Sampling

Sampling:

The procedure of draw or selecting a sample is known as sampling. There are two types of sampling.

1- Probability Sampling
2- Non- Probability Sampling

Probability Sampling:

A sampling procedure in which the probability of selecting of every unit is non-zero. Every unit in the population has a chance of being selected in the sample.

Probability sampling includes: Simple random sampling, Stratified sampling, Systematic sampling and cluster sampling. These various ways of probability sampling have two things in common.

i- Every element has a known non-zero probability of being sampled and
ii- involves random selection at some point.

Simple Random Sampling:

If the probability of selection of every unit is known and equal it is called simple random sampling.

Stratified sampling:
It is a sampling design. In it we make homogenous groups of hetrogenous units.

Systematic Sampling:
It is a sampling technique. It relies on arranging the study population according to some ordering scheme and then selecting elements at regular intervals through that ordered list.

Cluster sampling:
It is is a sampling technique often used in marketing research. In this technique the total population is divided into homogenous groups (or clusters) and a simple random sample of the groups is selected.

Non-Probability Sampling:
Non-probability sampling is any sampling method where some elements of the population have no chance of selection. It includes: Purposive sampling, Quota sampling, Snow ball sampling and convenient sampling.

Purposive Sampling:
It is also called judgement sampling. In it a person use his personal experience

to select a sample.

Snow-ball sampling:

In snow-ball sampling previous unit refers to the next unit.

Quota sampling:

In quota sampling, the population is first segmented into mutually exclusive sub-groups, just as in stratified sampling. Then judgement is used to select the units from each segment based on specified portion.

Sampling frame:

A sampling frame is the source material or device from which a sample is drawn. It is a list or map guide us to complete list of all sampling units. cover whole population.

Sampling with replacement:

If a sampling unit is drawn from the population and return to the population before the next selection then this process is called sampling with replacement.

Sampling without replacement.

If a sampling unit is drawn from the population and is not return to the population before the next selection then this process is called sampling without replacement. In this case a unit is selected only once.

Sampling design:

A definite plan for obtaining a sample from given population is called a sample design. It is always specified before any data are collected.

Sample Survey:

The collection of information from a part of the population is called making a sample survey.

Census:

The collection of information from all the units of population is called census.

Purpose of Sampling:

I- The basic purpose of sampling is to draw the inferences about the characteristics of population, without examining every unit of population.

II- To find the reliability of the estimate derived from sample.

Let the five students be identified as A, B, C, D and E. Then number of possible random samples of 2 students which can be *with replacement* from this population is $(5)^2 = 25$. Let X_1 denote marks of the student selected first and X_2, the marks of the selected on the second draw. Then the possible random samples $n = 2$ with values of $\bar{X}$ are given below:

Sample No.	Sample Students	Sample Marks (X_1, X_2)	Sample Mean Marks $(\bar{X})$
1	A, A	20, 20	20
2	A, B	20, 15	17.5
3	A, C	20, 12	16
4	A, D	20, 16	18
5	A, E	20, 18	19
6	B, A	15, 20	17.5
7	B, B	15, 15	15
8	B, C	15, 12	13.5
9	B, D	15, 16	15.5
10	B, E	15, 18	16.5
11	C, A	12, 20	16
12	C, B	12, 15	13.5
13	C, C	12, 12	12
14	C, D	12, 16	14
15	C, E	12, 18	15
16	D, A	16, 20	18
17	D, B	16, 15	15.5
18	D, C	16, 12	14
19	D, D	16, 16	16
20	D, E	16, 18	17
21	E, A	18, 20	19
22	E, B	18, 15	16.5
23	E, C	18, 12	15
24	E, D	18, 16	17
25	E, E	18, 18	18

(ii) the number of random samples of 2 students that can be drawn *without replacement* is $\binom{5}{2} = 10$. These samples with values of mean marks are given below:

Sample No.	Sample Students	Sample Marks (X_1, X_2)	Sample Mean Marks $(\bar{X})$
1	A, B	20, 15	17.5
2	A, C	20, 12	16
3	A, D	20, 16	18
4	A, E	20, 18	19
5	B, C	15, 12	13.5
6	B, D	15, 16	15.5
7	B, E	15, 18	16.5
8	C, D	12, 16	14
9	C, E	12, 18	15
10	D, E	16, 18	17

Example 14.2. The following frequency distribution gives the ages of a population of 1,000 college students:

Age (X_i)	14	15	16	17	18	19	20	Total
No. of Students	6	61	270	491	153	15	4	1,000

Using a random number table, select a simple random sample of 20 students. Find the sample mean age and compare it with the population mean age.

First of all we assign a number to each of the 1,000 students from the range 000 to 999. (We start with 000 so that the last student is assigned 999, a three digit number). The 6 students corresponding to the first class ($x = 14$) are assigned the numbers 000, 001, 002, 003, 004 and 005. The next class ($x = 15$) has 61 students to whom we assign the next 61 numbers 006 to 066 inclusive, and so forth. The sampling numbers can conveniently be assigned by compiling a cumulative frequency column. The assigned numbers are shown in column 3 in the following table:

Age (X_i)	No. of Students	Assigned Numbers	f_iX_i
14	6	000 – 005	84
15	61	006 – 066	915
16	270	067 – 336	4320
17	491	337 – 827	8347
18	153	828 – 980	2754
19	15	981 – 995	285
20	4	996 – 999	80
Σ	1,000	----	16,785

Next, we consult a table of random numbers to select a sample of 20 college students at random by finding 20 three-digit numbers between 000 and 999. Starting from the top left hand corner of Table 14.1, we find, going down the columns, the following 20 numbers: 537, 533, 353, 634, 982, 026, 645, 850, 585, 348, 039, 629, 084, 070, 018, 728, 887, 451, 967 and 433. Our sample students are those who correspond to these numbers. Thus the corresponding ages are: 17, 17, 17, 17, 19, 15, 17, 18, 17, 17, 15, 17, 16, 16, 15, 17, 18, 17, 18 and 17. These ages are grouped in the following frequency distribution.

Frequency Distribution of Sample Data

Age (x_i)	Tally	f_i	f_ix_i
15	III	3	45
16	II	2	32
17	𝍸 𝍸 I	11	187
18	III	3	54
19	I	1	19
Σ	---	20	337

Hence the sample mean age, $\bar{x} = \frac{1}{n}\Sigma f_ix_i = \frac{337}{20} = 16.85$ years, and the population mean age, $\mu = \frac{1}{N}\Sigma f_iX_i = \frac{16785}{1000} = 16.78$ years.

Example 14.3. Select a random sample of size 15, using a random number table, from a Poisson distribution with parameter $\mu = 3$.

Given X has a Poisson distribution with

$$P(X = x) = \frac{e^{-3}(3)^x}{x!}, \text{ for } x = 0, 1, 2, 3, \ldots.$$

We first calculate the probabilities associated with each value of x.

Now $P(X = 0) = e^{-3} = 0.0498$, and using the Poisson recurrence formula, we obtain probabilities for other values of x. We find the cumulative probabilities $P(X \leq x)$ and ignoring the decimal point, we assign 3-digit sampling numbers from the range 000 to 999, as 3 decimal places are used in calculating probabilities. The probabilities and the assigned numbers are shown below:

x	$P(X = x)$	Cumulative $P(X \leq x)$	Assigned Numbers
0	0.050	0.050	000 – 049
1	0.149	0.199	050 – 198
2	0.224	0.423	199 – 422
3	0.224	0.647	423 – 646
4	0.168	0.815	647 – 814
5	0.101	0.916	815 – 915
6	0.050	0.966	916 – 965
7	0.022	0.988	966 – 987
8+	0.012	1.000	988 – 999

We now consult a table of random numbers to select a sample of 15 by finding 15-three digit numbers. Let us select three columns, say columns 21, 22 and 23 of Table 14.1, page 8. Then going down the three columns we select the first 15 numbers. These numbers and the x-values (shown in brackets) corresponding to them, are listed below:

946 (6)	381 (2)	303 (2)
028 (0)	615 (3)	869 (5)
252 (2)	596 (3)	987 (7)
613 (3)	626 (3)	249 (2)
747 (4)	901 (5)	158 (1)

The sample results are shown in the following table:

x	0	1	2	3	4	5	6	7
$f(x)$	1	1	4	4	1	2	1	1