UNIT 1 PARAMETRIC AND NON-PARAMETRIC STATISTICS

Structure

1.0 INTRODUCTION

In this unit you will be able to know the various aspects of parametric and non-parametric statistics. A parametric statistical test specifies certain conditions such as the data should be normally distributed etc. The non-parametric statistics does not require the conditions of parametric stats. In fact non-parametric tests are known as distribution free tests.

In this unit we will study the nature of quantitative data and various descriptive statistical measures which are used in the analysis of such data. These include measures of central tendency, variability, relative position and relationships of normal probability curve etc. will be explained.

The computed values of various statistics are used to describe the properties of particular samples. In this unit we shall discuss inferential or sampling statistics, which are useful to a researcher in making generalisations of inferences about the populations from the observations of the characteristics of samples.

For making inferences about various population values (parameters), we generally make use of parametric and non-parametric tests. The concept and assumptions of parametric tests will be explained to you in this section along with the inference regarding the means and correlations of large and small samples, and significance of the difference between the means and correlations in large and small independent samples.

The assumptions and applications of analysis of variance and co-variance for testing the significance of the difference between the means of three or more samples will also be discussed.

In the use of parametric tests for making statistical inferences, we need to take into account certain assumptions about the nature of the population distribution, and also the type of the measurement scale used to quantify the data. In this unit you will learn about another category of tests which do not make stringent assumptions about the nature of the population distribution. This category of test is called distribution free or non-parametric tests. The use and application of several non-parametric tests involving unrelated and related samples will be explained in this unit. These would include chi-square test, median test, Man-Whitney U test, sign test and Wilcoxon-matched pairs signed-ranks test.

1.1 OBJECTIVES

After reading this unit, you will be able to:

- define the terms parametric and non-parametric statistics;
- differentiate between parametric and non-parametric statistics;
- describe the nature and meaning of parametric and non-parametric statistics;
- delineate the assumptions of parametric and non-parametric statistics; and
- list the advantages and disadvantages of parametric and non-parametric statistics.

1.2 DEFINITION OF PARAMETRIC AND NON-PARAMETRIC STATISTICS

Statistics is an Independent branch and its use is highly prevalent in all the fields of knowledge. Many methods and techniques are used in statistics. These have been grouped under parametric and and non-parametric statistics. Statistical tests which are not based on a *normal distribution* of data or on any other assumption are also known as distribution-free tests and the data are generally ranked or grouped. Examples include the *chi-square test* and *Spearman's rank correlation coefficient*.

The first meaning of *non-parametric* covers techniques that do not rely on data belonging to any particular distribution. These include, among others:

1) Distribution free methods: This means that there are no assumptions that the data have been drawn from a normally distributed population. This consists of non-parametric *statistical models*, *inference* and *statistical tests*.

2) Non-parametric statistics: In this the statistics is based on the *ranks* of observations and do not depend on any distribution of the population.

3) No assumption of a structure of a model: In non-parametric statistics, the techniques do not assume that the *structure* of a model is fixed. In this, the

individual variables *are* typically assumed to belong to parametric distributions, and assumptions about the types of connections among variables are also made. These techniques include, among others:

a) Non-parametric regression

b) Non-parametric hierarchical Bayesian models.

In non-parametric regression, the structure of the relationship is treated non-parametrically.

In regard to the Bayesian models, these are based on the *Dirichlet process*, which allows the number of *latent variables* to grow as necessary to fit the data. In this the individual variables however follow parametric distributions and even the process controlling the rate of growth of latent variables follows a parametric distribution.

4) The assumptions of a *classical* or *standard tests* are not applied to non-parametric tests.

Parametric tests

Parametric tests normally involve data expressed in absolute numbers or values rather than ranks; an example is the *Student's t-test*.

The parametric statistical test operates under certain conditions. Since these conditions are not ordinarily tested, they are assumed to hold valid. The meaningfulness of the results of a parametric test depends on the validity of the assumption. Proper interpretation of parametric test based on normal distribution also assumes that the scene being analysed results from measurement in at least an interval scale.

Let us try to understand the term population. Population refers to the entire group of people which a researcher intends to understand in regard to a phenomenon. The study is generally conducted on a sample of the said population and the obtained results are then applied to the larger population from which the sample was selected.

Tests like t, z, and F are called parametrical statistical tests.

T-tests: A T-test is used to determine if the scores of two groups differ on a single variable.

A t-test is designed to test for the differences in mean scores. For instance, you could use t-test to determine whether writing ability differs among students in two classrooms.

It may be mentioned here that the parametric tests, namely, t-test and F-test, are considered to be quite robust and are appropriate even when some assumptions are not met.

Parametric tests are useful as these tests are most powerful for testing the significance or trustworthiness of the computed sample statistics. However, their use is based upon certain assumptions. These assumptions are based on the nature of the population distribution and on the way the type of scale is used to quantify the data measures.

Let us try to understand what is a scale and its types. There are four types of scales used in measurement viz., nominal scale, ordinal scale, interval scale, and ratio scale.

1) Nominal scale deals with nominal data or classified data such as for example the population divided into males and females. There is no ordering of the data in that it has no meaning when we say male > female. These data are also given

arbitrary labels such as m / f and 1 //0 . These are also called as categorical scale , that is these are scales with values that are in terms of categories (i.e. they are names rather than numbers).

2) Ordinal scale deals with interval data. These are in certain order but the differences between values are not important. For example, degree of satisfaction ranging in a 5 point scale of 1 to 5, with 1 indicating least satisfaction and 5 indicating high satisfaction.

3) Interval scale deals with ordered data with interval. This is a constant scale but has no natural zero. Differences do make sense . Example of this kind of data includes for instance temperature in Centigrade or Fahrenheit. The dates in a calendar. Interval scale possesses two out of three important requirements of a good measurement scale, that is, magnitude and equal intervals but lacks the real or absolute zero point.

4) Ratio scale deals with ordered, constant scale with a natural zero. Example of this type of data include for instance, height, weight, age, length etc.

The sample with small number of items are treated with non-parametric statistics because of the absence of normal distribution, e.g. if our sample size is 30 or less; ($N \leq 30$). It can be used even for nominal data along with the ordinal data.

A non-parametric statistical test is based on model that specifies only very general conditions and none regarding the specific form of the distribution from which the sample was drawn.

Certain assumptions are associated with most non-parametric statistical tests, namely that the observation are independent, perhaps that variable under study had underlying continuity, but these assumptions are fewer and weaker than those associated with parametric tests.

More over as we shall see, non-parametric procedures often test different hypotheses about population than do parametric procedures.

Finally, unlike parametric tests, there are non-parametric procedures that may be applied appropriately to data measured in an ordinal scale, or in a nominal scale or categorical scale.

Non-parametric statistics deals with small sample sizes.

Non-parametric statistics are assumption free meaning these are not bound by any assumptions.

Non-parametric statistics are user friendly compared with parametric statistics and economical in time.

We have learnt that parametric tests are generally quite robust and are useful even when some of their mathematical assumptions are violated. However, these tests are used only with the data based upon ratio or interval measurements.

In case of counted or ranked data, we make use of non-parametric tests. It is argued that non-parametric tests have greater merit because their validity is not based upon assumptions about the nature of the population distribution, assumptions that are so frequently ignored or violated by researchers using parametric tests. It may be noted that non-parametric tests are less precise and have less power than the parametric tests.

1.3 ASSUMPTIONS OF PARAMETRIC AND NON-PARAMETRIC STATISTICS

1.3.1 Assumptions of Parametric Statistics

Parametric tests like, ‘t and f’ tests may be used for analysing the data which satisfy the following conditions :

The population from which the sample have been drawn should be normally distributed.

Normal Distributions refer to Frequency distribution following a normal curve, which is infinite at both the ends.

The variables involved must have been measured interval or ratio scale.

Variable and its types: characteristic that can have different values.

Types of Variables

Dependent Variable: Variable considered to be an effect; usually a measured variable. Independent Variable: Variable considered being a cause.

The observation must be independent. The inclusion or exclusion of any case in the sample should not unduly affect the results of study.

These populations must have the same variance or, in special cases, must have a known ratio of variance. This we call homosedasticity.

The samples have equal or nearly equal variances. This condition is known as equality or homogeneity of variances and is particularly important to determine when the samples are small.

The observations are independent. The selection of one case in the sample is not dependent upon the selection of any other case.

1.3.2 Assumptions of Non-parametric Statistics

We face many situations where we can not meet the assumptions and conditions and thus cannot use parametric statistical procedures. In such situation we are bound to apply non-parametric statistics.

If our sample is in the form of nominal or ordinal scale and the distribution of sample is not normally distributed, and also the sample size is very small, it is always advisable to make use of the non-parametric tests for comparing samples and to make inferences or test the significance or trust worthiness of the computed statistics. In other words, the use of non-parametric tests is recommended in the following situations:

Where sample size is quite small. If the size of the sample is as small as N=5 or N=6, the only alternative is to make use of non-parametric tests.

When assumption like normality of the distribution of scores in the population are doubtful, we use non-parametric tests.

When the measurement of data is available either in the form of ordinal or nominal scales or when the data can be expressed in the form of ranks or in the shape of + signs or – signs and classification like “good-bad”, etc., we use non-parametric statistics.

The nature of the population from which samples are drawn is not known to be normal.

The variables are expressed in nominal form.

The data are measures which are ranked or expressed in numerical scores which have the strength of ranks.

1.4 ADVANTAGES OF NON-PARAMETRIC STATISTICS

If the sample size is very small, there may be no alternative except to use a non-parametric statistical test.

Non-parametric tests typically make fewer assumptions about the data and may be relevant to a particular situation.

The hypothesis tested by the non-parametric test may be more appropriate for research investigation.

Non-parametric statistical tests are available to analyse data which are inherently in ranks as well as data whose seemingly numerical scores have the strength of ranks.

For example, in studying a variable such as anxiety, we may be able to state that subject A is more anxious than subject B without knowing at all exactly how much more anxious A is. Thus if the data are inherently in ranks, or even if they can be categorised only as plus or minus (more or less, better or worse), they can be treated by non-parametric methods.

Non-parametric methods are available to treat data which are simply classificatory and categorical, i.e., are measured in nominal scale.

Samples made up of observations from several different populations at times cannot be handled by Parametric tests.

Non-parametric statistical tests typically are much easier to learn and to apply than are parametric tests. In addition, their interpretation often is more direct than the interpretation of parametric tests.

1.5 DISADVANTAGES OF NON-PARAMETRIC STATISTICAL TESTS

If all the assumptions of a parametric statistical model are in fact met in the data and the research hypothesis could be tested with a parametric test, then non-parametric statistical tests are wasteful.

The degree of wastefulness is expressed by the power-efficiency of the non-parametric test. It will be remembered that, if a non-parametric statistical test has power-efficiency of say, 90 percent, this means that when all conditions of parametric statistical test are satisfied the appropriate parametric test would be just as effective with a sample which is 10 percent smaller than that used in non-parametric analysis.

Another objection to non-parametric statistical test has to do with convenience. Tables necessary to implement non-parametric tests are scattered widely and appear in different formats (The same is true of many parametric tests too).

1.6 PARAMETRIC STATISTICAL TESTS FOR DIFFERENT SAMPLES

Suppose we wish to measure teaching aptitude of M.A. Psychology Students(LARGE SAMPLE) by using a verbal aptitude teaching test.

It is not possible and convenient to measure the teaching aptitude of all the enrolled M.A. Psychology Students trainees and hence we must usually be satisfied with a sample drawn from this population.

However, this sample should be as large and as randomly drawn as possible so as to represent adequately all the M.A. Psychology Students of IGNOU.

If we select a large number of random samples of 100 trainees each from the population of all trainees, the mean values of teaching aptitude scores for all samples would not be identical.

A few would be relatively high, a few relatively low, but most of them would tend to cluster around the population mean.

The sample means due to 'sampling error' will not vary from sample to sample but will also usually deviate from the population mean. Each of these sample means can be treated as a single observation and these means can be put in a frequency distribution which is known as sampling distribution of the means.

An important principle, known as the 'Central Limit Theorem', describes the characteristics of sample means. According to this theorem, if a large number of equal-sized samples, greater than 30 in size, are selected at random from an infinite population:

The means of the samples will be normally distributed.

The average value of the sample means will be the same as the mean of the population.

The distribution of sample means will have its own standard deviation.

This standard deviation is known as the 'standard error of the mean' which is denoted as SE_M or $ó_M$.

It gives us a clue as to how far such sample means may be expected to deviate from the population mean.

The standard error of a mean tells us how large the errors are in any particular sampling situation.

The formula for the standard error of the mean in a large sample is:

$$SE_M \text{ or } ó_M = ó / \sqrt{N}$$

Where

ó = the standard deviation of the population

N = the size of the sample

In case of **small samples**, the sampling distribution of means is not normal. It was in about 1815 when William Seely Gosset developed the concept of small sample size. He found that the distribution curves of small sample means were some what different from the normal curve. This distribution was named as t-distribution. When the size of the sample is small, the t-distribution lies under the normal curve.

1.7 PARAMETRIC STATISTICAL MEASURES FOR CALCULATING DIFFERENCE BETWEEN MEANS

In some research situations we require the use of a statistical technique to determine whether a true difference exists between the population parameters of two samples. The parameters may be means, standard deviations, correlations etc. For example, suppose we wish to determine whether the population of male M.A. Psychology Students enrolled with IGNOU differs from their female counterparts in their attitude towards teaching… In this case we would first draw samples of male and female M.A. Psychology Students. Next, we would administer an attitude scale measuring attitude towards teaching on the selected samples, compute the means of the two samples, and find the difference between them. Let the mean of the male sample be 55 and that of the females 59. Then it has to be ascertained if the difference of 4 between the sample means is large enough to be taken as real and not due only to sampling error or chance.

In order to test the significance of the obtained difference of 4, we need to first find out the standard error of the difference of the two means because it is reasonable to expect that the difference between two means will be subject to sampling errors. Then from the difference between the sample means and its standard error we can determine whether a difference probably exists between the population means.

In the following sections we will discuss the procedure of testing the significance of the difference between the means and correlations of the samples.

1.7.1 Significance of the Difference between the Means of Two Independent Large and Small Samples

Means are said to be independent or uncorrelated when computed from samples drawn at random from totally different and unrelated groups.

Large Samples

You have leant that the frequency distribution of large sample means, drawn from the same population, fall into a normal distribution around the population mean (M_{pop}) as their measure of central tendency. It is reasonable to expect that the frequency distribution of the difference between the means computed from the samples drawn from two different populations will also tend to be normal with a mean of zero and standard deviation which is called the standard error of the difference of means.

The standard error is denoted by $ó_{dm}$ which is estimated from the standard errors of the two sample means, $ó_{m1}$ and $ó_{m2}$. The formula is:

$ó_{dM} = (ó_{M1}2 + ó_{M1}2)^{\text{Under root}}$

in which

$ó_{M1}$ = SE of the mean of the first sample

$ó_{Mw}$ = SE of the mean of the second sample

N_1 = Number of cases in first sample

N_2 = Number of cases in second sample

1.7.2 Significance of the Differences between the Means of Two Dependent Samples

Means are said to be dependent or correlated when obtained from the scores of the same test administered to the same sample upon two occasions, or when the same test is administered to equivalent samples in which the members of the group have been matched person for person, by one or more attributes.

$T = M_1 + M_2 \ / \ \text{ó}M^2_1 + - 2 \ r_{12} \ \text{ó}M_1 \ \text{ó}M_2$

in which

M_1 and M_2 = Means of the scores of the initial and final testing.

óM_1 = Standard error of the initial test mean.

óM_2 = Standard error of the final test mean.

r_{12} = Correlation between the scores on initial and final testing.

1.7.3 Significance of the Difference between the Means of Three or More Samples

We compute CR and t-values to determine whether there is any significant difference between the means of two random samples. Suppose we have N(N>2) random samples and we want to determine whether there are any significant differences among their means. For this we have to compute F value that is Analysis of Variance.

Analysis of variance has the following basic assumptions underlying it which should be fulfilled in the use of this technique.

The population distribution should be normal. This assumption, however, is not especially important.

Eden and Yates showed that even with a population departing considerably from normality, the effectiveness of the normal distribution still held.

All the groups of certain criterion or of the combination of more than one criterion should be randomly chosen from the sub-population having the same criterion or having the same combination of more than one criterion.

For instance, if we wish to select two groups in a population of M.A. Psychology Student trainees enrolled with IGNOU, one of males and the other of females, we must choose randomly from the respective sub populations. The assumption of randomness is the key stone of the analysis of variance technique. There is no substitute for randomization.

The sub-groups under investigation should have the same variability. This assumption is tested by applying F_{max} test.

F_{max} =Largest Variance / Smallest Variance

In analysis of variance, we have usually three or more groups i.e. there will be three or more variances.

Unless the computed value of F_{max} equals or exceeds the appropriate F critical value at .05 level in the Table N of the Appendix, (Statistics book) it is assumed that the variances are homogeneous and the difference is not significant.

1.8 PARAMETRIC STATISTICS MEASURES RELATED TO PEARSON'S 'r'

The mathematical basis for standard error of a Pearson's co-efficient of correlation 'r' is rather complicated because of the difficulty in its nature of sampling distribution.

The sampling distribution of r is not normal except when population r is near zero and size of the sample is large (N=30 or greater).

When r is high (0.80 or more) and N is small, the sampling distribution of r is skewed. It is also true when r is low (0.20 or less).

In view of this, a sound method for making the inference regarding Pearson's r, especially when its magnitude is very high or very low, is to convert r into Fisher's Z coefficient using conversion table provided in the Appendix (Statistics book) and find the standard error (SE) of Z.

The sampling distribution of Z co-efficient is normal regardless of the size of sample N and the size of the population r. Furthermore, the SE of Z depends only upon the size of sample N.

The formula for standard error of Z ($ó_z$) is:

$$SE_z = 1/\sqrt{N-3}$$

The method of determining the standard error of the difference between Pearson's co-efficient of correlation of two samples is first to convert the r's into Fisher's Z co-efficient and then to determine the significance of the difference between the two Z's.

When we have two correlations between the same two variables, X and Y, computed from two totally different and unmatched samples, the standard error of a difference between two corresponding Z's is computed by the formula:

$$SE_{dz} = ó_{z1-z2} = \sqrt{(1/N_{1-} 3 + 1/ N_{2-} 3)}$$

in which

N_1 and N_2 = sizes of the two samples

The significance of the difference between the two Z's is tested with the following formula:

$$CR = Z_1 - Z_2 / SE_{DZ}$$

1.8.1 Non-parametric Tests Used for Inference

The most frequently used non-parametric tests for drawing statistical inferences in case of unrelated or independent samples are:

1) Chi square test;

2) Median test; and

3) Mann-Whitney 'U' test.

The use and application of these tests are discussed below:

The Chi Square (X^2) Test

The chi square test is applied only to discrete data. The data that are counted rather than measured. It is a test of independence and is used to estimate the likelihood that some factor other than chance accounts for the observed relationship.

The Chi square (X^2) is not a measure of the degree of relationship between the variables under study.

The Chi square test merely evaluates the probability that the observed relationship results from chance. The basic assumption, as in case of other statistical significance, is that the sample observations have been randomly selected.

The formula for chi-square (X^2) is:

$$(X^2) = \sum[(fo-fe)2 / fe]$$

In which

Fo = frequency of occurrence of observed or experimentally determined facts.

Fe = expected frequency of occurrence.

The Median Test

The median test is used for testing whether two independent samples differ in central tendencies. It gives information as to whether it is likely that two independent samples have been drawn from populations with the same median. It is particularly useful when even the measurements for the two samples are expressed in an ordinal scale. In using the median test, we first calculate the combined median for all measures (scores) in both samples. Then both sets of scores at the combined median are dichotomized and the data are set in a 2 x 2 table with two rows one containing below median and the other row containing above median. On the column side we have two columns, one containing the sample 1 and the other column containing sample 2.

The Mann-Whitney U Test

The Mann-Whitney U test is more useful than the Median test. It is one of the most useful alternative to the parametric t test when the parametric assumptions cannot be met and when the measurements are expressed in ordinal scale values.

1.9 SOME NON-PARAMETRIC TESTS FOR RELATED SAMPLES

Various tests are used in drawing statistical inferences in case of related samples. In this section we shall confine our discussion to the use of Sign Test and Wilcoxon Matched-Pairs Signed-Ranks Test Only

The Sign Test

The sign test is the simplest test of significance in the category of non-parametric tests. It makes use of plus and minus signs rather than quantitative measures as its data. It is particularly useful in situations in which quantitative measurement is impossible or inconvenient, but on the basis of superior or inferior performance it is possible to rank with respect to each other, the two members of each pair.

The sign test is used either in the case of single sample from which observations are obtained under two experimental conditions. The researcher wants to establish that the two conditions are different.

The use of this test does not make any assumption about the form of the distribution of differences. The only assumption underlying this test is that the variable under investigation has a continuous distribution.

The Wilcoxon Matched Pairs Signed Ranks Test

The Wilcoxon matched pairs signed ranks test is more powerful than the sign test because it tests not only direction but also the magnitude of differences within pairs of matched groups.

This test, like the sign test, deals with dependent groups made up of matched pairs of individuals and is not applicable to independent groups. The null hypothesis would assume that the direction and magnitude of pair difference would be about the same.

1.10 LET US SUM UP

Parametric and non-parametric tests are important for students especially researchers working in any field. Parametric tests include all methods of statistics when the sample size is large whereas in non-parametric test the sample size is small. There are some advantages and disadvantages of both the tests. In this unit we discussed the statistical inference based on parametric tests. It included the assumptions on which the use of parametric tests are based; inferences regarding means of large and small samples; significance of the difference between the means of two large and small independent samples; significance of the difference between means of the two dependent samples; significance of the difference between means of three or more samples; significance of Pearson's coefficients of correlation; and significance of the difference between Pearson's coefficients of correlation of two independent samples. F test is used for testing the significance between the means of three or more samples. It involves the use of analysis of variance or analysis of co-variance. For testing the significance of Pearson's r, we make use of Fisher's Z transformation or t-test.

1.11 UNIT END QUESTIONS

1) Define parametric statistics.
2) Discuss non-parametric statistics?
3) Write various assumptions of parametric statistics?
4) What are the advantages of non-parametric statistics?
5) Differentiate between parametric and non-parametric statistics?
6) List the assumptions on which the use of Parametric Tests is base.
7) Describe the characteristics of Central Limit Theorem.
8) Define the standard error of mean.

1.12 GLOSSARY

Statistics : Measurement which are associated with sample

Parameters : Measurements which are associated with population

Assumptions	:	Prerequisite conditions
Population	:	Larger group of people to which inferences are made.
Sample	:	Small proportion of the population which we assert representing population.
Normal Curve	:	Bell shaped frequency distribution that is symmetrical and unimodel.
Distribution free tests	:	Hypothesis – testing procedure making non assumptions about population parameters.
Categorical Scale	:	Variable with values that are categories that is, they are name rather than numbers.
Test	:	Test is a tool to measure observable behaviour
Homosedasity	:	Populations must have some variance or in special cases must have a known ratio of variance.

1.13 SUGGESTED READINGS

Asthana H.S, and Bhushan. B. (2007) *Statistics for Social Sciences* (with SPSS Applications).

B.L. Aggrawal (2009). *Basic Statistics*. New Age International Publisher, Delhi.

Guilford, J.P. (1965); *Fundamental Statistics in Psychology and Education*. New York: McGraw Hill Book Company.

Siegel, S. (1956): *Non-parametric Statistics for Behavioural Sciences*. Tokyo: McGraw Hill Hoga Kunsa Ltd.

Sidney Siegel, & N. John Castetellan, Jr. (1958) *Non-parametric Statistics for the Behavioural Science*. McGraw Hill Books company, New Delhi