1. Introduction

Welcome to the series of E-learning modules on Construction of Simple Random samples With and Without Replacement. In this module, we are going to cover the basic problems of Simple Random Sampling With Replacement and Without Replacement applied to three types of data they are Raw data, Grouped data and data related to attributes (Contingency tables).

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

- Construct SRSWR and SRSWOR from an ungrouped data set
- Construct SRSWR and SRSWOR from a grouped data set
- Construct Simple random samples from a population based on attributes that is data in the form of contingency table

Problem 1:
The following data gives the grades of hundred students in Mathematics. Draw a Simple Random sample of size ten without replacement from the group of students and estimate the mean grade from the sample.

Just have a look at grades of 100 students. They are seventy five, fifty eight up to fifty six.

Solution:
Here, population size N is equal to one hundred.
Now, we will take down random numbers in groups of two, from the digit 5 in the second row and fifteenth column moving horizontally using random number table the numbers obtained are:
Fifty eight, ninety nine, eighty three, fifteen, seventy five, twenty four, fifty eight (rejected), ninety two, zero zero, eighty five, sixty six.

The seventh number fifty eight is rejected because it has already appeared once. The number zero zero is usually not considered and if we consider, it corresponds to the last observation of the population. The remaining ten random numbers and the grades having these as identity numbers are shown below:

Table 1

Random Numbers	58	99	83	15	75	24	92	00	85	66
Grades	87	58	79	68	90	58	52	56	83	59

A random sample of size 10 is therefore composed of grades:

Table 2

87	58	79	68	90	58	52	56	83	59

The mean of the sample which gives the estimate of the population mean is y bar is equal to summation, i runs from 1 to n yi by n is equal to six hundred and ninety by ten which is equal to sixty nine.
Identically the population mean of the hundred grades is:
Y bar is equal to summation Yi, i runs from 1 to N by N is equal to seventy two point six six.

2. Illustrations (Part 1)

Problem 2:
Draw a random sample of size ten (without replacement) from the following data:
Forty five , twenty four, forty three, twenty four, seventeen, five, twenty eight, twenty seven, twenty one, eleven, forty six, thirty three, twenty six, twenty four, fourteen, thirty four, twenty one, twenty five, forty eight, thirty five, thirty eight, twenty six, twenty seven, thirty five eight, thirty, twenty six, thirty, twenty eight, twenty one, twenty seven, twenty, thirteen, twenty three, thirty six, thirty eight, twenty, twenty five, thirty one, twenty four, eighteen, twelve.

By making use of the random sampling numbers given below
Five nine six seven, eight nine four one, seven nine eight nine, three three three five, seven five seven seven, nine seven three five, three zero four two, eight four zero nine, seven zero five three, five three six four, five eight seven two, one one four three.

Solution:
There are altogether N is equal to forty one members and these are serially numbered horizontally from 1 to forty one. Since the population size N is two digits, the identity numbers must also be two digits.
If we allot only one identity number to each member, many random numbers will be rejected before we get the required sample. Hence, we can go for the remainder method for the selection of the sample observations.

Let us allot 2 identity numbers to each member as follows:

Table 3

Sl. No.	**1**	**2**	**3**	**.**	**.**	**.**	**39**	**40**	**41**
Identity	01	02	03				39	40	41
Numbers	42	43	44				80	81	82

It will be noticed that identity numbers up to forty one are the same as the corresponding serial numbers and the other identity number when divided by forty one leave remainders which are equal to the serial numbers. With identity number 82, however, the remainder is 00 and the last observation of the population is referred for the sample.

Starting from the first digit of the given random sampling numbers and proceeding row wise we now taken down groups of 2 consecutive digits.
Fifty nine ,sixty seven, eighty nine(rejected),thirty three, thirty five, seventy five, seventy seven, ninety seven(rejected), thirty five, thirty, forty two, eighty four(rejected),zero nine, seventy etc.

The numbers greater than eighty two have been rejected because the identity numbers are from zero one to eighty two only. (If zero zero appeared, it will also have been rejected.) The

numbers greater than forty one are replaced by their remainders on division by forty one and we now have,
Eighteen, twenty six, forty one, thirty eight, thirty three, thirty five, thirty four, thirty six, thirty and one.
Note that these numbers all lie between 1 and forty one and none is repeated. The members having these as serial numbers are respectively,
Forty eight, twenty six, twelve, thirty one, twenty three, thirty eight, thirty six, twenty, twenty seven, and forty five.

3. Illustrations (Part 2)

Problem 3:
Draw SRSWR and SRSWOR samples of size twenty five from the following frequency distribution and find the mean and variance of the sampling distribution.

Table 4

Grade	90-100	80-89	70-79	60-69	50-59	40-49	30-39
No. of students	9	32	43	21	11	3	1

Solution:
All the one hundred and twenty students considered in the frequency distribution are serially numbered. From nine observations in the Class interval, ninety to hundred are numbered zero zero zero to zero zero nine, thirty two in the next interval are numbered zero one zero to zero four one and so on.

The table are shown as follows:

Table 5

				With Replacement	
Grade (Class interval)	No. of students (Frequency)	L.T.C.F	Inclusive C.I	Tally marks	Frequency
90-100	9	9	000-009	//	2
80-89	32	41	010-041	///	3
70-79	43	84	042-084	//// //// //	12
60-69	21	105	085-105	////	4
50-59	11	116	106-116	///	3
40-49	3	119	117-119	/	1
30-39	1	120	120		0
	N=120				

3 digit Random numbers are obtained from page No. One thirty four of Fisher and Yate's Random numbers table starting from column 1 proceeding vertically. The numbers above one twenty are discarded.
zero three four, zero zero three, one one two, zero five two, zero seven nine, zero nine seven, one one two, zero eight six, zero eight two, zero eight four, zero five zero, zero five seven, zero four seven, zero five nine,zero nine four, zero six zero , zero nine nine, one one four, zero two six, zero three three, zero zero three, zero six eight, zero seven eight, one one nine, zero eight three.

Sampling Distribution with Replacement scheme is,

Table 6

Grade (Class interval)	**No. of Sample values(f_i)**	**Mid points (y_i)**	**$f_i y_i$**	**$f_i y_i^2$**
90-100	2	95	190	18050
80-89	3	84.5	253.5	21420.75
70-79	12	74.5	894	66603
60-69	4	64.5	258	16641
50-59	3	54.5	163.5	8910.75
40-49	1	44.5	44.5	1980.25
30-39	0	34.5	0	0
	n=25		1803.5	133605.75

Sample mean y bar is equal to summation i runs from 1 to n fi into yi divided by n which is equal to one eight zero three point five divided by twenty five which is equal to seventy two point one four.
Variance sigma square is equal to summation fi into yi square divided by n minus y bar square which is equal to one three three six zero five point seven five divided by twenty five minus seventy two point one four square which is equal to one forty point zero five zero four.
3 digit Random numbers are obtained from page No. One thirty four of Fisher and Yate's Random numbers table starting from column 1 proceeding vertically. The numbers above one twenty are discarded and since we are in WOR scheme the numbers which are already copied are rejected. zero three four, zero zero three, one one two, zero five two, zero seven nine, zero nine seven, one one two, (rejected), zero eight six, zero eight two, zero eight four, zero five zero, zero five seven, zero four seven, zero five nine, zero nine four, zero six zero , zero nine nine, one one four, zero two six, zero three three, zero zero three(rejected), zero six eight, zero seven eight, one one nine, zero eight three, zero one three, one one one.

Table 7

				Without Replacement	
Grade (Class interval)	No. of students(Frequency)	L.T.C.F	Inclusive C.I	Tally marks	Frequency
90-100	9	9	000-009	/	1
80-89	32	41	010-041	////	4
70-79	43	84	042-084	~~////~~ ~~////~~ //	12
60-69	21	105	085-105	////	4
50-59	11	116	106-116	///	3
40-49	3	119	117-119	/	1
30-39	1	120	120		0
	N=120				

- In column 3 of this table less than cumulative frequencies are obtained for the

frequencies given in column 2

- In column 4 inclusive class intervals are obtained for the cumulative frequencies of column 3
- In column 5 , one by one random numbers are selected from the above selected list and then tally marks are allotted according as the class intervals of column 4
- In the next column total of the tally marks are entered for each row

Sampling Distribution without Replacement scheme is,

Table 8

Grade (Class interval)	No. of Sample values(f_i)	Mid points (y_i)	$f_i y_i$	$f_i y_i^2$
90-100	1	95	95	9025
80-89	4	84.5	338	28561
70-79	12	74.5	894	66603
60-69	4	64.5	258	16641
50-59	3	54.5	163.5	8910.25
40-49	1	44.5	44.5	1980.25
30-39	0	34.5	0	0
	n=25		1793	131721

Sample mean y bar is equal to summation i runs from 1 to n fi into yi divided by n
Which is equal to one seven nine three divided by twenty five which is equal to seventy one point seven two
Variance sigma square is equal to summation fi into yi square divided by n minus y bar square which is equal to one three one seven two one divided by twenty five minus seventy one point seven two square which is equal to one twenty five point zero eight one six.

4. Illustrations (Part 3)

Problem 4:
Draw a Random sample of size 15 using SRSWOR from the following contingency Table relating to casting of votes by democrats and republicans on a particular proposal of national importance.

Table 9

	Favour	**Oppose**	**Undecided**
Democrats	85	78	37
Republicans	180	61	25

Solution:
Less than Cumulative Frequencies are obtained as follows:

Table 10

	Favour	**Oppose**	**Undecided**
Democrats	85 85	78 163	37 200
Republicans	180 318	61 379	25 404

N is equal to four hundred and four which has 3 digits.
Therefore, we shall select 3 digit random numbers leaving the numbers zero zero zero and numbers greater than four not four under Without replacement scheme from Fisher & Yates table from page one hundred and thirty six, Column 1 proceeding vertically.

The random numbers are:
Two twenty one, one nine three, one six seven, zero three two, two three six, one five three, three six nine, one eight eight, zero nine seven, one two nine, three eight six, zero two one, three five seven, two six two, three seven five.

Inclusive class intervals for the cumulative frequencies and tally marks are obtained for each cell which are as follows:

Table 11

	Favour	**Oppose**	**Undecided**
Democrats	000-085 // (2)	086-163 /// (3)	164-200 /// (3)
Republicans	201-318 /// (3)	319-379 /// (3)	380-404 / (1)

Hence a required sample from the given data is,

Table 12

	Favor	Oppose	Undecided
Democrats	2	3	3
Republicans	3	3	1

5. Illustrations (Part 4)

Problem 5:
Show how to select 5 random samples of 4 students each with replacement from the table below using random numbers.

Table 13

Height	Frequency
60-62	5
63-65	18
66-68	42
69-71	27
72-74	8

Solution:
As explained before, less than cumulative frequencies and inclusive class intervals are obtained for the given cumulative frequencies.

Table 14

Height	No. of students (Frequency)	L.T.C.F	Inclusive C.I
60-62	5	5	000-005
63-65	18	23	006-023
66-68	42	65	024-065
69-71	27	92	066-092
72-74	8	100	093-100
	N=100		

Three digit Random numbers (because N equals to hundred has 3 digits) are obtained from page Number one thirty five of Fisher and Yate's table proceeding column wise from column 1.
From the first line we find the sequence zero two six, zero three nine, zero eight four, zero seven zero etc. which we take as a random sampling numbers each which yields the height of a particular student. Thus, zero two six corresponds to a student having height sixty six to sixty eight in which we take as sixty seven in the class mark. Similarly, zero three nine, zero eight four and zero seven zero yields the height of sixty seven, seventy and seventy inches respectively.

By this process we obtain the table which shows the sample numbers drawn, the corresponding heights and the mean height of each of 5 samples. It should be mentioned that although we have entered the random number table on the first line, we could have started anywhere and chosen any specified pattern.

For the first sample four 3 digit random numbers are drawn from the random number table
Find out the Inclusive class interval to which the corresponding random numbers belongs.
Consider the corresponding height, which are represented in the form of class intervals in the previous table and obtain the midpoints of the heights. Note down the value in column 3 and continue till you complete all the 5 samples.
Last column corresponds to the mean (that is sum of four values given in the column3 divided by 4) of corresponding heights for each row.

Table 15

Sl. no	Sample drawn	Corresponding heights	Mean height
1	026, 039, 084, 070	67, 67, 70, 70	68.5
2	018, 072, 029, 058	64, 70, 67, 67	67
3	073, 071, 092, 062	70, 70, 70, 67	69.25
4	082, 074, 034, 055	70, 70, 67, 67	68.5
5	045, 024, 079, 018	67, 67, 70, 64	66.25

This table gives the frequency distribution of sample mean of heights .This is a sampling distribution of the means.

Table 16

Sample mean ($\bar{y}_i$)	Tally marks	Frequency
68.5	/	1
67	/	1
69.25	/	1
68.5	/	1
66.25	/	1
	Total	05

Here's a summary of our learning in this session:

- Demonstrated a technique to select Simple Random samples With and Without replacement from a raw data
- Discussed a technique to select Simple Random samples With and Without replacement from a frequency distribution (grouped data)

- Demonstrated a procedure to construct Simple Random samples With and Without replacement from a population based on attributes