1. Introduction

Welcome to the series of E-learning modules on problems on SRS for variables and attributes. In this module, we are going to cover the basic problems of simple random sampling with replacement and without replacement to obtain the estimates of the population parameters.

By the end of this session, you will be able to:

- Obtain an estimate of population mean, total, variance of the estimators and its Standard error under SRSWR
- Obtain an estimate of population mean, total, variance of the estimators and its standard error under SRSWOR
- Obtain the estimates of population parameters of a population based on attributes

Problem 1:
A population consists of five numbers 2, 3, 6, 8 and 11. Consider all possible samples of size 2 that can be drawn using SRSWR from this population. Find

i. The mean of the population
ii. Standard deviation of the population
iii. The mean of the sampling distribution of the means
iv. The Standard error of the means

Solution:
Mean of the population is equal to Y bar is equal to (summation Yi divided by N) is equal to (thirty by five) which is equal to (six).
Variance is equal to (sigma square) is equal to (summation Yi square divided by N) minus (Y bar square) which is equal to (ten point eight)
And sigma is equal to (three point two nine).
There are 5 squares, which is equal to twenty-five possible samples of size 2 that can be drawn under SRSWR.

Table 1

Sl. No	Samples	Means	Sl. No.	Samples	Means	Sl. No.	Samples	Means
1	(2,2)	2	9	(3,8)	5.5	17	(8,3)	5.5
2	(2,3)	2.5	10	(3,11)	7	18	(8,6)	7
3	(2,6)	4	11	(6,2)	4	19	(8,8)	8
4	(2,8)	5	12	(6,3)	4.5	20	(8,11)	9.5
5	(2,11)	6.5	13	(6,6)	6	21	(11,2)	6.5
6	(3,2)	2.5	14	(6,8)	7	22	(11,3)	7
7	(3,3)	3	15	(6,11)	8.5	23	(11,6)	8.5
8	(3,6)	4.5	16	(8,2)	5	24	(11,8)	9.5
						25	(11,11)	11

In the first, third and fifth columns of the table, samples of size 2 are drawn using SRS with replacement scheme based on the population observations 2, 3, 6, 8 and 11. Since we are under with replacement scheme, the same sample units may be repeated within a sample. Hence starting from the sample (2, 2) as the first sample, we get twenty fifth sample as (11,11) that is corresponding to the last observation of the population.
In the second, fourth and sixth columns corresponding means are obtained for each of the 25 samples.
For example: For the eleventh sample (6,2), the sample mean is obtained as (6 plus 2) divided by (2) is equal to (eight divided by 2) which is equal to (4)

The mean of the sampling distribution of the mean is sum of all means in the above divided by twenty five, which is equal to one hundred and fifty divided by twenty five, which is equal to 6.
The variance sigma square of the sampling distribution of means is obtained by subtracting the mean 6 from each of the above means, squaring the result, adding all the 25 numbers and dividing by 25.

Variance of y bar is equal to (summation i runs from 1 to n) (yi minus y bar whole square) divided by (n) which is equal to (one hundred and thirty five) divided by (twenty five) which is equal to (5 point four zero).
Thus, standard deviation is (square root of variance of y bar) which is equal to (square root of 5 point four zero) which is equal to (2 point three two)
This illustrates, the fact that for finite populations involving SRSWR, sigma y bar square is equal to (sigma square by n) which is equal to (ten point eight by 2) which is equal to (5 point four zero) agreeing with the above value.

2. Illustrations (Part 1)

PROBLEM 2:
The number of boats to catch the fish in a random sample of 30 landing centres was drawn from a coastal area consisting of five thousand landing centres using SRSWOR. The observed number of boats are as follows:
11, 14, 4, 8, 10, 4, 2, 1, 7, 12, 10, 13, 2, 2, 13, 9, 2, 7, 2, 6, 8, 5, 3, 4, 5, 2, 4, 5, 4, 4.

Then, estimate

i. The average number of boats
ii. The total number of boats
iii. Find the estimate of variance of the estimated population total
iv. The standard errors of the mean and population total

Solution:
Given: sample size n is equal to 30 and population size N is equal to five thousand
i)Sample mean y bar is equal to summation yi by n
which is equal to one eighty three divided by thirty
which is equal to six point one
ii) Estimate of population total Y cap is equal to N into y bar
which is equal to five thousand into six point one
which is equal to thirty thousand and five hundred

iii) variance of y cap is equal to (N square into N minus n by N into S square by n)
Which is equal to (five thousand square) into (five thousand minus thirty) into (fourteen point eight five one seven) divided by (five thousand into thirty)
Which is equal to three thousand five hundred and seven point four four three

iv) Estimate of Standard error of the estimate of the population mean
Standard error of the sample mean y bar is equal to (square root of variance of y bar) which is equal to (square root of N minus n into s square) divided by (N into n)
Which is equal to square root of (five thousand minus thirty) into (fourteen point eight five one seven) divided by (five thousand into thirty)
Which is equal to zero point seven zero one four eight nine
Estimate of Standard error of estimated population total is given by
Standard error of Y cap which is equal to (square root of N square into variance of y bar)
Which is equal to square root of (N square into N minus n into s square divided by N into n)
Which is equal to (square root of three thousand five hundred and seven point four four three)
which is equal to (fifty nine point two two three six seven)

RESULT:
i)The average number of boats is equal to 6
ii)The total number of boats is equal to thirty thousand and five hundred
iii)The estimate of variance of the estimated population total is three thousand five hundred and seven point four four three
iv)Estimate the standard errors of the mean is zero point seven zero one four eight nine and population total is fifty nine point two two three six seven

3. Illustrations (Part 2)

Problem 3:
The number of diseased plants in a random sample of 30 farms was drawn from a district consisting of 10,000 farms using SRSWR. The observed numbers of diseased plants of 30 farms are as follows:
1, 4, 11, 2, 5, 1, 1, 1, 7, 2, 3, 3, 2, 2, 3, 1, 2, 7, 2, 6, 3, 5, 3, 4, 5, 3, 4, 5, 1, 4, 5, 4, 1

Then, estimate

i. The average number of diseased plants
ii. The total diseased plants
iii. Find the estimate of variance of the estimated population total
iv. Estimate the standard errors of the mean and population total

Solution:
i) An estimate of the average number of diseased plants
sample mean y bar is equal to (summation yi by n) which is equal to (ninety-one by thirty) which is equal to three point zero three three three
ii) Estimate of total diseased plants Y cap is equal to (N into y bar) which is equal to (ten thousand into three point zero three three three)
which is equal to (thirty thousand three hundred and thirty three point three) approximately (thirty thousand three hundred and thirty three)

iii) An estimate of variance of the estimated population total
Estimate of the (cap) variance of estimated population total under SRSWR is equal to (N square into s square) divided by (n)
Where s square is equal to (summation yi square minus n into y bar square) divided by (n minus 1)
Which is equal to (three hundred and seventy five minus thirty) into (three point zero three three three square) divided by (thirty minus one)
Which is equal to (three point four one two nine)
By substituting the value of s square in the above formula we get,
Estimate of the variance of estimated population total is equal to (one thousand square) into (three point four one two nine square) divided by (thirty).
Which is equal to (eleven lakhs three hundred and seventy six thousand and hundred and seventy six point two four one three)

iv) Estimate the standard error of the population mean is
(square root of variance of y bar) which is equal to (square root of s square by n) which is equal to (square root of 3 point four one two nine by thirty) is equal to (zero point three three seven four)
Estimate the standard error of the population total is
(square root of variance of Y cap) which is equal to (three thousand three hundred and seventy two point eight five eight seven six)

Problem 4:
A population consists of 4 units 5, 6, 3, 7. A random sample of size 2 is taken under SRSWR.

i) Write down all possible samples
ii) Verify that
a)Expected value of y bar is equal to the population mean Y bar
b) Variance of y bar is equal to sigma square by n
c) Expected value of s square is equal to sigma square

Solution:
Here population size N is equal to 4 and the sample size n is equal to 2.
Under SRSWR the total number of possible samples of size that can be drawn from a population of size N is N to the power n. Hence, in this case we will get 4 square, which is equal to 16.

Table 2

Samples	$\bar{y}$	s^2	$\bar{y}^2$
(5, 5)	5	0	25
(5, 6)	5.5	0.5	30.25
(5,3)	4	2	16
(5, 7)	6	2	36
(6,5)	5.5	0.5	30.25
(6,6)	6	0	36
(6, 3)	4.5	4.5	20.25
(6, 7)	6.5	0.5	42.25
(7, 5)	6	2	36
(7, 6)	6.5	0.5	42.25
(7, 3)	5	8	25
(7, 7)	7	0	49
(3, 5)	4	2	16
(3, 6)	4.5	4.5	20.25
(3, 7)	5	8	25
(3, 3)	3	0	9

Under SRSWR the samples can be obtained as follows:
i) Under SRSWR all possible 16 sample are given in the first column of the table.
Second column gives the mean of each of the 16 samples. For example mean of the sample (5,5) is 5 plus 5 by 2 which is equal to 5.
Third column gives the values of s square for each of the samples, which is calculated using the formula s square is equal to (summation y square) minus (n) into (y bar square) divided by (n minus 1)
For example, for the sample (5, 5), s square is equal to (5 square plus 5 square) minus (2) into (5 square) divided by (1), which is equal to (twenty five) minus (twenty five by 1) which is equal to 0.
Last column gives the values of y bar square obtained by squaring the second column values.

Population mean Y bar is equal to summation Yi by N
Where Yi‘s are 5, 6, 7 and 3
Which is equal to five point two five
And the population variance sigma square is equal to (summation Yi square by N) minus (Y

bar square)
Which is equal to (one hundred and nineteen by 4) minus (five point two five square)
Which is equal to two point one eight seven five

ii)

a) Expected value of y bar is equal to (summation y bar divided by N to the power n), which is equal to (eighty four by sixteen), which is equal to (five point two five) which is nothing but the population mean Y bar

Hence, expected value of y bar is equal to the population mean Y bar
From the above table, summation y bar is equal to eighty four
Summation y bar square is equal to four hundred and fifty eight point five
And summation s square is equal to thirty five

b) Variance of y bar is equal to (expected value of y bar square) minus (expected value of y bar whole square)
Which is equal to (four hundred and fifty eight point five by sixteen) minus (5 point two five whole square)
Which is equal to one point zero nine four
Which is equal to sigma square by n
Sigma square by n is equal to (2 point one eight by 2) which is equal to (1 point zero nine four)
Hence, Variance of y bar is equal to (sigma square by n)

c) Expected value of s square is equal to (summation s square by N to the power n), which is equal to two point one eight seven five, which is equal to sigma square.
Hence, expected value of s square is equal to sigma square.

4. Illustrations (Part 3)

Problem 5:

From the five hundred names and addresses, a simple random sample of one hundred names is selected without replacement and twenty-five wrong addresses were found. Estimate the total number of addresses, which needs correction in the list and estimate the standard error of the estimate.

Solution:

Total number of addresses needing correction in the list can be estimated as follows

Let X denote the total number of addresses needing correction in the list

N is the population size is equal to five hundred

Sample size n is equal to one hundred

Let x denote the number of addresses needing correction in the sample selected is equal to twenty five

Then, the sample proportion p is equal to (x by n), which is equal to (twenty five by one hundred) which is equal to (zero point two five)

q is equal to (1 minus p), which is equal to (zero point seven five)

Estimate of number of addresses needing correction in the list

X cap is equal to (N into p) which is equal to (five hundred into zero point two five) which is equal to (one hundred and twenty five)

Standard error of the estimate is obtained as follows:

Standard Error of X cap is equal to (square root of N square into Variance of p)

Which is equal to (square root of N square) into (N minus n) into (p into q) divided by (N into n minus 1)

Which is equal to (square root of five hundred square) into (five hundred minus one hundred) into (zero point two five) into zero (point seven five) divided by (five hundred) into (one hundred minus 1)

Which is equal to nineteen point four six two five

Problem 6:

A random sample of one hundred and seventy boys from eight thousand five hundred and two boys in an area showed that twenty-one had some nutritional deficiency. Estimate the proportion of nutritionally deficient boys and standard error of your estimate.

Solution:

Proportion of nutritionally deficient boys can be estimated as follows:

Let X denote the total number of nutritionally deficient boys in an area.

N is the population size which is equal to (eight thousand five hundred and two)

Sample size n is equal to one hundred and seventy

Let x denote the number of boys who had nutritional deficiency in the sample selected, which is equal to twenty one

Then, the sample proportion p is equal to x by n which is equal to (twenty one by one hundred and seventy) which is equal to (zero point one two)

q is equal to (1 minus p) which is equal to (point eight eight)

Hence, an estimate of proportion of nutritionally deficient boys is equal to (zero point one

two)

Standard error of the estimate can be obtained as follows:
Standard error of p is equal to (square root of variance of p)
Which is equal to square root of (N minus n) into (p into q) divided by (N into n minus 1)
Which is equal to square root of (eight thousand five hundred and two minus one hundred and seventy) into (zero point one two) into (zero point eight eight) divided by (eight thousand five hundred and two) into (one hundred and seventy minus one)
Which is equal to zero point zero two four seven five

5. Illustrations (Part 4)

Problem 7:
A random sample of size 2 households was drawn from a small population of 5 households having weekly income in rupees as follows:

Table 3

Household	1	2	3	4	5
Income	159	149	166	164	155

Draw all the possible SRSWOR samples of size 2 and verify that sample mean is an unbiased estimate of the population mean.

Solution:
Here population size N is equal to 5 and the sample size n is equal to 2
Under SRSWOR, we can draw N C n possible samples, that is 5 C 2 which is equal to ten possible samples.

Table 4

Samples	Corresponding Income(y_i)	$\bar{y}_i$	S^2	$\bar{y}_i$
(1,2)	(159,149)	154	50	23716
(1,3)	(159,166)	162.5	24.5	26406.25
(1,4)	(159,164)	161.5	12.5	26082.25
(1,5)	(159,155)	157	8	24649
(2,3)	(149,166)	157.5	144.5	24806.25
(2,4)	(149,164)	156.5	112.5	24492.25
(2,5)	(149,155)	152	18	23104
(3,4)	(166,164)	165	2	27225
(3,5)	(166,155)	160.5	60.5	25760.25
(4,5)	(164,155)	159.5	40.5	25440.25
	Total	1586	473	251681.5

Ten possible samples of size 2 drawn under SRSWOR is shown in the first column of the table. Since we are under SRS without replacement scheme, the same sample unit cannot be selected again.
In the second column, corresponding to the sample units, the corresponding income is selected from the previous table. For example, sample unit 1 represents the income one five nine and sample unit 2 represents the income one four nine. Hence, corresponding to the

sample (1, 2), we get the income yi as (one five nine, one four nine)
In the third column, means of each of the samples of size 2 are obtained for all the ten samples.

Population mean Y bar is equal to (summation Yi by N) which is equal to (seven hundred and ninety three – which is got by adding all the income figures) divided by 5 which is equal to (one fifty eight point six)
Expected value of y bar is equal to summation (y bar by N C n)
Which is equal to (one thousand five hundred and eighty six by ten)
Which is equal to one hundred fifty eight point six, which is equal to the population mean.
Therefore, sample mean is an unbiased estimate of the population mean.

Here's a summary of our learning in this session, where we have understood:

- The method of drawing SRSWR from a given population and estimate the parameters of the population with the standard error
- The method of drawing SRSWOR from a given population and estimate the parameters of the population
- The SRS procedure to estimate the parameters of a population based on attributes