SOLVED PROBLEMS

INTRODUCTION TO STATISTICAL THEORY

PART- II

Can be had from:

ILMI KITAB KHANA

KABIR STREET, URDU BAZAR, LAHORE

Published by:

Markazi Kutub Khana,
Kabir Street,
Urdu Bazar, Lahore.

A NOTE TO READERS

THOSE QUESTIONS WHOSE ANSWERS/ SOLUTIONS DIRECTLY APPEAR IN THE TEXT, UNDER DIFFERENT SECTIONS, HAVE BEEN OMITTED IN ORDER TO DECREASE THE BULK OF THE BOOK

Printed by:

Al-Hijaz Printing Press,
Darbar Market, Lahore.

CONTENTS

CHAPTER 14

SURVEY SAMPLING AND SAMPLING DISTRIBUTIONS

Q.14.1. (a) Population. In Statistics, *a population* **(or universe) is defined as an aggregate or totality of units of interest or objects, whether animate or inanimate, concrete or abstract,** *e.g.* **population of College students, population of heights, population of botanical plants, population of opinions, etc. A statistical population thus consists of all measurements or counts of a variable.**

A population may be finite or infinite. A *finite population* contains a finite number of units or objects such as the population of students in a college, the population of all licensed motor drivers, the population of houses in a country, etc. The size of a finite population is usually denoted by the letter N. A population containing an infinite number of units, is called an *infinite population* which is usually regarded as a conceptual device because nobody can ever actually enumerate all the units. The population of all points on a line, the population of all the results obtained by throwing of two dice, the population of all heights between 5 feet and 6 feet, etc. are the examples of infinite population.

Sample. A sample is a part of a population, selected for study. The number of units included in the sample is called the *size* of the sample and is denoted by the letter n. Sometimes, a sample may include the entire population. Samples are selected from population to provide estimates of population parameters. One cannot obtain valid conclusions if the sample is not representative of the population.

Sampling Frame. A list of elements or form of identification of the elements in the population that is used to select a sample, is called *sampling frame*.

Parameter. A numerical quantity used to describe the characteristics of a population is called a parameter.

A Statistic is a numerical descriptive measure computed from a sample.

(b) **Sampling.** The procedure of selecting a sample from the population is called *sampling*, and different procedures give different types of sampling.

Objects of Sampling. The fundamental objects of sampling are:

(i) to get the maximum information about a population without examining each and every unit in it; and

(ii) to find the reliability of the estimates obtained from the sample.

Q.14.16. The population consists of $X_1=2$, $X_2=4$, $X_3=6$, $X_4=8$, $X_5=10$. The mean, μ and the variance, σ^2 of the population are computed as below:

$$\mu = \frac{\sum X_i}{N} = \frac{2+4+6+8+10}{5} = \frac{30}{5} = 6, \text{ and}$$

$$\sigma^2 = \frac{\sum(X_i-\mu)^2}{N} = \frac{(2-6)^2+(4-6)^2+(6-6)^2+(8-6)^2+(10-6)^2}{5}$$

$$= \frac{16+4+0+4+16}{5} = \frac{40}{5} = 8.$$

These five members of the population can be used to simulate an infinite population, if samples are drawn with replacement. Theoretically an infinite number of samples could be drawn, many of them would be identical. All the possible random samples of size 2 that could result when sampling is with replacement, from the population are $(5)^2 = 25$.

But the number of all possible *distinct* samples (*i.e.* when sampling is without replacement) is ${}^5c_2$, *i.e.* $m=10$. These ten distinct samples of size 2 together with their means and variances are given as follows:

No.	Members of Samples		Sample Mean ($\bar{x}_i$)	Sample Variance (S^2)
1	X_1, X_2	2, 4	$\frac{2+4}{2} = 3$	$\frac{(2-3)^2+(4-3)^2}{2} = 1$
2	X_1, X_3	2, 6	4	4
3	X_1, X_4	2, 8	5	9
4	X_1, X_5	2, 10	6	$\frac{(2-6)^2+(10-6)^2}{2} = 16$
5	X_2, X_3	4, 6	5	1
6	X_2, X_4	4, 8	6	4
7	X_2, X_5	4, 10	7	9
8	X_3, X_4	6, 8	7	1
9	X_3, X_5	6, 10	8	4
10	X_4, X_5	8, 10	9	1
Total	--	--	60	50

Now the mean of sample means, $\mu_{\bar{x}}$, is

$$\mu_{\bar{x}} = \frac{\sum \bar{x}_i}{m} = \frac{60}{10} = 6 = \mu$$

Thus the mean of sample means is equal to the mean of the population. This property is of fundamental importance.

The mean of sample variance is

$$E(S^2) = \frac{50}{10} = 5$$

which is not identical with the population variance. The sampling variance is a biased estimator.

Q.14.17. We are required to select a random sample of 10 households from a list of 250 households, by using a table of random sampling numbers.

We first list the households and assign a 3-digit serial number to each household from 001 to 250 as 250 is a 3-digit number. Then using the first 3 columns from Random Numbers