantity  for this selection procedure is given as:

 (9.6.1)

Substituting the values from (9.6.1), (9.6.2) and (9.7.1) in (9.4.6), Shahbaz and Hanif (2003) has obtained following modified Murthy estimator:

 

=  (9.6.2)

This estimator is a slight modification of estimator given by Shahbaz (2004).

The unbiasedness and variance of the estimator is proved in the following theorem.

**THEOREM 9.8:** The modified Murthy estimator under Shahbaz–Hanif–Samiuddin procedure is unbiased for population total with variance given as:



where  and  are defined appropriately.

**PROOF:** The unbiasedness of can be proved in a straight forward manner. The variance is obtained in the following:



= 

= 



= 

= 



= 

(9.6.3)

Where  (9.6.4)

 (9.6.5)

 (9.6.6)

**EXAMPLE 9.7:**

Considering data of Example–9.2, obtain modified Murthy estimator under Shahbaz–Hanif–Samiuddin procedure for all possible samples of size 2. Show that the estimator is unbiased. Also, find its variance.

**SOLUTION:**

The population given in example 9.2 is:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **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 |

The expected value and variance of estimator  can be obtained by using:

 (9.6.7)



=  (9.6.8)

where  for Durbin procedure is given in (9.6.1).

Necessary calculations to obtain the expectation and variance of this estimator are given in table 9.7 below:

**Table 9.7**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Unit** | **yi** | **yj** | **pi** | **pj** |  |  |  |  |
| 1 , 2 | 0.5 | 1.2 | 0.1 | 0.2 | 0.0162 | 9.4152 | 0.1525 | 1.4356 |
| 1 , 3 | 0.5 | 2.1 | 0.1 | 0.3 | 0.0365 | 8.7826 | 0.3208 | 2.8174 |
| 1 , 4 | 0.5 | 3.2 | 0.1 | 0.4 | 0.1165 | 6.6313 | 0.7723 | 5.1212 |
| 2 , 3 | 1.2 | 2.1 | 0.2 | 0.3 | 0.0871 | 8.8716 | 0.7723 | 6.8514 |
| 2 , 4 | 1.2 | 3.2 | 0.2 | 0.4 | 0.2610 | 6.7981 | 1.7743 | 12.0615 |
| 3 , 4 | 2.1 | 3.2 | 0.3 | 0.4 | 0.4828 | 6.6447 | 3.2079 | 21.3158 |
|  |  |  |  |  | 1.0000 |  | 7.0000 | 49.6029 |

Using (9.6.7) we get, , so estimator  is unbiased for population total. Again using (9.6.8) we get



Some other estimators are also proposed in literature by Samiuddin et. el. (1992) following the idea of Raj (1956a) and Murthy (1957). One set of estimators in this class is given as:

 (9.6.9)

These two estimators can be symmetries by using the idea of Murthy (1957). The symmetrized estimator is:

 (9.6.10)

In the following numerical example we have shown that the estimator (9.6.10) performs well as compared to Murthy estimator given in (9.3.7).

**Table 9.7**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Unit** | **yi** | **yj** | **pi** | **pj** |  |  |  |  |
| 1 , 2 | 0.5 | 1.2 | 0.1 | 0.2 | 0.0162 | 6.1660 | 0.2912 | 1.7953 |
| 1 , 3 | 0.5 | 2.1 | 0.1 | 0.3 | 0.0365 | 6.5281 | 0.4974 | 3.2470 |
| 1 , 4 | 0.5 | 3.2 | 0.1 | 0.4 | 0.1165 | 6.8379 | 0.7598 | 5.1951 |
| 2 , 3 | 1.2 | 2.1 | 0.2 | 0.3 | 0.0871 | 6.8041 | 1.0935 | 7.4403 |
| 2 , 4 | 1.2 | 3.2 | 0.2 | 0.4 | 0.2610 | 7.0667 | 1.6489 | 11.6522 |
| 3 , 4 | 2.1 | 3.2 | 0.3 | 0.4 | 0.4828 | 7.2942 | 2.7093 | 19.7622 |
|  |  |  |  |  | 1.0000 |  | 7.0000 | 49.0921 |

Now 

So this symmetrized estimator is more efficient than the Murthy estimator.

**9.7 General Modified Murthy Estimator with Shahbaz–Hanif (2003)**

**General Procedure [Butt and Shahbaz (2004)]**

Butt and Shahbaz (2004) have obtained a general Murthy estimator by using the Shahbaz–Hanif (2003) general selection procedure in (9.3.1). For this modification we consider the Murthy (1957) estimator for a sample of size 2 given in (9.3.5). Also, for Shahbaz–Hanif (2003) general selection procedure the quantities  and  are the same for Yates-Yates (1953) .The quantity  for this selection procedure is given as:

 (9.7.1)

Substituting the value from (9.6.1), (9.6.2) and (9.8.1) in (9.4.6) we have:



 (9.7.2)

The estimator (9.7.2) provides number of estimators for estimation of population total for different choices of constant “*a*”. Some of the special cases of estimator (9.7.2) can be easily derived and are given below:

The General modified Murthy (1957) estimator is given as:

 (9.7.2)

Some of the special cases of this estimator are obtained for selected values of the constant “*a*”. The first special case is obtained by using *a* = 0.5. For this value of *a* the constant *c* of (9.7.2) is 1 so estimator (9.7.2) transform to:





That is the General modified Murthy Estimator transform to usual Murthy (1957) estimator for *a* = 0.5. Another special case is obtained for *a* = 1. For this choice of *a* constant *c* of (9.7.2) is same as constant *k* of Brewer procedure given in (8.4.12). Also the estimator (9.7.2) transform to:



That is for *a* = 1, the general modified Murthy estimator transform to the estimator given by Shahbaz and Hanif (2003b).

The variance of general modified Murthy estimator is obtained in the following theorem.

**THEOREM 9.9:** The General Murthy estimator is unbiased for population total with:

****

where  and  are defined appropriately.

**PROOF:** The unbiasedness can be proved in a straight forward manner. For variance we have:



****

****

**** (9.7.3)

Where  (9.7.4)

 (9.7.5)

 (9.7.6)

The variance expression given in (9.7.3) reduces to (9.3.9) for *a*=0.5 and to (9.4.2) for *a*=1.0.

**9.8 GENERALIZATION OF MURTHY’S ESTIMATOR:**

The generalization of Murthy’s estimator is given for n=2 and without replacement. We start with a general set up and define the probability of selecting ith unit at the first draw as qi such that The probability of drawing jth unit (j ≠ i) at the second draw given that ith unit was drawn at the first draw is q(j | i),  Proceeding as in Raj (1956) we can define

 (9.8.1)

A linear combination of t1 and t2 is given as:

 (9.8.2)

Murthy (1957) observed that (yi, yj) ≡ (yj, yi) ≡ S is sufficient i.e. the order in which units i and j are drawn are irrelevant. Using the process of Rao- Blackwellisation we have two estimator depending on their order of drawing

 with probability 

and  with probability 

Now  where l, m = 1,2 and l ≠ m



(9.8.3)

Of course it would be ideal to free the estimator from c so that any linear combination will lead to the same estimator. This avoids the arbitraries of choosing c = 1/2 in Raj’s scheme. For this to happen the expression in (9.9.3) should be free from c for which the condition is



for all yi and yj

Equating Coefficient of yi and yj to zero we get

 as in Raj (1956) (9.8.4)

However there are two essential differences. These are

1. qi need not be equal to  and this can be exploited
2. If  then Murthy’s (1957) estimator turns out to be   
    which is precisely Murthy’s estimator if 

Pathak (1961) and independently Samiuddin and Hanif (1978) proved strictly in relation to Raj’s estimator that any linear combination would lead to the same Murthy’s estimator. The condition is thus sufficient. We now see it to be necessary as well in a more general context. We have considered several ways of generalizing Murthy’s estimator. These are as follows:

1. Select first unit i with probability  and second unit j ≠ i with probability . Note that in this case starting with Raj’s estimate we get the same Murthy estimator for all values of c. For each selected value of **a** we get a Murthy estimator.
2. Select first unit i with probability  and second unit j with probability . Here c = 1/2 can be chosen to construct Murthy’s estimator starting from Raj’s estimator. Each value of **a** leads to a distinct Murthy estimator.
3. Based on Shahbaz et al (2003) selection procedure. The first unit i is selected with probability, the second unit j ≠ i with probability proportional to . Here we have selected different values of c to construct Murthy’s estimator starting from Raj.
4. Following Samiuddin et al (1992) we select first unit i with probability proportional to pi and the second unit j ≠ i with equal probability without replacement from the remaining units and use the generalized Murthy’s estimator  where having selected the first unit i, the rest of the sample S is selected in what ever way we like. In the above case the resulting estimator is  which is the well-known ratio estimator. Notice that this estimator is both design and model unbiased in this particular case. We also note that there are other ways in which the second unit j can be selected (such as  which will lead to usual Murthy’s estimator).

**9.9 RAO, HARTLEY AND COCHRAN ESTIMATOR [RAO, HARTLEY AND COCHRAN (1962)]**

Rao–Hartley–Cochran (1962) proposed this estimator, by using their own selection procedure, and this can be used for sample of any size. In this selection procedure, if sample of size n is to be selected, the population is divided into n random groups, having  units in each group, where number of units in each group is predetermined. The groups are formed so that number of units in each group is equal. To select a sample of size “*n*” one unit is selected from each group. In this selection procedure, and hence in the estimator, the probabilities of selection are the normed measures of size within the group.

By using the above stated methodology, Rao–Hartley–Cochran proposed the following unbiased estimator of population total:

 (9.9.1)

where  is the initial probability of the sample unit selected from the n random groups, being the initial probability of the Tth unit of the ith group such that  Since will not be equal for all the groups, the probabilities of selection is not proportional to size.

The derivation of variance formula of  is very simple. In this selection procedure randomization is involved at two stages, firstly when groups are formed, secondly at stage of selecting a sample unit. We can use the well known formula given in Chapter 1, that is:



Since , which is constant and so the variance of this term will be zero so we will left with:

 (9.9.2)

Now, since one unit is selected from each group with probability proportional to group size, the concept of sampling with probability proportional to size with replacement can be used, hence the variance formula will

 (9.9.3)

= 

Now the probability that two specified units falls in a group of size Ni is , therefore:

 (9.9.4)

Using (9.9.4) in (9.9.2) we get:





 (9.9.5)

If  where  and *R* is a positive integer one choice of formation of groups is  and (that is to make *k* groups of size *R*+1 and the remaining *n – k* of size *R*). In this case (9.9.5) reduces to

 (9.9.6)

If *N* is a multiple *n* then *K* = 0 and (9.9.10) further reduces to

 (9.9.7)

Unbiased estimators of (9.9.6) and (9.9.7) are respectively

 (9.9.8)

If  this reduced

 (9.9.9)

Note that  reduces to zero if .

This selection procedure is simple and applicable for any sample size.

**EXAMPLE 9.8:**

From the data given in Example 1, find the variance of  for n =2

# SOLUTION: In this example N = 4, n = 2 and



We know that 

Hence 

This method suffers slightly in precision but it has its own advantage i.e. simplicity and applicability for any n.

**EXAMPLE 9.9:**

A population of agriculture holders of ten villages for 1967 and 1971 is given. Select a sample or with n = 2, under RHC sampling scheme and estimate total number of agriculture holders for 1971. Find the variance of this estimator.

|  |  |  |  |
| --- | --- | --- | --- |
| **Village** | **1971 (Yi)** | **1971 (Zi)** | **Cumulated (Zi)** |
| 1 | 60 | 56 | 56 |
| 2 | 55 | 50 | 106 |
| 3 | 60 | 45 | 151 |
| 4 | 70 | 60 | 211 |
| 5 | 75 | 62 | 273 |
| 6 | 65 | 65 | 338 |
| 7 | 50 | 51 | 389 |
| 8 | 60 | 55 | 444 |
| 9 | 65 | 53 | 497 |
| 10 | 80 | 70 | 567 |
| **Total** | 640 |  |  |

**Solution:** Since n = 2, the population given under RHC scheme will be divided at random into two groups with random number tables as

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **GROUP 1** | | | | | **GROUP 2** | | | | |
| **Sr.**  **No.** | **1971**  **Yi1** | **1967**  **Zi1** | **Cumulated**  **Z­i1** |  | **Sr.**  **No.** | **1971**  **Yi2** | **1967**  **Zi2** | **Cumulated**  **Z­i2** |  |
| 5 | 75 | 62 | 62 | 0.1093 | 1 | 60 | 56 | 56 | 0.0988 |
| 6 | 65 | 65 | 127 | 0.1146 | 2 | 55 | 50 | 106 | 0.0882 |
| 9 | 65 | 53 | 170 | 0.0935 | 4 | 70 | 60 | 166 | 0.1058 |
| 3 | 60 | 45 | 225 | 0.0794 | 7 | 50 | 51 | 217 | 0.0899 |
| 10 | 80 | 70 | 295 | 0.1235 | 8 | 60 | 55 | 272 | 0.0970 |
|  |  |  |  |  |  |  |  |  |  |

From group 1, one unit is selected by selecting a random number between 1 and 295. From this group 4th unit is selected with initial probability, . Similarly, from second group 3rd unit with probability  Now



Also 







**EXAMPLE:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sr.**  **No.** | **Number of**  **Schools** | **Enrolment**  **(000)** | **Sr.**  **No.** | **Number of**  **Schools** | **Enrolment**  **(000)** |
| 1 | 38 | 6 | 31 | 94 | 21 |
| 2 | 85 | 12 | 32 | 50 | 9 |
| 3 | 43 | 5 | 33 | 33 | 9 |
| 4 | 56 | 9 | 34 | 28 | 7 |
| 5 | 22 | 2 | 35 | 74 | 21 |
| 6 | 38 | 6 | 36 | 132 | 59 |
| 7 | 51 | 10 | 37 | 39 | 11 |
| 8 | 53 | 9 | 38 | 50 | 12 |
| 9 | 36 | 6 | 39 | 17 | 4 |
| 10 | 64 | 14 | 40 | 49 | 19 |
| 11 | 40 | 6 | 41 | 109 | 27 |
| 12 | 99 | 24 | 42 | 108 | 29 |
| 13 | 54 | 10 | 43 | 117 | 33 |
| 14 | 53 | 11 | 44 | 53 | 14 |
| 15 | 38 | 4 | 45 | 55 | 17 |
| 16 | 16 | 1 | 46 | 71 | 26 |
| 17 | 7 | 1 | 47 | 52 | 17 |
| 18 | 18 | 4 | 48 | 130 | 39 |
| 19 | 51 | 11 | 49 | 98 | 25 |
| 20 | 30 | 6 | 50 | 92 | 24 |
| 21 | 46 | 12 | 51 | 223 | 93 |
| 22 | 64 | 13 | 52 | 156 | 56 |
| 23 | 27 | 4 | 53 | 50 | 8 |
| 24 | 23 | 2 | 54 | 80 | 21 |
| 25 | 9 | 1 | 55 | 78 | 23 |
| 26 | 111 | 69 | 56 | 56 | 10 |
| 27 | 105 | 28 | 57 | 149 | 41 |
| 28 | 83 | 21 | 58 | 44 | 12 |
| 29 | 83 | 20 | 59 | 55 | 14 |
| 30 | 58 | 8 | 60 | 56 | 11 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sr.**  **No.** | **Number of**  **Schools** | **Enrolment**  **(000)** | **Sr.**  **No.** | **Number of**  **Schools** | **Enrolment**  **(000)** |
| 61 | 56 | 10 | 81 | 59 | 14 |
| 62 | 154 | 50 | 82 | 19 | 4 |
| 63 | 34 | 6 | 83 | 28 | 8 |
| 64 | 26 | 11 | 84 | 164 | 42 |
| 65 | 31 | 7 | 85 | 37 | 8 |
| 66 | 64 | 15 | 86 | 32 | 8 |
| 67 | 63 | 18 | 87 | 19 | 3 |
| 68 | 54 | 12 | 88 | 11 | 2 |
| 69 | 27 | 7 | 89 | 60 | 14 |
| 70 | 49 | 12 | 90 | 16 | 3 |
| 71 | 65 | 18 | 91 | 37 | 7 |
| 72 | 51 | 5 | 92 | 91 | 19 |
| 73 | 16 | 4 | 93 | 64 | 16 |
| 74 | 18 | 5 | 94 | 129 | 33 |
| 75 | 106 | 26 | 95 | 37 | 9 |
| 76 | 43 | 12 | 96 | 115 | 28 |
| 77 | 67 | 23 | 97 | 27 | 3 |
| 78 | 22 | 6 | 98 | 18 | 2 |
| 79 | 24 | 5 | 99 | 37 | 2 |
| 80 | 97 | 23 | 100 | 65 | 10 |

***solution***

**RANDOMLY ARRANGED DATA**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Sr.**  **No.** | **No. of**  **Schools** | **Enrolment**  **(000)**  **(Yi)** | **Pi** | **Sr.**  **No.** | **No. of**  **Schools** | **Enrolment**  **(000)**  **(Yi)** | **Pi** |
| 1 | 46 | 12 | 0.00762 | 16 | 38 | 4 | 0.00630 |
| 2 | 43 | 5 | 0.00713 | 17 | 80 | 21 | 0.01326 |
| 3 | 83 | 21 | 0.01376 | 18 | 51 | 11 | 0.00845 |
| 4 | 51 | 5 | 0.00845 | 19 | 34 | 6 | 0.00563 |
| 5 | 132 | 59 | 0.02188 | 20 | 24 | 5 | 0.00398 |
| 6 | 56 | 10 | 0.00928 | 21 | 130 | 39 | 0.02154 |
| 7 | 164 | 42 | 0.02718 | 22 | 56 | 11 | 0.00928 |
| 8 | 31 | 7 | 0.00514 | 23 | 33 | 9 | 0.00547 |
| 9 | 63 | 18 | 0.01044 | 24 | 54 | 10 | 0.00895 |
| 10 | 92 | 24 | 0.01525 | 25 | 19 | 3 | 0.00315 |
| 11 | 27 | 3 | 0.00447 | 26 | 37 | 2 | 0.00613 |
| 12 | 54 | 12 | 0.00895 | 27 | 51 | 10 | 0.00845 |
| 13 | 55 | 14 | 0.00912 | 28 | 16 | 3 | 0.00265 |
| 14 | 50 | 8 | 0.00829 | 29 | 7 | 1 | 0.00116 |
| 15 | 49 | 12 | 0.00812 | 30 | 111 | 69 | 0.01840 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Sr.**  **No.** | **No. of**  **Schools** | **Enrolment**  **(000)**  **(Yi)** | **Pi** | **Sr.**  **No.** | **No. of**  **Schools** | **Enrolment**  **(000)**  **(Yi)** | **Pi** |
| 31 | 86 | 20 | 0.01425 | 66 | 18 | 4 | 0.00298 |
| 32 | 38 | 6 | 0.00630 | 67 | 64 | 14 | 0.01061 |
| 33 | 64 | 15 | 0.01061 | 68 | 22 | 2 | 0.00365 |
| 34 | 97 | 23 | 0.01608 | 69 | 223 | 93 | 0.00696 |
| 35 | 37 | 7 | 0.00613 | 70 | 108 | 29 | 0.01790 |
| 36 | 53 | 14 | 0.00878 | 71 | 58 | 8 | 0.00961 |
| 37 | 154 | 50 | 0.02552 | 72 | 105 | 28 | 0.01740 |
| 38 | 71 | 26 | 0.01177 | 73 | 40 | 6 | 0.00663 |
| 39 | 117 | 33 | 0.01939 | 74 | 65 | 18 | 0.01077 |
| 40 | 55 | 17 | 0.00912 | 75 | 17 | 4 | 0.00282 |
| 41 | 91 | 19 | 0.01508 | 76 | 18 | 2 | 0.00298 |
| 42 | 65 | 10 | 0.01077 | 77 | 53 | 11 | 0.00878 |
| 43 | 94 | 21 | 0.01558 | 78 | 11 | 2 | 0.00182 |
| 44 | 52 | 17 | 0.00862 | 79 | 19 | 4 | 0.00315 |
| 45 | 156 | 56 | 0.02585 | 80 | 49 | 19 | 0.00812 |
| 46 | 26 | 11 | 0.00431 | 81 | 39 | 11 | 0.00646 |
| 47 | 30 | 6 | 0.00497 | 82 | 22 | 6 | 0.00365 |
| 48 | 115 | 28 | 0.01906 | 83 | 50 | 9 | 0.00829 |
| 49 | 67 | 23 | 0.01110 | 84 | 16 | 4 | 0.00265 |
| 50 | 129 | 33 | 0.02138 | 85 | 37 | 8 | 0.00613 |
| 51 | 59 | 14 | 0.00978 | 86 | 36 | 6 | 0.00597 |
| 52 | 23 | 2 | 0.00381 | 87 | 38 | 6 | 0.00630 |
| 53 | 53 | 9 | 0.00878 | 88 | 32 | 8 | 0.00530 |
| 54 | 27 | 7 | 0.00447 | 89 | 74 | 21 | 0.01226 |
| 55 | 149 | 41 | 0.02469 | 90 | 18 | 5 | 0.00298 |
| 56 | 99 | 24 | 0.01641 | 91 | 56 | 10 | 0.00928 |
| 57 | 106 | 26 | 0.01757 | 92 | 85 | 12 | 0.01409 |
| 58 | 56 | 9 | 0.00928 | 93 | 9 | 1 | 0.00149 |
| 59 | 28 | 8 | 0.00464 | 94 | 50 | 12 | 0.00829 |
| 60 | 16 | 1 | 0.00265 | 95 | 109 | 27 | 0.01806 |
| 61 | 37 | 9 | 0.00613 | 96 | 98 | 25 | 0.01624 |
| 62 | 78 | 23 | 0.01293 | 97 | 43 | 12 | 0.00713 |
| 63 | 64 | 13 | 0.1061 | 98 | 27 | 4 | 0.00447 |
| 64 | 44 | 12 | 0.00729 | 99 | 60 | 14 | 0.00994 |
| 65 | 28 | 7 | 0.00464 | 100 | 64 | 16 | 0.01061 |

We next form ten groups of ten each because a sample of size 10 is required. These groups are given below:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Group Number 1** | | | | **Group Number 2** | | | |
| **Sr.**  **No.** | **Zi** | **Yi** | **Pi** | **Sr.**  **No.** | **Zi** | **Yi** | **Pi** |
| 1 | 46 | 12 | 0.00762 | 11 | 27 | 3 | 0.00447 |
| 2 | 43 | 5 | 0.00713 | 12 | 54 | 12 | 0.00895 |
| 3 | 83 | 21 | 0.01376 | 13 | 55 | 14 | 0.00912 |
| 4 | 51 | 5 | 0.00845 | 14 | 50 | 8 | 0.00829 |
| 5 | 132 | 59 | 0.02188 | 15 | 49 | 12 | 0.00812 |
| 6 | 56 | 10 | 0.00928 | 16 | 38 | 4 | 0.00630 |
| 7 | 164 | 42 | 0.02718 | 17 | 80 | 21 | 0.01326 |
| 8 | 31 | 7 | 0.00514 | 18 | 51 | 11 | 0.00845 |
| 9 | 63 | 18 | 0.01044 | 19 | 34 | 6 | 0.00563 |
| 10 | 92 | 24 | 0.01525 | 20 | 24 | 5 | 0.00398 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Group Number 3** | | | | **Group Number 4** | | | |
| **Sr.**  **No.** | **Zi** | **Yi** | **Pi** | **Sr.**  **No.** | **Zi** | **Yi** | **Pi** |
| 21 | 130 | 39 | 0.02154 | 31 | 86 | 20 | 0.01425 |
| 22 | 56 | 11 | 0.00928 | 32 | 38 | 6 | 0.00630 |
| 23 | 33 | 9 | 0.00547 | 33 | 64 | 15 | 0.01061 |
| 24 | 54 | 10 | 0.00895 | 34 | 97 | 23 | 0.01608 |
| 25 | 19 | 3 | 0.00315 | 35 | 37 | 7 | 0.00613 |
| 26 | 37 | 2 | 0.00613 | 36 | 53 | 14 | 0.00878 |
| 27 | 51 | 10 | 0.00845 | 37 | 154 | 50 | 0.02552 |
| 28 | 16 | 3 | 0.00265 | 38 | 71 | 26 | 0.01177 |
| 29 | 7 | 1 | 0.00116 | 39 | 117 | 33 | 0.01939 |
| 30 | 111 | 69 | 0.01840 | 40 | 55 | 17 | 0.00912 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Group Number 5** | | | | **Group Number 6** | | | |
| **Sr.**  **No.** | **Zi** | **Yi** | **Pi** | **Sr.**  **No.** | **Zi** | **Yi** | **Pi** |
| 41 | 91 | 19 | 0.01508 | 51 | 59 | 14 | 0.00978 |
| 42 | 65 | 10 | 0.01077 | 52 | 23 | 2 | 0.00381 |
| 43 | 94 | 21 | 0.01558 | 53 | 53 | 9 | 0.00878 |
| 44 | 52 | 17 | 0.00862 | 54 | 27 | 7 | 0.00447 |
| 45 | 156 | 56 | 0.02585 | 55 | 149 | 41 | 0.02469 |
| 46 | 26 | 11 | 0.00431 | 56 | 99 | 24 | 0.01641 |
| 47 | 30 | 6 | 0.00497 | 57 | 106 | 26 | 0.01757 |
| 48 | 115 | 28 | 0.01906 | 58 | 56 | 9 | 0.00928 |
| 49 | 67 | 23 | 0.01110 | 59 | 28 | 8 | 0.00464 |
| 50 | 129 | 33 | 0.02138 | 60 | 16 | 1 | 0.00265 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Group Number 7** | | | | **Group Number 8** | | | |
| **Sr.**  **No.** | **Zi** | **Yi** | **Pi** | **Sr.**  **No.** | **Zi** | **Yi** | **Pi** |
| 61 | 37 | 9 | 0.00613 | 71 | 58 | 8 | 0.00961 |
| 62 | 78 | 23 | 0.01293 | 72 | 105 | 28 | 0.01740 |
| 63 | 64 | 13 | 0.01061 | 73 | 40 | 6 | 0.00663 |
| 64 | 44 | 12 | 0.00729 | 74 | 65 | 18 | 0.01077 |
| 65 | 28 | 7 | 0.00464 | 75 | 17 | 4 | 0.00282 |
| 66 | 18 | 4 | 0.00298 | 76 | 18 | 2 | 0.00298 |
| 67 | 64 | 14 | 0.01061 | 77 | 53 | 11 | 0.00878 |
| 68 | 22 | 2 | 0.00365 | 78 | 11 | 2 | 0.00182 |
| 69 | 223 | 93 | 0.03696 | 79 | 19 | 4 | 0.00315 |
| 70 | 108 | 29 | 0.01790 | 80 | 49 | 19 | 0.00812 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Group Number 9** | | | | **Group Number 10** | | | |
| **Sr.**  **No.** | **Zi** | **Yi** | **Pi** | **Sr.**  **No.** | **Zi** | **Yi** | **Pi** |
| 81 | 39 | 11 | 0.00646 | 91 | 56 | 10 | 0.00928 |
| 82 | 22 | 6 | 0.00365 | 92 | 85 | 12 | 0.01409 |
| 83 | 50 | 9 | 0.00829 | 93 | 9 | 1 | 0.00149 |
| 84 | 16 | 4 | 0.00265 | 94 | 50 | 12 | 0.00829 |
| 85 | 37 | 8 | 0.00613 | 95 | 109 | 27 | 0.01806 |
| 86 | 36 | 6 | 0.00597 | 96 | 98 | 25 | 0.01624 |
| 87 | 38 | 6 | 0.00630 | 97 | 43 | 12 | 0.00713 |
| 88 | 32 | 8 | 0.00530 | 98 | 27 | 4 | 0.00447 |
| 89 | 74 | 21 | 0.01226 | 99 | 60 | 14 | 0.00994 |
| 90 | 18 | 5 | 0.00298 | 100 | 64 | 16 | 0.01061 |

The necessary calculations to estimate total number of schools are given below:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Selected**  **Units** | **Enrolment**  **yi (000)** | **Pi** | **π** | **yit/Pit \*π** | **(yit/Pit-y’)2** | **π(yit/Pit-y’)2** |
| 1 | 12 | 0.0076 | 0.1261 | 198.5217 | 11571.8549 | 1459.4268 |
| 13 | 14 | 0.0091 | 0.0766 | 117.6000 | 4818.1469 | 368.9068 |
| 28 | 3 | 0.0027 | 0.0852 | 96.3750 | 112318.417 | 9567.7273 |
| 33 | 15 | 0.0106 | 0.1279 | 180.9375 | 2734.8350 | 349.8993 |
| 43 | 21 | 0.0156 | 0.1367 | 184.3085 | 14040.6202 | 1919.7070 |
| 57 | 26 | 0.0176 | 0.1021 | 151.0943 | 182.8806 | 18.6699 |
| 62 | 23 | 0.0129 | 0.1137 | 202.2821 | 97807.5649 | 11119.6536 |
| 79 | 4 | 0.0032 | 0.0721 | 91.5789 | 38493.8947 | 2775.0819 |
| 88 | 8 | 0.0053 | 0.0600 | 90.5000 | 1762.7906 | 105.7558 |
| 96 | 25 | 0.0162 | 0.0996 | 153.3163 | 5295.6642 | 527.4601 |
|  |  |  |  | 1466.5144 |  | 28212.289 |





From above table we have:

Estimated Total = 1466.5144 Thousand

Sample Variance = 28212.289

Standard Error of Estimate = 167.97

Population Total = 1557 Thousand

Actual Variance of Estimate = 233094.5289

Actual Standard Error = 482.7986

**9.10 Lahiri’s Estimator [Lahiri (1951)]**

In chapter 6, we have shown that ratio estimator with equal probability sampling is biased. Lahiri (1951) has given a sampling selection procedure of with probability proportional to the aggregate of the sizes of the sample units. The procedure may be described as a set of n units using simple random sampling without replacement as selected and size measure of those units are aggregated. A random number between, zero and sum of the sizes of n largest units (or any number greater than this) is chosen. If this random number exceeds the aggregate size of the sample random sampling without replacement of n units, the sample is rejected as a whole otherwise accepted .The process is repeated until a sample is accepted.

A simple procedure of selecting a sample with probability proportional to the aggregate of the size has been proposed by Midzuno (1952), Sen (1952). In this procedures the unit is selected with probability proportional to size and the remaining (n – 1) units from the (N – 1) with simple random sampling without replacement. The probability will be

 **(9.10.1)**

as the first unit is selected with probability and the remaining (n – 1) units of the sample with probability 

When the selection of a sample is proportional to its aggregate measure of size, the conventional ratio estimator

  **(9.10.2)**

is unbiased as



as each population unit occur time of the  samples. For n = 2 the variance of  is:

 **(9.10.3)**

Raj (1954) and Sen. (1955) provided an unbiased variance estimator for n = 2 and is given as.

 **(9.10.4)**

This can take negative values. Rao and Vijayan (1977) have proposed new unbiased estimators, which for some populations are non negative. For the case n = 2 these estimators coincide and take the form

 **(9.10.5)**

where  **(9.10.6)**

For n > 2 both the estimator are different. For details see Rao and Vijayan (1975).

**EXAMPLE 9.10:**

Draw all possible distinct samples for n = 2 from the data given in Example 9.5. Calculate Lahiri estimator for each sample and show that and calculate the variance of .

**SOLUTION:**

The Lahiri estimator for a sample of size 2 is defined as:



The necessary calculations to prove the unbiasedness and for calculation of variance are given in the following table.

**Table 9.8**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Unit** | **yi** | **yj** | **pi** | **pj** |  |  |  |  |
| 1 , 2 | 0.5 | 1.2 | 0.1 | 0.2 | 0.1000 | 5.6667 | 0.5667 | 3.2111 |
| 1 , 3 | 0.5 | 2.1 | 0.1 | 0.3 | 0.1333 | 6.5000 | 0.8667 | 5.6333 |
| 1 , 4 | 0.5 | 3.2 | 0.1 | 0.4 | 0.1667 | 7.4000 | 1.2333 | 9.1267 |
| 2 , 3 | 1.2 | 2.1 | 0.2 | 0.3 | 0.1667 | 6.6000 | 1.1000 | 7.2600 |
| 2 , 4 | 1.2 | 3.2 | 0.2 | 0.4 | 0.2000 | 7.3333 | 1.4667 | 10.7556 |
| 3 , 4 | 2.1 | 3.2 | 0.3 | 0.4 | 0.2333 | 7.5714 | 1.7667 | 13.3762 |
|  |  |  |  |  | 1.0000 |  | 7.0000 | 49.3629 |

Now expected value and variance is given as:

 (9.10.7)

 (9.10.8)

Now from above table we can see that  So the Lahiri estimator is unbiased. Also using (9.10.8):



* 1. **Poisson Sampling [Hajek (1964)]**

Poisson sampling is a name given to a selection procedure in which all the population units have independent, and in general un-equal, probabilities of selection πi, Poisson sampling as defined by Hajek (1964) gives each unit in the population a certain probabilities of inclusion in the sample which is denoted by πi for the ith unit. To select a sample, a set of N Bernoulli trials is carried out to determine whether each unit in tern is to be included in the sample or not. It was actually used by the U.S. Bureau of Census for its Annual Survey of Manufacturer (Ogus and Clark 1971). Poisson sampling is known in forestry as 3–P sampling (Sethumadhavi and Rajagopalen 1974). For the Bureau, the chief virtue of Poisson sampling lay in the simple manner in which control could be exercised over the decision as to which units were to be included in various samples. This property of Poisson sampling has been considered in detail by Brewer, Early and Joyce (1972) and Sunter (1977). Mathematical treatment was given by Brewer, Early and Hanif (1980). Poisson sampling suffers from the draw back that the sample size m is a random variable.

The unbiased Horvitz-Thompson estimator of the population total Y is

 **(9.11.1)**

where s is the set of units in the sample.

Since the joint probability of inclusion  takes the simple form  (because of independence), the variance of (9.11.1) in case of Poisson Sampling takes the simple form:

 **(9.11.2)**

An unbiased estimator of (7.12.37) is from (7.8.4)

 **(9.11.3)**

Because the sample size m varies in sampling procedure, the ratio estimator

 (9.11.4)

is more efficient for large samples. If Yi are roughly proportional to πi, it is more efficient for samples of any size.

It can be shown that for large n the mean square error of is given

 **(9.11.5)**

where 

The conventional estimator of (9.11.5) is

 **(9.11.6)**

Although the probability of selecting an empty sample is trivially small when n is large, the problem cannot be ignored in all large scale surveys. This is because such surveys typically employ a fairly detailed stratification by type of unit, and the target sample size within some of the smaller strata is often quite modest. The modified form has been suggested by Ogus and Clark (1971) which ensures that an empty sample is never selected. The name to this sort of sampling given by them MODIFIED POISSON SAMPLING. In this an ordinary Poisson sample is drawn first, but if there are zero units in that sample, a second Poisson sample is drawn and so on repeatedly until a non empty sample is achieved.

The only advantage of modified Poisson sampling over ordinary Poisson sampling is that it ensures a non-empty sample. If the sample selected is much smaller (or much larger) than the target size, modified Poisson sampling provides no remedy. A procedure which ensures a much more stable sample size is called COLLOCATED SAMPLING. This is similar to Poisson sampling but reduces the variation in sampling size by requiring the random variable ri to be uniformly spaced instead uniformly distributed over the interval (0, 1). A random number L (Li = 1, 2, 3, …. N) is chosen with equal probabilities, and a random variable γ is also selected from a uniform distribution over the interval (0, 1). For each i we then define.

 **(9.11.7)**

The Horvitz-Thompson estimator still used, but now no simplification of its variance formula is possible. (For details see Brewer, Early and Hanif 1980).

**Example 9.11**: Given the population of size 100. Draw a Poisson Sample of expected size 12 with  where *xi* denote the digit at *i*th place.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 6 | 3 | 4 | 9 | 3 | 9 | 6 | 4 | 5 | 6 | 3 | 1 | 3 | 4 | 6 | 4 | 3 | 3 | 9 | 3 |
| 6 | 8 | 6 | 5 | 2 | 8 | 2 | 3 | 1 | 2 | 4 | 7 | 3 | 3 | 1 | 1 | 1 | 5 | 2 | 1 |
| 2 | 4 | 8 | 1 | 9 | 2 | 2 | 8 | 6 | 6 | 2 | 7 | 1 | 2 | 8 | 1 | 2 | 9 | 2 | 4 |
| 9 | 9 | 5 | 6 | 4 | 6 | 4 | 9 | 6 | 6 | 8 | 1 | 2 | 4 | 6 | 7 | 6 | 4 | 3 | 8 |
| 2 | 2 | 2 | 9 | 7 | 7 | 9 | 5 | 7 | 9 | 7 | 3 | 5 | 8 | 3 | 2 | 8 | 2 | 3 | 3 |

**Solution:**

Here we have . Hence .

Also we have  which is satisfied for all *i*.

Now for each unit *i*, sample we select a random number, say *Ri*, between 1 and T and select unit *i* if  otherwise unit *i* is not selected. Now we first obtain the value of T in the following table:

|  |  |  |  |
| --- | --- | --- | --- |
| *k* |  | *Fk* |  |
| 1 | 1.0000 | 10 | 10.000 |
| 2 | 1.4142 | 17 | 24.042 |
| 3 | 1.7321 | 15 | 25.981 |
| 4 | 2.0000 | 11 | 22.000 |
| 5 | 2.2361 | 6 | 13.416 |
| 6 | 2.4495 | 14 | 34.293 |
| 7 | 2.6458 | 7 | 18.520 |
| 8 | 2.8284 | 9 | 25.456 |
| 9 | 3.0000 | 11 | 33.000 |
|  | Total | 100 | 206.708 |

Following table shows the selection of random sample from first 20 population units:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Unit** |  | **Random Number** | **Reminder** | **Unit Selected** | **Unit** |  | **Random Number** | **Reminder** | **Unit Selected** |
| 1 | 29.394 | 415598 | 2.182 | 1 | 11 | 20.785 | 049226 | 49.227 | - |
| 2 | 20.785 | 101252 | 101.252 | - | 12 | 12.000 | 598687 | 185.271 | - |
| 3 | 24.000 | 411886 | 205.178 | - | 13 | 20.785 | 351695 | 144.987 | - |
| 4 | 36.000 | 277020 | 70.312 | - | 14 | 24.000 | 689867 | 69.743 | - |
| 5 | 20.785 | 399908 | 193.200 | - | 15 | 29.394 | 057261 | 57.261 | - |
| 6 | 36.000 | 804908 | 184.784 | - | 16 | 24.000 | 482689 | 69.273 | - |
| 7 | 29.394 | 384424 | 177.716 | - | 17 | 20.785 | 288765 | 82.057 | - |
| 8 | 24.000 | 587282 | 173.866 | - | 18 | 20.785 | 481217 | 67.801 | - |
| 9 | 26.833 | 277797 | 71.089 | - | 19 | 36.000 | 232877 | 26.169 | 19 |
| 10 | 29.394 | 558827 | 145.411 | - | 20 | 20.785 | 297646 | 90.938 | - |

**9.13 Empirical and Semi Empirical Comparison**

In the preceding sections we have mentioned several estimators for use with unequal probability sampling without replacement excluding the Horvitz–Thompson estimator. These estimators have some sort of edge over each other in some context. Some perform well in some sort of populations and other in other sort of populations. To decide about the performance of these estimators we have given empirical and semi empirical study using artificial as well as natural populations. We have used following two artificial populations, given by Yates and Grundy (1953), for this empirical study:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Probability of Selection |  | 0.1 | 0.2 | 0.3 | 0.4 |
| Population – 1 |  | 0.8 | 1.4 | 1.8 | 2.0 |
| Population – 2 |  | 0.2 | 0.6 | 0.9 | 0.8 |

We have also given the empirical study using same fifty natural populations that were used in the empirical study of selection procedures given in chapter 8.

The sampling variances of various estimators for artificial populations are given in table 9.9 below:

**Table 9.9**

|  |  |  |
| --- | --- | --- |
| **Selection Procedures** | **Population 1** | **Population 2** |
| **Hansen–Hurwitz Estimator** | 0.500 | 0.125 |
| **Horvitz–Thompson Estimator** |  |  |
| Yates and Grundy | 0.057 | 0.059 |
| Brewer–Durbin–Sampford. | 0.275 | 0.058 |
| Random Systematic | 0.367 | 0.033 |
| Midzuno | 0.384 | 0.240 |
| Raj’s estimator | 0.365 | 0.088 |
| Murthy’s estimator | 0.312 | 0.070 |
| Rao–Hartley–Cochran estimator | 0.333 | 0.083 |
| Lahiri’s estimator | 0.510 | 0.101 |

From this empirical studies with N = 4, n = 2, Pi = 0.1, 0.2, 0.3, 0.4, it is different rather not far to draw any definite conclusion, which estimator is more efficient.

Rao and Bayless (1969) and Bayless and Rao (1970) conducted both empirical and semi empirical studies using the linear model (7.9.1) with α = 0. They found that there were not applicable differences of the efficiency of the Horvitz and Thompson estimator in practice from one selection procedure to another. Brewer and Hanif (1969a) for n = 2, reached the same conclusion.

Hanurav (1962), Vijayan (1966) compared the relative efficiencies under the model (7.9.1) ignoring α of and  and came to the conclusion that was better than  for γ = 0.5. Rao (1966) further proved that  was better than  for all values of γ. Rao and Bayless (1969) and Bayless and Rao (1970) in their semi empirical studies, they found that ­ was nearly always more efficient than . They used the value of γ = 0.5, 0.75, 0.85 and 1. For all values of γ and for n = 2, they found that Murthy’s estimator was consistently more efficient than Raj’s estimator. For γ > 0.5. Murthy’s estimator was more efficient than for γ < 0.875 and less efficient for γ = 1. For n = 3, and 4, Raj’s estimator was less efficient than  for all values of γ. Murthy’s estimator was again more efficient than the  for γ < 0.875. Samiuddin, Hanif and Asad (1978) studied the behaviour of ,  and several other estimator with six artificial and six semi empirical populations. They found that Harvitz-Thompson estimator was reasonably more efficient in almost all cases.

In the semi–empirical studies carried out by the Rao and Bayless (1969), the was found to be consistently less efficient than both  and , its efficiency was greatly affected by the small value of γ. As with Murthy’s and Raj’s estimators, most of the differences were only of the order of a few percent. They also concluded that Murthy’s variance estimator was consistently more stable than the Sen–Yates–Grundy variance estimator for all value of γ. Murthy’s variance estimator also tended to be more stable than Raj’s variance estimator especially for large value of γ and of n. The RHC variance estimator was consistently more stable than Raj, Murthy and Sen–Yates–Grundy variance estimator for all values of γ, however the gains over Murthy’s estimator was not large. For n = 3 and 4 RHC variances estimator was still almost always more stable than Murthy’s variance estimator for γ = 0.875 but for γ = 1 the case was reverse. They found that the Lahiri’s estimator was more efficient than the , ,  and when either (i) few units in the population had large sizes relative to the sizes of remaining units in the population, and sample containing those units gave good estimates of Y, or (ii) the coefficient of variation of the benchmark variable was small Bayless and Rao (1970) extended their investigation for n = 3 and 4, both with respect to the efficiency of the estimator of total and poor performance of the variance estimator.

Next we have given the results of empirical study for fifty natural populations. Following estimator have been used in this empirical study:

* Hansen–Hurwitz Estimator
* Horvitz–Thompson Estimator under Yates–Grundy draw–by–draw procedure
* Horvitz–Thompson Estimator under Brewer draw–by–draw procedure
* Raj’s Ordered Estimator
* Murthy’s Unordered Estimator
* Modified Murthy’s Estimator under Brewer procedure
* Modified Murthy’s Estimator under Durbin’s procedure
* Rao–Hartley–Cochran estimator

For this empirical study the sampling variance of these estimators has been calculated. After calculating the variance we have assigned rank to each estimator within a population depending upon its performance. An estimator that has smallest sampling variance in a population has been assigned a rank of 1, the estimator with second smallest variance is assigned a rank of 2 and so on. The ranks of various estimators in fifty natural populations, along with the ranks of coefficient of variation and correlation coefficient, are given in table 9.10 below. The ranking of coefficient of variation and correlation coefficient has been done on the basis of their magnitude across population. The smallest coefficient of variation and correlation coefficient has been given rank of 1 and so on.

**Table 9.10: Ranks of Various Estimators with Ranks of Coefficient of Variation of X**

**And Correlation Coefficient between X and Y**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Pop. | CV (X) | Rho | HH | HT  (YG) | HT  Brew. | RHC | Raj | Murthy | MM  Brew. | MM  Durb. |
| 1 | 20 | 7 | 7 | 1 | 2 | 6 | 5 | 4 | 8 | 2 |
| 2 | 46 | 49 | 7 | 2 | 6 | 3 | 4 | 1 | 8 | 5 |
| 3 | 38 | 35 | 7 | 1 | 6 | 4 | 3 | 2 | 8 | 5 |
| 4 | 10 | 41 | 7 | 8 | 5 | 3 | 6 | 2 | 1 | 4 |
| 5 | 9 | 10 | 8 | 1 | 5 | 2 | 6 | 3 | 7 | 4 |
| 6 | 45 | 50 | 6 | 7 | 5 | 3 | 2 | 1 | 8 | 4 |
| 7 | 44 | 44 | 7 | 1 | 3 | 6 | 5 | 4 | 8 | 2 |
| 8 | 4 | 9 | 8 | 6 | 3 | 5 | 7 | 4 | 1 | 2 |
| 9 | 12 | 21 | 8 | 7 | 3 | 5 | 6 | 4 | 1 | 2 |
| 10 | 33 | 23 | 7 | 1 | 6 | 4 | 3 | 2 | 8 | 5 |
| 11 | 39 | 25 | 7 | 1 | 3 | 6 | 5 | 4 | 8 | 2 |
| 12 | 15 | 28 | 8 | 7 | 3 | 6 | 5 | 4 | 1 | 2 |
| 13 | 48 | 29 | 7 | 1 | 3 | 6 | 5 | 4 | 8 | 2 |
| 14 | 34 | 45 | 7 | 1 | 6 | 4 | 3 | 2 | 8 | 5 |
| 15 | 37 | 40 | 7 | 6 | 3 | 5 | 4 | 1 | 8 | 2 |
| 16 | 19 | 27 | 8 | 7 | 3 | 6 | 5 | 4 | 1 | 2 |
| 17 | 25 | 12 | 8 | 1 | 6 | 2 | 4 | 3 | 7 | 5 |
| 18 | 40 | 32 | 7 | 1 | 6 | 4 | 3 | 2 | 8 | 5 |
| 19 | 43 | 13 | 7 | 1 | 5 | 6 | 3 | 2 | 8 | 4 |
| 20 | 22 | 16 | 8 | 7 | 3 | 6 | 5 | 4 | 1 | 2 |
| 21 | 29 | 33 | 7 | 1 | 3 | 6 | 5 | 4 | 8 | 2 |
| 22 | 23 | 11 | 8 | 1 | 6 | 3 | 4 | 2 | 7 | 5 |
| 23 | 18 | 14 | 8 | 1 | 6 | 3 | 4 | 2 | 7 | 5 |
| 24 | 50 | 4 | 7 | 1 | 4 | 6 | 5 | 2 | 8 | 3 |
| 25 | 31 | 30 | 7 | 1 | 5 | 6 | 3 | 2 | 8 | 4 |
| 26 | 36 | 26 | 8 | 2 | 5 | 6 | 3 | 1 | 7 | 4 |
| 27 | 21 | 8 | 8 | 1 | 4 | 6 | 5 | 2 | 7 | 3 |
| 28 | 35 | 2 | 8 | 1 | 6 | 2 | 4 | 3 | 7 | 5 |
| 29 | 16 | 20 | 7 | 1 | 4 | 5 | 6 | 2 | 8 | 3 |
| 30 | 17 | 22 | 8 | 7 | 4 | 6 | 5 | 2 | 1 | 3 |
| 31 | 13 | 1 | 8 | 1 | 5 | 3 | 6 | 2 | 7 | 4 |
| 32 | 5 | 3 | 8 | 5 | 2 | 7 | 6 | 4 | 3 | 1 |
| 33 | 11 | 24 | 8 | 7 | 6 | 2 | 4 | 3 | 1 | 5 |
| 34 | 14 | 37 | 7 | 1 | 6 | 2 | 4 | 3 | 8 | 5 |
| 35 | 27 | 18 | 8 | 1 | 6 | 3 | 4 | 2 | 7 | 5 |
| 36 | 30 | 42 | 7 | 6 | 5 | 2 | 3 | 1 | 8 | 4 |
| 37 | 41 | 38 | 7 | 1 | 5 | 6 | 3 | 2 | 8 | 4 |
| 38 | 1 | 19 | 8 | 1 | 5 | 2 | 6 | 3 | 7 | 4 |
| 39 | 24 | 15 | 8 | 7 | 6 | 4 | 3 | 2 | 1 | 5 |
| 40 | 28 | 17 | 7 | 1 | 5 | 6 | 3 | 2 | 8 | 4 |

**Table 9.10: Ranks of Various Estimators with Ranks of Coefficient of Variation of X**

**And Correlation Coefficient between X and Y**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Pop. | CV (X) | Rho | HH | HT  (YG) | HT  Brew. | RHC | Raj | Murthy | MM  Brew. | MM  Durb. |
| 41 | 3 | 5 | 8 | 1 | 5 | 2 | 7 | 2 | 6 | 4 |
| 42 | 47 | 48 | 7 | 8 | 4 | 6 | 5 | 2 | 1 | 3 |
| 43 | 49 | 47 | 7 | 4 | 6 | 3 | 2 | 1 | 8 | 5 |
| 44 | 7 | 39 | 7 | 1 | 3 | 5 | 6 | 2 | 8 | 2 |
| 45 | 26 | 34 | 7 | 1 | 6 | 2 | 4 | 3 | 8 | 5 |
| 46 | 8 | 6 | 7 | 1 | 6 | 2 | 4 | 3 | 8 | 5 |
| 47 | 6 | 46 | 7 | 1 | 5 | 3 | 4 | 2 | 8 | 4 |
| 48 | 32 | 43 | 7 | 8 | 3 | 6 | 5 | 4 | 1 | 2 |
| 49 | 42 | 31 | 7 | 1 | 3 | 6 | 5 | 4 | 8 | 2 |
| 50 | 2 | 36 | 8 | 7 | 5 | 2 | 6 | 3 | 1 | 4 |

Table 9.11 below contains the frequency of ranks of various selection procedures

**Table 9.11: Frequency of Ranks of Various Procedures**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Rank | HH | HT  (YG) | HT  Brew. | RHC | Raj | Murthy | MM  Brew. | MM  Durb. |
| 1 | 0 | 31 | 0 | 0 | 0 | 6 | 12 | 1 |
| 2 | 0 | 2 | 2 | 11 | 2 | 21 | 0 | 12 |
| 3 | 0 | 0 | 13 | 9 | 11 | 10 | 1 | 6 |
| 4 | 0 | 1 | 5 | 5 | 12 | 13 | 0 | 14 |
| 5 | 0 | 1 | 14 | 5 | 14 | 0 | 0 | 16 |
| 6 | 1 | 3 | 16 | 19 | 9 | 0 | 1 | 1 |
| 7 | 28 | 9 | 0 | 1 | 2 | 0 | 10 | 0 |
| 8 | 21 | 3 | 0 | 0 | 0 | 0 | 26 | 0 |
| **Average Rank** | 7.4 | 2.98 | 4.58 | 4.30 | 4.46 | 2.58 | 5.98 | 3.60 |

From table 9.11 we can see that Murthy (1957) estimator clearly outperform all other estimator involved in the study and is closely followed by the Horvitz–Thompson (1952) estimator under Yates–Grundy (1953) draw–by–draw procedure. The modified Murthy estimator under Durbin (1967) draw–by–draw procedure is next best estimator in the study.

The performance of various estimators for various ranges of coefficient of variation and correlation coefficient is also studied and is given in Table 9.11 and Table 9.12 below. From these tables we can see that Murthy (1957) estimator clearly outperform all other estimators involved in the study for almost all ranges of coefficient of variation. Further for small correlation coefficients the Horvitz–Thompson (1952) estimator under Yates–Grundy (1953) draw–by–draw procedure. For other ranges the Murthy (1957) estimator outperforms other estimators involved in the study.

**Table 9.11: Average Ranks of Various Procedures for Different Ranges**

**of Coefficient of Variation**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Rank  CV (Z) | HH | HT  (YG) | HT  Brew. | RHC | Raj | Murthy | MM  Brew. | MM  Durb. |
| 1 – 10 | 7.6 | 3.2 | 4.4 | 3.3 | 5.8 | 2.8 | 5.0 | 3.4 |
| 11 – 20 | 7.7 | 4.0 | 4.2 | 4.4 | 5.0 | 3.0 | 4.3 | 3.3 |
| 21 – 30 | 7.6 | 2.7 | 5.0 | 4.0 | 4.0 | 2.5 | 6.2 | 4.0 |
| 31 – 40 | 7.2 | 2.3 | 4.9 | 4.7 | 3.6 | 2.3 | 7.1 | 3.9 |
| 41 – 50 | 6.9 | 2.7 | 4.4 | 5.1 | 3.9 | 2.3 | 7.3 | 3.4 |

**Table 9.12: Average Ranks of Various Procedures for Different Ranges**

**of Correlation Coefficient**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Rank  ρYZ | HH | HT  (YG) | HT  Brew. | RHC | Raj | Murthy | MM  Brew. | MM  Durb. |
| 1 – 10 | 7.7 | 1.9 | 4.2 | 4.1 | 5.5 | 2.9 | 6.2 | 3.3 |
| 11 – 20 | 7.7 | 2.2 | 5.2 | 4.0 | 4.2 | 2.4 | 6.1 | 4.2 |
| 21 – 30 | 7.6 | 4.1 | 4.1 | 5.3 | 4.4 | 3.0 | 4.4 | 3.1 |
| 31 – 40 | 7.1 | 2.1 | 4.6 | 4.2 | 4.3 | 2.6 | 7.3 | 3.6 |
| 41 – 50 | 6.9 | 4.6 | 4.8 | 3.9 | 3.9 | 2.0 | 5.9 | 3.8 |

The regression summary for various estimators for model

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

is given in table 9.13 below:

**Table 9.13: Regression Summaries for Ranks of Various Estimators**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | HH | HT  (YG) | HT  Brew. | RHC | Raj | Murthy | MM  Brew. | MM  Durb. |
| β0 | 8.190 | 2.588 | 4.406 | 3.602 | 5.595 | 3.237 | 4.295 | 3.469 |
| p-value | 0.000 | 0.005 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| β1 | -0.014 | -0.055 | 0.005 | 0.044 | -0.047 | -0.014 | 0.088 | 0.005 |
| p-value | 0.001 | 0.052 | 0.712 | 0.013 | 0.000 | 0.201 | 0.004 | 0.696 |
| β2 | -0.017 | 0.071 | 0.002 | -0.017 | -0.012 | -0.012 | -0.022 | -0.0002 |
| p-value | 0.000 | 0.014 | 0.908 | 0.337 | 0.276 | 0.245 | 0.446 | 0.987 |
| F | 23.112 | 3.840 | 0.108 | 3.322 | 14.268 | 2.478 | 4.726 | 0.088 |
| p-value | 0.000 | 0.029 | 0.898 | 0.045 | 0.000 | 0.095 | 0.013 | 0.916 |

From table we can readily see that the average rank of Horvitz–Thompson (1952) estimator under Yates – Grundy (1953) draw–by–draw procedure is least and is followed by the Murthy (1957) estimator and modified Murthy estimator under Durbin (1967) procedure. Further, from this table we can see that coefficient of variation has inverse effect on the rank for Hansen–Hurwitz (1943) estimator, Horvitz–Thompson (1952) estimator under Yates–Grundy (1953) draw–by–draw procedure, Raj’s (1956) estimator and Murthy’s (1957) estimator. The results of table 9.13 also show that correlation coefficient and rank of an estimator are positively related for Horvitz–Thompson (1952) estimator under both the selection procedures. This clearly shows that the performance of this estimator will be worst for populations having larger correlation coefficients. It is also worth noting that correlation coefficient has significant effect only on Hansen–Hurwitz (1943) estimator and Horvitz–Thompson (1952) estimator under Yates–Grundy (1953) draw–by–draw procedure performs. For rest of the estimators the correlation coefficient has insignificant effect.