



Dr S Amaladoss Xavier, SJ Assistant Professor in Physical Science at St Xavier's College of Education, Palayamkottai, TamilNadu. He has published more than thirty research articles in both National and International Journals and presented thirty five papers in National and International seminars. He has edited two International conference and two National conference proceedings. He has guided forty M. Ed., and 30 M.Phil. scholars. Seven of his PhD scholars have been awarded the degree. Six PhD scholars are waiting for the viva and six are doing PhD under his guidance. His areas of interest are research, soft skills and educational psychology.

Mr. A. John Lawrence is an Associate Professor in English, at St. Xavier's College of Education (Autonomous), Palayamkottai. The experiences gained in teaching Special and General English and Educational Psychology for more than 16 years, 4 years in R.C.T.T.I, Trichy, and 2 years abroad have turned him an effective teacher. He holds M.A. and M. Phil. in English, M.Sc. and M. Phil. Psychology, Post Graduate Diploma in Guidance and Counselling, Post Graduate Diploma in Methodology of Teaching English, Certificate in Teaching of English and is pursuing his research in ELT- Education.



Rev. Dr. I. Jesudoss, sj M.Sc(Maths); M.A(Eng); M.Ed; M.Sc(PSY); MBA; NET; SLET; Ph.D., Director of SXCE Library and Fr. Bohnure Research Center. Expert in NET-coaching programmes, Soft Skill Development for Teacher Training students, Thesis writing and Basic Educational Research for M.Ed. students.

Dr. M. Antony Raj is an Assistant Professor in History Education at St. Xavier's College of Education (Autonomous), Palayamkottai. He holds M.A., M.Phil. (His), M.Ed, M.Phil., NET, Ph.D. (Edn.). He has 17 years of teaching experience in school and 10 years teaching experience in college of education. His specialized areas are History Education, Teacher Education and Education for Special Children. He is guiding M.Ed., M.Phil. and Ph.D. scholars.

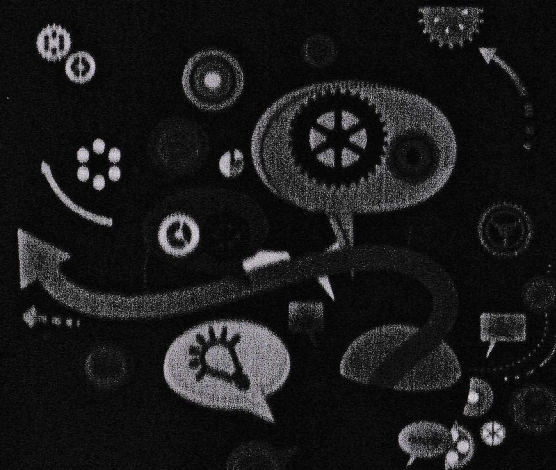


Dr. A. Punitha Mary has been working as Assistant Professor in Education at St. Xavier's College of Education (Autonomous) since 2005. She had organised two UGC sponsored National Level Seminar. She guides M.Ed. M.Phil., Scholars every year. She is handling educational research for M.Ed. scholars

Dr. A. Michael J Leo has been working as Assistant Professor in Education at St. Xavier's College of Education (Autonomous), since 2006. His areas of interests are Educational Statistics, Teacher Education, Educational Psychology, Educational Technology and Soft Skills. He had guided more than 25 M.Ed. scholars and organised two UGC sponsored seminars. He edited more than 5 monograph books with ISBN and published more than 8 articles in National and International Journals with ISBN numbers.



# FUNDAMENTALS ON EDUCATIONAL RESEARCH



EDITOR: REV. DR S. AMALADOSS XAVIER S.J.

PUBLISHED BY

LEARNING CENTRE FOR RESEARCH

BY XAVIER'S COLLEGE OF EDUCATION (AUTONOMOUS)


It is accredited (3<sup>rd</sup> Cycle) at 'A' Grade by NAAC with CGPA 3.67

PALAYAMKOTTAI - 627 002

*Handwritten signature*  
Principal  
Xavier's College of Education  
(Autonomous)  
Palayamkottai - 627 002

## CONTENT

Sl. No.	Title	Page No.
1.	Introduction to Research and Research problems	1
2.	Review of Related Studies	10
3.	Formulation of Objectives and Hypotheses	19
4.	Population and Sample	24
5.	Preparation of a Research Tool	43
6.	Quantitative Data Analysis	53
7.	Research Report	72

  
Principal  
St. Xavier's College of Education  
(Autonomous)  
Palayamkottai - 627 002

## I. INTRODUCTION TO RESEARCH AND RESEARCH PROBLEMS

Rev. Dr. I. Jesudoss, SJ.

### Introduction

Research is considered to be more formal, systematic, intensive process of carrying on the scientific methods of analysis. Research is literally speaking, a kind of human behavior, and an activity in which people are engaged. It involves a more systematic structure of investigation, usually resulting in some sort of formal record of procedures and conclusion. It is also the activity of collecting information in an orderly and systematic fashion. Research is the voyage of discovery. It is the quest for answers to unsolved problems.

Research is required in any field to come up with new theories to modify, accept, or nullify the existing theory. According to Advanced Learner's Dictionary, Research is 'a careful investigation or inquiry specially through search for new facts in any branch of knowledge'. Redman and Mory define Research as, 'a systematized effort to gain new knowledge.'

### Objectives of Research

Research attempts to achieve new insights into the problem and discover facts which are existing in the world. It invents new solutions to the problems. Research highlights the significance of the phenomenon under study and develops new tools, concepts and theories for a better study. Also it tests the hypothesis of casual relationship between variables and writes the report based upon the findings of the research. Research aims, at planning and thus contributes to national development.

### Meaning of Educational Research

Educational Research is cleansing of educational process. Many experts think that Educational Research is the systematic application of scientific method for solving educational problems. In education, teachers, administrators, scholars or others engage in educational

## ii) Conduct face-to-face interviews

This can be a good way to ensure that you are reaching your target demographic and can reduce missing information in your questionnaires, as it is more difficult for a respondent to avoid answering a question when you ask it directly.

## iii) Try using the telephone

While this can be a more time-effective way to collect your data, it can be difficult to get people to respond to telephone questionnaires.

## iv) Include a deadline

Ask your respondents to have the questionnaire completed and returned to you by a certain date to ensure that you have enough time to analyze the results.

## v) Make your deadline reasonable

Giving respondents up to two weeks to answer should be more than sufficient. Anything longer and you risk your respondents forgetting about your questionnaire.

## vi) Consider providing a reminder

A week before the deadline is a good time to provide a gentle reminder about returning the questionnaire. Include a replacement of the questionnaire in case it has been misplaced by your respondent.

## Conclusion

Question designing remains primarily a matter of common sense and experience and of avoiding known pitfalls, as there are no hard and fast rules relating to it. Hence alternative versions of questions must be rigorously tested in pre-tests. Test-revision-retests play a crucial role in questionnaire construction. Every researcher must remember that the knowledge of procedures involved in research is perhaps the most important tool because without this, the best computers and most reliable tests would be useless.

## VI. QUANTITATIVE DATA ANALYSIS

Dr. A. Michael J. Leo

Data analysis in any educational research is an important task of the researcher. There are quantitative and qualitative analyses depending upon the nature of the methodology. This chapter explains precisely the process of quantitative data analysis in educational research. The following are the steps involved in the process of quantitative data analysis.

1. Data Collection
2. Data Entry
3. Analysis of Data

Dr. A. Michael J. Leo  
Assistant Professor in Education  
St. Xavier's College of Education (Autonomous)  
Palayamkottai - 627 002.

### 1. Data Collection

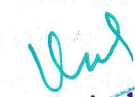
Data collection methods in educational research are used to gather information that is then analyzed and interpreted. As such, data collection is a very important step in conducting research and can influence results significantly. Once the research question and sources of data are identified, appropriate methods of data collection are determined. Data collection includes a broad range of more specific techniques.

Historically, much of the data collection performed in educational research depended on methods developed for studies in the field of psychology, a discipline which took what is termed a 'quantitative' approach. This involves using instruments, scales, tests, structured observation and interviewing. As contemporary educational researchers also draw from fields such as business, political science and medicine, the data collection in education has become a multidisciplinary phenomenon.

### Need for Data Collection

From the following points, a researcher can understand the need for data collection in educational research.

- i) It provides a definite direction to a research inquiry.

  
Principal  
St. Xavier's College of Education  
(Autonomous)  
Palayamkottai - 627 002

- ii) The data are needed to substantiate the various arguments in research findings.
- iii) It is to test and verify the hypotheses.
- iv) Through data collection, one can generalize his results.
- v) The data collection will be helpful to formulate new theory or principles.
- vi) Data are necessary to provide solution for the problem.

## 2. Data Entry

After the data collection, the data entry process may be done in three steps as follows

- a) Scoring of the Data
- b) Coding of the Data
- c) Entering of the Data

### a. Scoring of the Data

After or during data collection, the scoring of the tools can be done. By following the scoring procedure prescribed by the author(s) of the tool or designed by the investigator, the numerical values could be allotted to the items of the tools. The item-wise or dimension-wise scoring of the tool will be helpful for the investigator to adopt additional data analysis.

### b. Coding of the Data

Any instrument or tool for data collection in educational survey will have two parts,

- (i) Personal data sheet according to the sample group.
- (ii) Actual instrument what it measures.

A sample personal data sheet may have the following information (For Instance Teacher Educators are the target group).

#### Personal Data Sheet

It includes general information about the respondents regarding name of the college of education, age, gender, marital status, locality of college, nature of college, account in social network, access of social

network, teaching experience, hobbies, subject and so on. After the data collected, the coding can be done as follows,

- i) Allot codes (Avoid allotting '0') to each demographic variable in the personal sheet (Ex: Male - 1 and Female - 2 or Female - 1 and Male -2).
- ii) Verification of data given by the respondent in the data sheet.
- iii) The missing responses from the respondents may be collected well in advance, before the data entry.
- iv) Note the code in the personal data sheet in an order.
- v) Entering the codes in the Excel Sheet or in the SPSS.

For the questionnaires, inventories and rating scales, the item-wise or dimension-wise or total scores (When no dimension) could be entered in the tool.

### c. Entering of the Data

Entering data correctly at the first time can help us to avoid problems later on and make it easier to use a number of tools in MS Excel and features such as formulas and charts.

Further it will help us to export to any other software to analyse the data. When we enter the data in the MS excel, the following points may be considered.

#### The DOs and DON'Ts

- a. Plan the spreadsheet in terms of number of rows and columns to be utilized.
- b. Don't use numbers as column headings and don't include units with the data.
- c. Kindly use the strings as headings that show the column information.
- d. Do use cell references and named ranges in formulas.
- e. Don't leave cells containing formulas unlocked.
- f. Don't leave blank rows or columns when entering related data.
- g. Do save frequently and save in two places.  
After entering the data in the cells of MS Excel, the data analysis could be done as follows.

### 3. Analysis of Data Data

are collected for both variables as well as attributes. These are gathered in terms of frequency and scores. It depends on the type of instrument employed for its measurement. Generally tests yield the data in the form of scores and questionnaires provide the data in the form of frequency.

#### *Difference between Facts and Data*

The facts and data have been distinguished in the following manner,

- a. The facts are organized in the original form whereas the data are organized in a systematic order.
- b. The facts are difficult to interpret. But the data can be interpreted easily.
- c. The facts are mysterious in nature but data have no mystery.
- d. The facts are descriptive in nature whereas data are explanatory.
- e. The facts are collected in historical or survey research whereas data are gathered in a scientific and experimental study.
- f. The facts may not be directly the basis of findings or research conclusions, but the data are directly linked with research conclusions.
- g. The facts are not amenable to objective statistical treatments whereas data can be easily subjected to objective statistical treatments.

#### **Nature of Data**

The data can be classified into two broad categories.

1. Qualitative Data or Attributes
2. Quantitative Data or Variables

#### *Qualitative Data*

The characteristics or traits for which numerical value cannot be assigned are called attributes, E.g. Motivation. The responses to open-ended questions of a questionnaire or a schedule, first hand information

from people about their experiences, ideas, beliefs, etc. and selected content or excerpts from documents, case histories, personal diaries and letter are other examples of qualitative data.

#### *Quantitative Data*

Data are classified on the basis of the characteristics possessed by the different group of units of a universe. When the data are classified on the basis of qualities, which are incapable of direct measurement, the classification is said to be according to the attributes. When the data are classified on the basis of quantitative measurement, the classification is said to be according to class intervals. The classification can be,

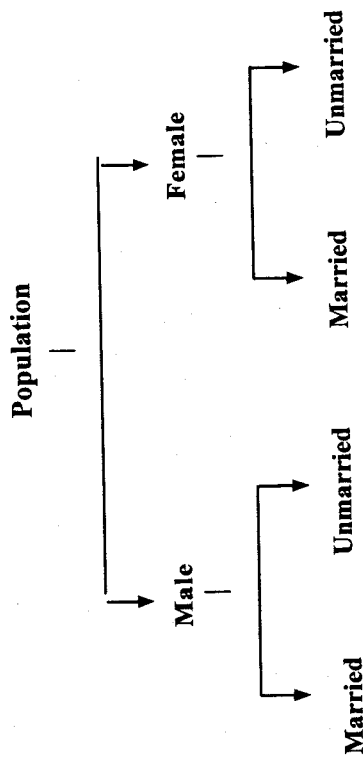
- a. Geographical (Village, Town, Districts).
- b. Chronological (Time Basis).
- c. Qualitative (Attributes).
- d. Quantitative (Magnitude).

#### *Qualitative Classification*

In qualitative classification, data are classified on the basis of some attributes or qualities such as gender, colour, religion, etc. The point to note, in this attribute under study cannot be measured, one can only find out how many persons are living in urban area and how many in rural area. When only one attribute is studied, two classes are formed namely classes possessing the attributes and not possessing the attribute. This type of classification is known as simple classification. For an example, a population is divided into two categories as Rural and Urban.

The type of classification where two classes are formed is called *2-fold or dichotomies classification*. The type of classification where several classes are formed is called *manifold classification*.

**Figure 1**



E.g.

**Quantitative Classification**

This type of classification is applicable in the case, where direct quantitative measurement of the data is possible. Eg. Height, weight etc.

E.g. The students of the college may be classified according to their weight. Such distribution is empirical distribution. There are two elements

- a. Variable : The weight of the students.
- b. The frequency : The number of students in each class.

Term variable refers to the distribution. A variable may be either continuous or discrete. A continuous variable is capable of manifesting every conceivable fractional value within the range of product. In continuous variable data are obtained by numerical measurement rather than continuous. E.g. When a student grows say from 90 cm to 150 cm his height pass through all values between these limits.

Discrete variable is that which can vary only by finite 'jumps' cannot manifest every conceivable fractional value. E.g. Number of rooms in a house values should be 1, 2, 3 etc.

**Continuous Variable**

Continuous variable takes the real value (even in decimal points). It aspects the numerical value within a range. With this type of data,

one can develop more and more accurate measurements depending on the instrument used. A continuous variable is that which can assume any numerical value within a specific range.

E.g.

- Height in centimetres (2.7 cm or 2.746 cm or 2.743216 cm).
- Temperature in degrees Celsius (39.21° C or 39.209999° C).

**Discrete Variable**

Discrete variable is the variable which takes the integer value only; these are variables which can have full values. A variable for which the individual values fall on the scale only with distinct gap is called discrete variable.

E.g.

- Number of patients visits the park (0, 1, 2, 3, 4, etc).
- Number of books in the Library (0, 1, 2, 3, 4, 5, etc.)

**Dependent Variable**

The dependent variables are the conditions that appear, disappear or change as the experimenter introduces, removes or changes independent variable. The dependent variable may be a test score, the number of errors or speed in performing a task. Thus the dependent variables are the measured changes in pupil performance attributable to the influence of the independent variables.

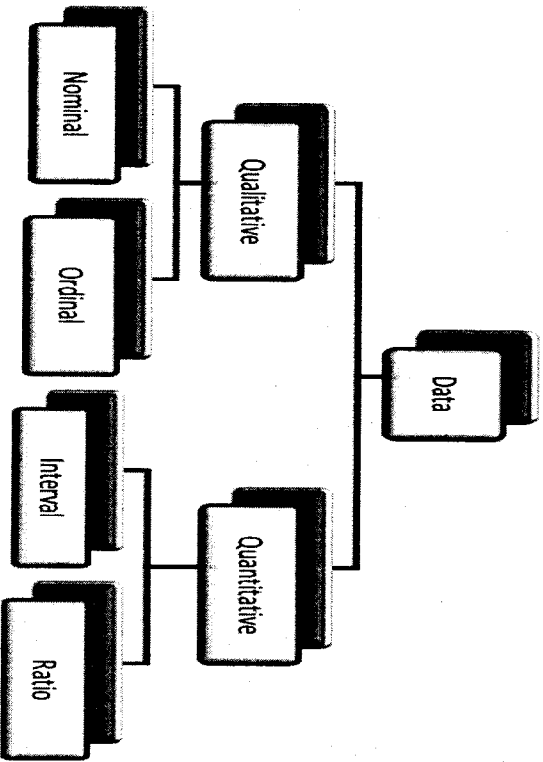
**Independent Variable**

The independent variable are the conditions that the experimenter manipulates, controls or observes in the investigator's attempt to ascertain conditions that he manipulates in his attempts to ascertain their relationship to observed phenomena. The experimenter constructs the experiment in which he attempts to control all conditions except the independent variables which he manipulates. In educational research, an independent variable may be particular teaching method, a type of teaching material, a reward or a period of exposure to particular conditions.

## Scales of Measurement

Measurement is defined as the assignment of numbers to objects and events according to logically accepted values. Or measurement has been defined as a numerical method of describing observations of materials or characteristics. When a defined portion of the material or characteristic is used as a standard for measuring any sample, a valid and precise method of data description is provided. A convenient way of classification of the measurement levels is made by Stevens, S.S. (1951) in his book Scales of Measurement. According to this classification there are four levels of measurement scales.

Figure 2  
Classification of Data and Scale



### 1. Nominal Scale

A nominal scale is the least precise method of measurement. Nominal scales of measurement are used when a set of objects among two or more categories are to be differentiated on the basis of qualitative differences. A nominal scale describes differences between things by assigning them to categories-such as Business man, Professional, Farmer, and Coolie and to subsets such as male or female.

Nominal data are counted data. Each individual can be a member of only one category and all the members of the category have the same defined characteristics such categories as nationality, gender, socio economic status, race, occupation or religious affiliation provides examples. Nominal scales are non orderable and only arithmetical operation applicable to such scales is counting.

### 2. Ordinal Scale

The ordinal scale of measurement corresponds to quantitative classification of a set of objects. Sometimes it is possible to indicate not only that things differ, but that they differ in amount or degree. Ordinal scales permit the ranking of items or individuals from highest to lowest the ranking of students in class for height, weight, or scholastic achievement are the examples of ordinal scale of measurement. The criterion for highest to lowest ordering is expressed as relative position or rank in a group: 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>, ..., n<sup>th</sup>. Ordinal measures have no absolute values, and the real differences between adjacent ranks may not be equal. Ranking spaces them equally, though may not actually be equally spaced.

### 3. Interval Scale

An arbitrary scale based on equal units of measurements indicates how much of a given characteristic is present. The difference in amount of the characteristic possessed by persons with scores of 95 and 96 is assumed to be equivalent to that between persons with scores of 67 and 68 (In the interval scale, the differences between consecutive numbers on the scale are equal over the entire scale but there is no true zero point on it). The interval scale represents a decided advantage over nominal and ordinal scales because it indicates the relative amount of a trait or characteristic. Its primary limitation is the lack of a true zero. It does not have the capacity to measure the complete absence of the trait, most psychological test and inventories are based on interval scales.

### 4. Ratio Scale

A ratio scale has the equal interval properties of an interval scale but has two additional features.

- i. The ratio scale has a true zero. It is possible to indicate the complete absence of a property. For an example the zero point on a meter scale indicates the complete absence of length or height.
- ii. The numerals of the ratio scales have the qualities of real numbers and can be added, subtracted, multiplied, and divided and expressed in ratio relationships. For example 5 grams is one-half of 10 grams, 15 grams is three times 5 grams, and on a laboratory weighing scales two 1 gram weights will balance a 2 gram weight.

Proceeding from the nominal scale (the least precise) to ratio scale (the most precise) increasingly relevant information is provided. If the nature of the variables permits, the scale that provides the most precise description should be used.

In quantitative approach of educational research, the data could be analyzed according to the objectives framed and the nature of the problem undertaken for the investigation. In general, the quantitative data analysis may be classified as follows:

- a) Descriptive Statistical Analysis
- b) Inferential Statistical Analysis

**a) Descriptive Statistical Analysis**

Statistics is the science of making effective use of numerical data relating to groups of individuals or experiments. It deals with all aspects of this, including not only the collection, analysis and interpretation of such data, but also planning of the collection of data, in terms of the design of surveys and experiments. In descriptive analysis, the measures of central tendency and dispersion are calculated and the data is presented accordingly, in graphical and tabular form.

After the conversion of the raw scores into T-scores (if needed), generally in descriptive analysis, the levels are fixed by the mean and standard deviation. Here the researcher attempts to divide the sample into three subgroup based on the measures of central tendency and dispersion in terms of the performance or outcome what the study intended to measure. In standardized and validated instruments, the given standard scores by the author(s) of the instrument could be used to fix the levels.

**Levels**

After calculating the mean and standard deviation, the scores were corrected into T scores

$$T = 50 + 10 Z$$

Where,

$$Z = (\text{Raw Score} - \text{Mean}) / \text{Standard Deviation}$$

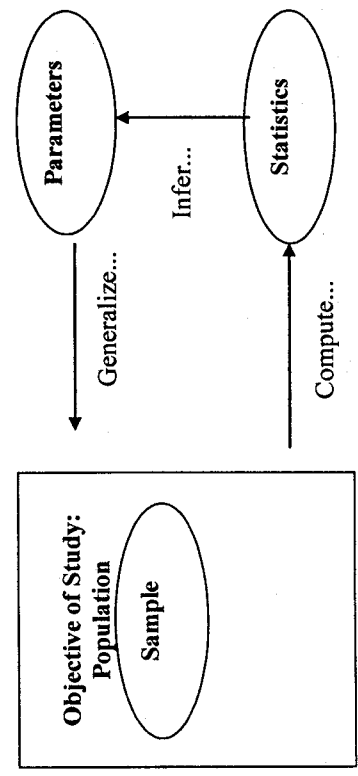
The levels could be fixed by the guidelines in the manual of the tools used or as follows.

**Table 1**  
**Fixing the levels**

Score	Levels
Above Mean + 1 S.D	High
Between Mean -1 S.D and mean +1 S.D	Average/Moderate
Below Mean -1 S.D	Low

**b. Inferential Statistical Analysis**

Inferential statistical analysis is using sample information to draw inferences or test hypotheses about a characteristic of a population. It helps to make inductive generalizations from the particular (the sample) to the general (the population). The generalization is done through the process of hypotheses testing and estimation. This process is explained as follows





## Hypotheses Testing

Many statistical tests are based on the assumption of normality. The statistical tests may be classified into two categories namely Parametric and Non-Parametric. According to the scales of measurement the appropriate statistical test is selected. Inferential statistics is performed for testing of hypothesis.

The method in which the researcher selects samples to learn more about characteristics in a given population is called hypothesis testing. Hypothesis testing is really a systematic way to test claims or ideas about a group or population.

*The following are the steps involved in hypothesis testing:*

### 1. Making Assumptions

In hypothesis testing, the researcher makes the assumptions about the level of measurement of the variable, the sampling method, the shape of the population distribution, and the sample size. In the following example, the assumptions are:

- a) Sample collection : Random Sampling.
- b) The scale of measurement of the variable under this study is interval.
- c) The size of the sample assumes a normal population.
- d) The distribution of the sample is normal.

### 2. Stating the Hypotheses

The null hypothesis ( $H_0$ ) is a statement about a population parameter, such as the population mean, that is assumed to be true. The null hypothesis is a starting point. The researcher will test whether the null hypothesis is likely to be true. An alternative hypothesis ( $H_1$ ) is a statement that directly contradicts a null hypothesis. Two tailed test is applied in the case of Null hypothesis.

### 3. Set the Criteria for a Decision

To set the criteria for a decision, the researcher states the level of significance for a test. Level of significance, or significance level, refers to a criterion of judgment upon which a decision is made regarding the value stated in a null hypothesis. The criterion is based on the probability of obtaining a statistic measured in a sample if the value stated in the null hypothesis is true.

In behavioural science, the criterion or level of significance is typically set at 5%. When the probability of obtaining a sample mean is less than 5% if the null hypothesis is true, then the investigator can reject the null hypothesis. The alternative hypothesis determines whether to place the level of significance in one or both tails of a sampling distribution.

### 4. Computing the Test Statistic

A test statistic tells us how far, or how many standard deviations, a sample mean is from the population mean. The larger the value of the test statistic, the further the distance, or number of standard deviations, a sample mean is from the population mean stated in the null hypothesis. The value of the test statistic is used to make a decision.

The test statistic is a mathematical formula that allows researchers to determine the likelihood of obtaining sample outcomes, if the null hypothesis is true. The value of the test statistic is used to make a decision regarding the null hypothesis.

### 5. Making a Decision

The decision to reject or retain the null hypothesis is called significance. When the researcher uses the value of the test statistic to make a decision about the null hypothesis, the decision is based on the probability of obtaining a sample mean, given that the value stated in the null hypothesis is true. If the probability of obtaining a sample mean is less than 5% when the null hypothesis is true, then the decision is to reject the null hypothesis. If the probability of obtaining a sample mean is greater than 5% when the null hypothesis is true, then the decision is to retain the null hypothesis.

Because the researcher observes a sample and not an entire population, it is possible that a conclusion may be wrong. There are four decision alternatives regarding the truth and falsity of the decision a researcher makes about a null hypothesis:

- i. The decision to retain the null hypothesis could be correct.
- ii. The decision to retain the null hypothesis could be incorrect.
- iii. The decision to reject the null hypothesis could be correct.
- iv. The decision to reject the null hypothesis could be incorrect.

Type I error is the probability of rejecting a null hypothesis that is actually true. Researchers directly control the probability of committing

this type of error. An alpha ( $\alpha$ ) level is the level of significance or criterion for a hypothesis test. It is the largest probability of committing a Type I error that a researcher will allow and still decide to reject the null hypothesis.

The incorrect decision is to retain a false null hypothesis. This decision is an example of a Type II error, or  $\beta$  error. With each test the researcher makes, there is always some probability that the decision could be a Type II error. In this decision, the researcher decides to retain previous notions of truth that are in fact false. While it's an error, he still did nothing; he retained the null hypothesis. He can always go back and conduct more studies. It is explained in the following table.

**Table 2**  
**Type I and Type II Error**

		Decision	
		Retain the null	Reject the null
Truth in the population	True	CORRECT $1 - \alpha$	TYPE I ERROR $\alpha$
	False	TYPE II ERROR $\beta$	CORRECT $1 - \beta$ POWER

### 6. Interpretation of Results

Based on the result or decision in the previous step the results could be interpreted. In general, the researcher should always 'translate' the statistics into some understandable form for the reader. Remember that inferential statistics can never "prove" anything. The researcher should think of statistics as a body of evidence (much like a fingerprint at a crime scene) that provides support for the argument. Sometimes it can be used as primary evidence or sometimes it is used in a more supporting role. Graphs are excellent alternatives to tables, and they are used by virtually everyone.

### Parametric and Non-Parametric Tests

Based on the scales of measurement, the statistical tests to verify the hypotheses are classified into two major categories namely

Parametric and Non-parametric tests which are given as follows.

**Table 3**  
**Difference between Parametric and Non-Parametric Tests**

Sl. No.	Parametric Tests	Non-Parametric Tests
1	If the information about the population is completely known by means of its parameters then the statistical test is called parametric tests.	If there is no knowledge about the population or parameters, but still it requires testing the hypothesis of the population. Then it is called non-parametric tests.
2	Specific assumptions are made regarding the population	No assumptions are made regarding the population
3	Null Hypothesis is made on the parameters of the population distributions	Null Hypothesis is free from parameters
4	The test statistic is based on the distribution	The test statistic is arbitrary
5	It is applicable for only variables	It is applicable for variables and attributes
6	It is applicable for Ratio and Interval Data	It is applicable for Nominal and Ordinal Data
7	It is based on sampling theory	It is free from sampling theory
8	t test, Paired t test, ANOVA, Product moment Correlation	Chi-square tests, Mann-Whitney Test, Wilcoxon Signed-Rank Test, Kruskal-Wallis Test and the Friedman Test and Rank Correlation

The following table explains the appropriate statistical tests the researcher could select, while testing the hypotheses.

**Table 4**  
**Appropriate Statistical Tests**

Analysis Type	Parametric Procedure	Non-Parametric Procedure
Compare means between two distinct/independent groups	Two-sample t-test	Wilcoxon rank-sum test
Compare two quantitative measurements taken from the same individual	Paired t-test	Wilcoxon signed-rank test
Compare means between three or more distinct/independent groups	ANOVA	Kruskal-Wallis test
Estimate the degree of relationship between two quantitative variables	Pearson coefficient of correlation	Spearman's rank correlation
Estimate association between two quantitative variables		Chi-Square test

**A Few Statistical Techniques**

**i. 't' test**

't' test or test of significance of the difference between means for large independent samples (Garret, 1969) is used to compare the means between any two groups on any of the variables. If the 't' value is below a cut-off point at 5% level, the differences in means is considered as not significant and the null hypothesis is accepted. When 't' value exceeds a cut-off point at 5% level, then the difference side is considered to be significant and null hypothesis is rejected.

**ii. ANOVA**

Analysis of Variance (ANOVA) is an extremely useful technique for testing the difference between the means of multiple independent samples. The basic principle for ANOVA is to test the difference among the means of samples by examining the amount of variation between the samples relative to the amount of variation among samples. The value of ANOVA is compared in the 'F' limit for given degrees of freedom at 5% level. If the 'F' value worked out is equal or exceeds the 'F' limit value from the table indicated, then there are significant differences among the samples between the means.

**Post ANOVA test - Waller-Duncan test**

Duncan's multiple comparison procedure based on the Standardized range test. For tests that are used for detecting homogeneity subsets of means, non-empty group means are sorted in ascending order. Means that are not significantly different are included together to form a homogeneity subset. The significance for each homogeneity subset of means is displayed. For unequal sample sizes, the harmonic mean  $n_h$  is used instead of  $n$ .

**iii. Chi-Square test**

Chi-square distribution ( $\chi^2$ -distribution) with  $k$  degrees of freedom is the distribution of a sum of the squares of  $k$  independent standard normal random variables. It is one of the most widely used inferential statistics. It is used for goodness of fit on an observed distribution to a theoretical one of qualitative data. The value of  $\chi^2$  is compared in the ' $\chi^2$ ' limit for given degrees of freedom at 5% level. If the ' $\chi^2$ ' value worked out is equal or exceeds out the ' $\chi^2$ ' limit value from the table indicated, then there are significant association between the samples.

**iv. Pearson Product Moment Correlation**

Pearson ' $\gamma$ ' is used for estimating the extent of relation existing among different variables taken in pairs for all the different groups. Garret (1973) presents the following classification for interpreting the various values of ' $\gamma$ ', which is adopted for the study.

- $\gamma$  from 0.00 to +/- 0.20      denotes negligible correlation.
- $\gamma$  from +/-0.20 to +/- 0.40      denotes low correlation.
- $\gamma$  from +/-0.40 to +/- 0.70      denotes substantial correlation.
- $\gamma$  from +/-0.70 to +/- 1      denotes high correlation.

The correlation is interpreted only after the statistical significance of co-efficient correlation is considered from the tables.

**v. Multiple Correlations**

Multiple correlations are used for estimating the inter-correlations among independent variables as well as to their correlations with the dependent variable. The co-efficient of multiple correlation indicates the strength of relationship between one variable and two or more other variables taken together.

#### vi. Factor Analysis

Harman (1960) defines the procedure of factor analysis as follows; "The principal concern of factor analysis is the resolution of a set of variables linear in terms of a smaller number of categories or 'factors'. This resolution can be accomplished by the analysis of correlation among the variables. A satisfactory solution will yield factors, which convey all the essential information of the original set of variables. Thus, the chief aim is to attain scientific parsimony or economy of description".

- Factor analysis is a statistical technique to study the inter-relationships among the variables in an effort to find a new set of factors, fewer in number than the original variables so that the factors are common among the original variables.
- There is difference between factor analysis and principal component analysis.
- In principal component analysis the components are so selected that they can explain maximum variation in the original data set.
- In factor analysis a small number of common factors are extracted so that these common factors are sufficient to study the relationships of original variables.

#### Aims of Factor Analysis

The following points could be considered as the aims of factor analysis:

- Factor analysis helps the researcher to reduce the number of variables to be analyzed, thereby making the analysis easier.
- To identify the hidden dimensions or construct which may not be apparent from direct analysis.
- To identify relationships between variables.
- It helps in data reduction.
- It helps the researcher to cluster the product and population being analyzed.
- Using Factor Analysis, the researcher can reduce the large number of variables into a few dimensions called factors that summarize the available data.
- It aims at grouping the original input variables into factors which underlie the input variables.

- For example, age, gender, marital status can be combined under a factor called demographic characteristics. The income level, education, employment status can be combined under a factor called socioeconomic status.

*Guilford (1973) outlines the different steps in factor analytic study in the following terms,*

- a. Select appropriate domain for investigation.
- b. Develop a hypothesis concerning the factor.
- c. Select or construct suitable tests.
- d. Select suitable population.
- e. Obtain a sample of adequate size.
- f. Extract factors with communalities in the diagonal cells of the correlation matrix.
- g. Rotate the reference axes.
- h. Interpret the rotated factors.