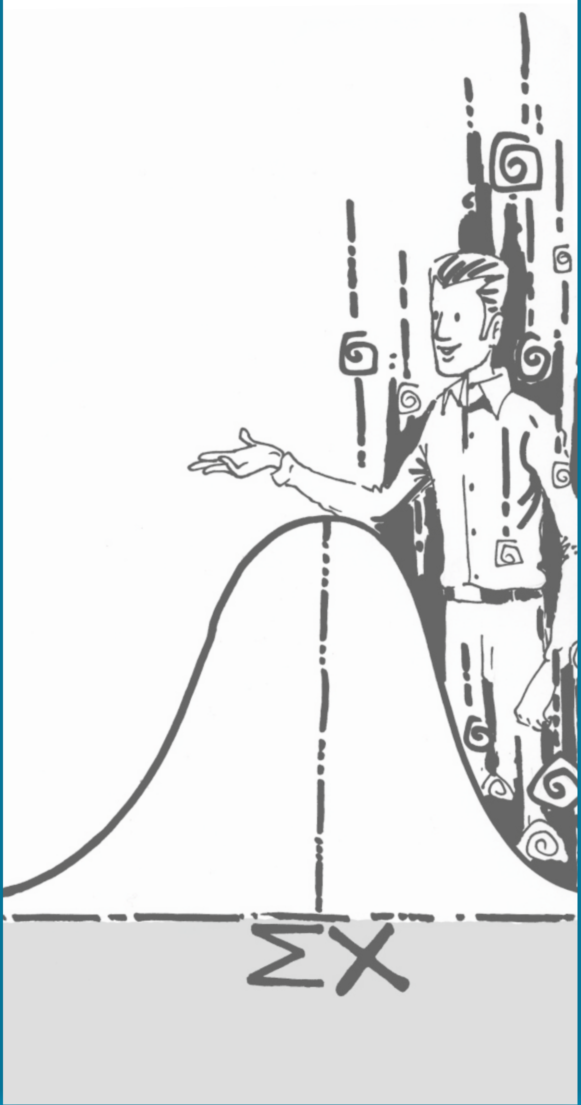
Module 14



Basic Statistics in Guidance and Counselling-II

Department of Educational Psychology and Foundations of Education

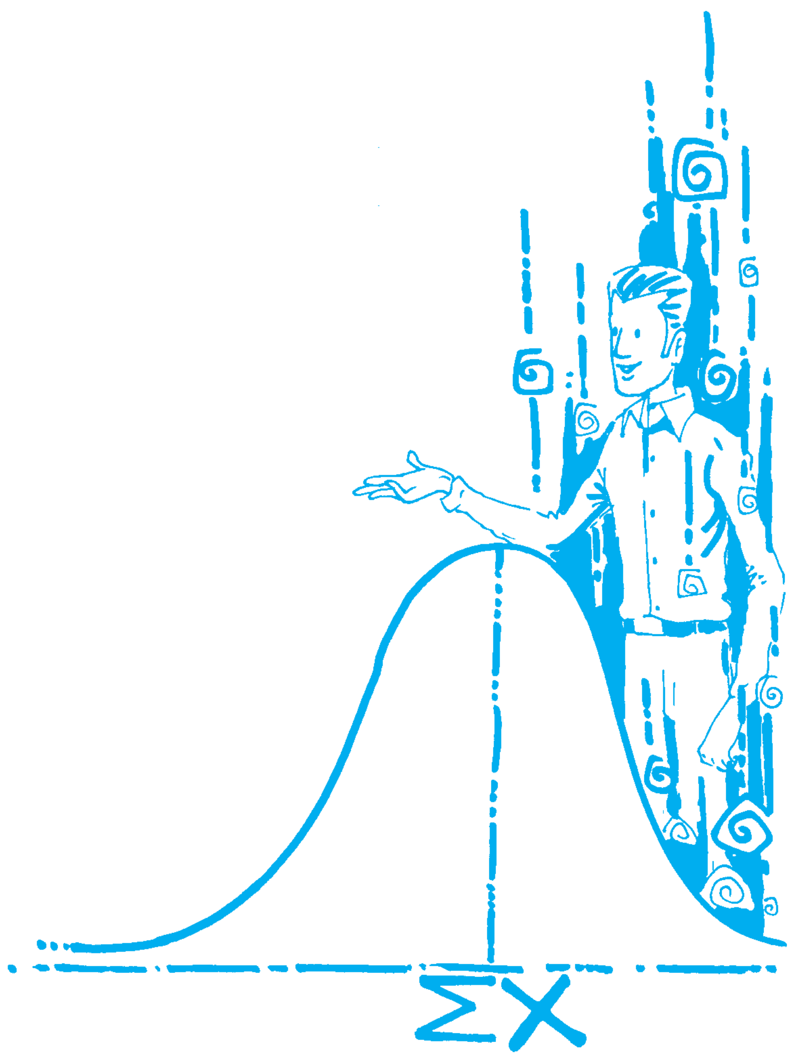
National Council of Educational Research and Training

"When you can measure what you are speaking about and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of the meager and unsatisfactory kind."

— LORD KELVIN

Basic Statistics in Guidance and Counselling-II

Module 14



राष्ट्रीय शैक्षिक अनुसंधान और प्रशिक्षण परिषद्
NATIONAL COUNCIL OF EDUCATIONAL RESEARCH AND TRAINING

First Edition
May 2009 Jyaishtha 1931

Reprinted
January 2017 Pausha 1938
January 2021 Pausha 1942

PD 1T RPS

© National Council of Educational Research and Training, 2009

₹ 60.00

Printed on 80 GSM paper

Published at the Publication Division by the Secretary, National Council of Educational Research and Training, Sri Aurobindo Marg, New Delhi 110 016 and printed at Sagar Offset Printer India (P.) Ltd., 518, Ecotech III, Udyog Kendra II, G.B. Nagar, Greater Noida (U.P.)

ISBN 978-81-7450-956-7

ALL RIGHTS RESERVED

- No part of this publication may be reproduced, stored in a retrieval system or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without the prior permission of the publisher.
- This book is sold subject to the condition that it shall not, by way of trade, be lent, re-sold, hired out or otherwise disposed of without the publisher's consent, in any form of binding or cover other than that in which it is published.
- The correct price of this publication is the price printed on this page, Any revised price indicated by a rubber stamp or by a sticker or by any other means is incorrect and should be unacceptable.

OFFICES OF THE PUBLICATION DIVISION, NCERT

NCERT Campus
Sri Aurobindo Marg
New Delhi 110 016 **Phone : 011-26562708**

108, 100 Feet Road
Hosdakere Halli Extension
Banashankari III Stage
Bengaluru 560 085 **Phone : 080-26725740**

Navjivan Trust Building
P.O.Navjivan
Ahmedabad 380 014 **Phone : 079-27541446**

CWC Campus
Opp. Dhankal Bus Stop
Panihati
Kolkata 700 114 **Phone : 033-25530454**

CWC Complex
Maligaon
Guwahati 781 021 **Phone : 0361-2674869**

Publication Team

Head, Publication Division	:	*Anup Kumar Rajput*
Chief Editor	:	*Shveta Uppal*
Chief Production Officer	:	*Arun Chitkara*
Chief Business Manager (In charge)	:	*Vipin Dewan*
Production Assistant	:	*Rajesh Pippal*

Cover and Layout
Blue Fish

Illustrations
Joel Gill

About the Module

This is the fourteenth module of the course in Guidance and Counselling. It aims at developing basic statistical literacy especially in the context of guidance and counselling. The units covered in this module will help you to understand the statistical data given in the test manuals, and to analyse and interpret test scores.

The first unit deals with the most widely used methods of transforming raw scores into more interpretable, and therefore, more useful derived scores. A raw score, by itself is not interpretable. The solution to the problem of making scores meaningful involves providing an adequate frame of reference to them. The second unit briefly covers some initial concepts of sampling and some tests of significance, which help us to draw some conclusions about the population under consideration, after studying a small part of population, i.e. the sample. The t-test is a very useful method for testing hypotheses about group level differences in outcomes or measurements. This unit also covers some non-parametric tests, i.e. Chi-Square Test and Spearman's Rank-Order Correlation Coefficient Test, which are used when data have been categorized or ranked.

There are self-check exercises and activities in every unit which will help you evaluate your progress through the module. Summary given at the end provides an overview of the unit and suggested readings give additional sources of information.

Module Development Team

Contributors
Indu Gupta, *Reader*, Daulat Ram College, Delhi
Sridhar Srivastava, *Professor*, DES&DP, NCERT, New Delhi

Consulting Editor (Instruction design)
Bruce Thompson, British Columbia, Canada

Editor
A. K. Srivastava, *Professor,* DERPP, NCERT, New Delhi

Team Leader and Editor
Prabhat Kr. Mishra, *Senior Lecturer*, DEPFE, NCERT, New Delhi

Co-editor
Shraddha Dhiwal, *Lecturer*, DEPFE, NCERT, New Delhi

Project In-charge
Nirmala Gupta, *Professor*, DEPFE, NCERT, New Delhi

Members of the Review Team (October, 2016)
N.C.Ojha, *Assistant Professor*, RIE, Bhopal
Prem Sahajpal, *Reader*, LSR College for Women (Retd.), New Delhi
T.V.Somashekar, *Assistant Professor*, RIE, Mysore

Member Coordinators
Anjum Sibia, *Professor and Head*, DEPFE, NCERT, New Delhi
Prabhat K.Mishra, *Associate Professor*, DEPFE, NCERT, New Delhi

Acknowledgements

National Council of Educational Research and Training (NCERT) gratefully acknowledges the partnership and support of Commonwealth of Learning (COL), Vancouver, Canada for development of course material. This has been a gigantic task which has been possible with the help and cooperation of a large number of persons whose contribution we wish to acknowledge.

We gratefully acknowledge the continued support and encouragement provided by Professor Krishna Kumar, Director, NCERT all through the different stages. Special thanks are due to Professor Sushma Gulati, Head, DEPFE, NCERT for her constant guidance and leadership in steering the work through its various stages and to Professor D. K. Bhattacharjee the former Head of the Department for his help in initiating this work.

We thank Mukesh Kumar, Computer Assistant and Tanveer Ahmed, DTP Operator for typing, formatting and preparing graphics for this module.

Our grateful thanks are also due to Mrs. Usha Nair for language editing. The help provided by the Publication Department for preparing illustrations, layout and designing, and getting the material printed is also gratefully acknowledged.

Contents

1

Transformation of Scores

Transformation of Scores

1

1.0 INTRODUCTION

In the previous units, you looked at ways of describing a set of scores. For instance, you could summarise and organise the data in the form of a frequency distribution and present the data using graphs. Various statistical methods help assign meaning to the data or scores and make sense of it. For example, measures of central tendency are used to find out the central or typical value of a distribution. Variability measures tell us how clustered or scattered the scores are around a central value. Similarly, the shape of the distribution indicates the region in which most of the scores lie. These methods summarise and describe groups of scores. However, often in psychological and educational measurement, it is necessary to describe an individual score or specifically, where a score is located in the distribution. Suppose, you are told that your examination score in statistics is 123. Can you tell how well did you do in the test on the basis of your score only? Obviously, the raw score does not give you much information. To answer this question, you need to know something about the entire distribution of the scores or the performance of the group. It would help if you know the mean and the range of the test scores. This information will allow you to evaluate the location of your score in the distribution. If you are told that the class mean was 105, you are happy because your score is above the mean. On further enquiry you are told that the lowest score in the distribution was 80, and the highest score was 125. This information indicates that your score is on the higher side in the distribution.

The above example illustrates that a raw score, by itself, is not interpretable and to give meaning to it we need a frame of reference. Relating the score to a measure of central tendency and a measure of variability can provide this frame of reference. Such problems require transforming the given raw score into derived scores. In the process of transformation of raw scores into derived scores, the properties that exist in the original distribution may or may not be retained in the new system. Depending upon this characteristic, the derived scores can be classified into two types: linear transformations and non-linear transformations. In this unit, the relevance and techniques of derived scores will be discussed.

1.1 Objectives

After going through this unit, you would be able to

- *describe* the need and relevance of transforming raw scores into derived scores.
- *explain* the various techniques of linear transformations of raw scores e.g. Standard Scores – Z scores and others.
- *accomplish* non-linear transformations of raw scores such as Percentiles and Normalised Standard Scores – Stanines and T scores.

1.2 Need and Relevance of Derived Scores

In psychological and educational measurement, you may come across problems where you find that the raw score, by itself, does not communicate any meaning to you and is not interpretable. To give meaning to the raw score you need a frame of reference to determine whether the given score indicates a good performance or a poor one. The raw score originally without meaning becomes interpretable by relating it to a measure of central tendency and a measure of variability. Such problems of unequivocal meaning can be dealt with by using derived scores or score transformations. A derived score relates the position of a raw score either to (a) other scores in the same distribution, or (b) the distribution of raw scores obtained by a representative group. Scores on psychological tests are most commonly interpreted by reference to norms. Norms are summarised statistics that describe the test performance of reference groups in the standardisation samples. Grades, age, sex, standard scores, and percentile scores are common types of norms. You will come across these different types of norms in your guidance activities.

Derived scores serve dual purposes. First, they provide a standard frame of reference such that the meaning of a score can be better understood. Second, derived scores also provide comparable measures that permit a direct comparison of an individual's performance in different tests.

1.3 Types of Derived Scores

You will read in this section about various kinds of derived scores and also the methodology of converting raw scores into derived scores. Derived scores can be catogerised into two types – linear transformations and non-linear transformations.

1.3.1 Linear Transformations: Standard Scores

Linear transformations are defined by rules that include only a combination of multiplication, division, addition, and subtraction to set up one-to-one correspondence between raw and derived scores. Being linear transformations of scores these also preserve the proportional relation of inter-score distances in the distribution and, thus, do not change the shape of the distribution. All standard scores – Z scores and others – are linear transformations of the original raw scores. Latest tests are making increasing use of standard scores, which are the most satisfactory type of derived scores. A standard score provides information about location of a score in a distribution using mean and standard deviation as frames of reference. It conveys how far above or below a raw score is, from the mean of the distribution in relation to standard deviation. An important property of standard scores is that they have a fixed mean and a fixed standard deviation and that is what is 'standard' about them. The most important type of standard score is called a Z score. In addition, there are several other varieties of standard scores also. In the following account we will, however, only discuss the characteristics of Z scores in detail and the methodology of converting raw scores into Z scores.

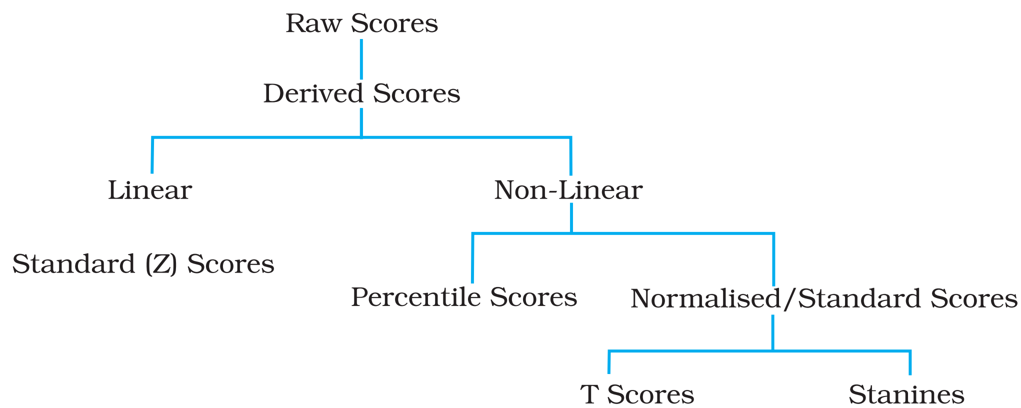


Fig. 1.1 : *Representation of Different Types of Derived Scores*

1.3.1.1 *Z scores*

One of the most familiar linear transformations of raw scores is the Z score. Z scores are standard scores that state the position of the raw scores in relation to the mean of the distribution, using standard deviation as the unit of measurement. Thus, Z scores provide information about location of a score in a distribution. The formula for Z scores can be expressed as follows:

Z score in a population: $$Z = \frac{X - \mu_x}{\sigma_x}$$

Z score in a sample: $$Z = \frac{X - \bar{X}}{S_x}$$

In this formula you will find that in order to calculate the Z score you require a set of raw scores (X), a mean ($\bar{X}$), and a standard deviation (S_x). If the value of a particular raw score is more than the mean, the resulting Z score would be positive (+); but if the raw score is less than the mean, the resulting Z score will be negative (-). Thus, a positive Z score shows that the raw score is more than the mean and negative Z score shows that the raw score is less than the mean. The location of any raw score can be understood by its Z scores. The higher the Z scores, the better the performance; the lower the Z score, the poorer the performance.

Consider the following example of converting raw scores into Z scores:

Example:

A distribution of achievement test scores has a mean of 60 and a standard deviation of 4. What will be the Z score for a student who received a score of 66?

Let us convert the raw score 66 into Z score using mean of 60 and standard deviation of 4 by using the formula given above for a sample:

$$Z = \frac{X - \bar{X}}{S_x} = \frac{66 - 60}{4} = \frac{6}{4} = +1.5$$

You can see that a raw score of 66 has a Z score of +1.5. It means that it lies 1.5 standard units (Z score) above the mean value of 60. This is graphically shown below in Figure 1.2.

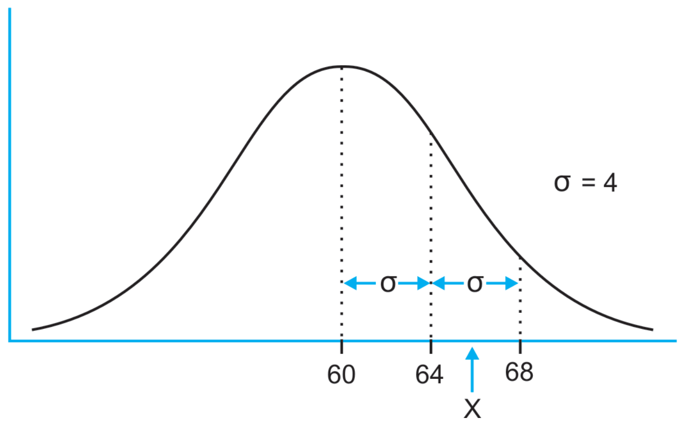


Fig. 1.2 : *Showing Location of Z score Equivalent of a Raw Score of 66 in a Distribution with a Mean of 60 and Standard Deviation of 4*

Self-check Exercise 1

1. What are standard scores?
2. A distribution has a mean of 100 and standard deviation of 16. Convert the following raw scores into Z scores:

   (a) X = 108, (b) X = 100, (c) X = 120

The Z scores as described above have three important properties:

(i) The mean of any distribution of scores converted to Z scores is always zero: mx = 0

(ii) The standard deviation of any distribution expressed in Z scores is always one: sx = 1

(iii) Transforming raw scores to Z scores changes the mean to zero and standard deviation to one, but it does not change the shape of the distribution. This means that being a linear transformation, the proportional relation that exists among the distances between scores remains the same.

This property can be understood by the following example of four raw scores and their Z score equivalents.

Example: Raw Scores and Z score Equivalents

Four scores from a distribution where $\mu = 80$ and $\sigma = 10$

Raw Scores	70		75		85		100
Inter-score differences among raw scores		5		10		15	
Equivalent Z scores	-1		$-\frac{1}{2}$		$+\frac{1}{2}$		$+2$
Inter-score differences among Z scores		$\frac{1}{2}$		1		$1\frac{1}{2}$	

In the above example you will note that the proportional magnitude of difference between successive raw scores, 5:10:15, is exactly the same as in Z score terms, $\frac{1}{2}:1:1\frac{1}{2}$. It is because of this reason that the shape of the distribution in the new system will be same as that of the raw score distribution.

1.3.1.2 *Importance of Z scores*

In educational and psychological measurements, you are often required to compare scores from different distributions. These comparisons may involve different raw scores of the same group, called 'intra-individual comparisons' or comparing raw scores from different groups, called 'inter-individual comparisons.' Since raw scores from different distributions are not directly comparable, they must be standardised by converting to Z scores and then interpreted accordingly. 'One element necessary to make an appropriate comparison is that the reference groups used to enable the standard scores be comparable and have roughly the same shape.' Unless the distributions from which standard scores are computed have the same shape, equal standard scores may not have equal rank in their respective distributions. Consider the following examples, one each of intra-individual and inter-individual comparisons, to understand how Z score transformations enable a comparison of scores from different distributions.

Example : Intra-individual Comparison

In an entrance test with the following means and standard deviations, Naresh has scored 56 in English, 72 in History, and 38 in Mathematics. Find out in which subject he has done better than the other two, so that proper educational guidance can be given.

	English	*History*	*Mathematics*
Mean	50	66	30
Standard Deviation	8	12	10

Since the three raw scores have different means and standard deviations in the respective subjects, direct comparison of performance cannot be made. The comparison can only be made by converting these raw scores into standard scores. The formula for converting raw scores into Z scores (standard scores) is: $Z = \dfrac{X - \bar{X}}{S_x}$

Now, Z scores of Naresh in three subjects can be obtained by substituting values in the above formula:

- English Test: $\dfrac{X - \bar{X}}{S_x} = \dfrac{56 - 50}{8} = \dfrac{6}{8} = 0.75\,Z$

- History Test: $\dfrac{X - \bar{X}}{S_x} = \dfrac{72 - 66}{12} = \dfrac{6}{12} = 0.50\,Z$

- Mathematics Test: $\dfrac{X - \bar{X}}{S_x} = \dfrac{38 - 30}{10} = \dfrac{8}{10} = 0.80\,Z$

The above calculations show that though Naresh had the highest score of 72 in History, but its Z score equivalent is only 0.50. On the other hand, the lowest score of 38 in Mathematics resulted in highest Z score of 0.80. Thus, we find that Naresh's best performance is in Mathematics (0.80Z) followed by English (0.75Z) and then History (0.50Z). He can now be given guidance accordingly.

Example: **Inter-individual Comparison**

To test achievement in Mathematics two different question papers are set and are given to students of section A and B. Ramesh, a student of section A got 80 marks, and Suresh, a student of section B, got 60 marks. Which of these two students stands better in terms of achievement in Mathematics? Means and standard deviations of the distributions of scores for sections A and B are as follows:

	Section A	*Section B*
Mean	70	50
Standard Deviation	20	10

Since marks of Ramesh and Suresh are on different tests, it cannot be concluded that Ramesh with 80 marks is a better student in Mathematics as compared to Suresh who has only 60 marks. In order to compare the two scores, their raw scores need to be transformed into standard scores (Z scores). The formula for transformation is:

$$Z = \frac{X - \bar{X}}{S_x}$$

Now, Z scores of Ramesh and Suresh can be obtained by substituting values in the above formula:

$$\text{Z score of Ramesh} = \frac{80 - 70}{20} = \frac{10}{20} = 0.5$$

$$\text{Z score of Suresh} = \frac{60 - 50}{10} = \frac{10}{10} = 1.0$$

Since Z score of Suresh (1.0) is better than Ramesh (0.5), it can be concluded that Suresh has achieved better in Mathematics than Ramesh.

Self-check Exercise 2

1. A student of class V has earned a score of 80 in History and 50 in Arithmetic. The counsellor wants to provide educational guidance on the basis of the performance in the two subjects. The mean of the History scores is 65 and standard deviation is 15. The mean of the Arithmetic scores is 65 and standard deviation is 7.5. Compare the performance of the student in the two subjects.
2. In Psychology, a student has received a score of 37, while the mean of the class was 28 and standard deviation was 6. In another section, her friend has received a score of 46, while the mean for that class was 35 and standard deviation was 10. Who has a higher standing in the class?

To conclude this section, it can be said that Z score transformations are extremely useful and have a number of advantages:

(i) Z scores indicates each person's standing as compared to the group mean. Those with positive Z scores receive raw scores that are above the mean, whereas those with negative Z scores receive raw scores that are below the mean.

(ii) Z scores can be safely used for making comparisons of scores from different distributions.

(iii) When the distribution of raw scores is reasonably normal, Z score can be directly converted into percentiles. Details about this will be discussed in the next section on percentiles.

1.3.2 Non-linear Transformations

One of the most important reasons for transforming raw scores into any type of derived scores is to render them comparable. In the previous section, you looked at standard scores (Z scores). In order to achieve comparability of scores from dissimilarly shaped distributions, non-linear transformations may be used. These transformations fall into two categories.

(a) *Percentile Ranks*

These are independent of the shape of the original distributions and, thus, may be used for comparison purposes even though the distributions differ in shape.

(b) *Normalised Standard Scores*

These concern converting scores to derived score distributions which have identical shapes, as well as identical means and standard deviations. In this case, scores are forced into a normal distribution, in addition to being given standard values for the mean and standard deviation and, therefore, the name normalised standard scores. Distributions treated in this fashion yield scores that are equal in rank as well as in score value and have the well-known properties of the normal distribution. This latter characteristic is particularly useful in advanced statistical procedures based on the assumption of normally distributed data. The two most important types of normalised standard scores are T scores and stanines. In the account given below, you will read about percentile scores and normalised standard scores such as T scores and stanines.

1.3.2.1 *Percentile Scores*

Percentile scores are also derived scores but their properties differ in some respects from those of standard scores. The percentile rank of a raw score describes its location relative to the other scores in the distribution. These are expressed in terms of the percentage of persons in the standardisation sample who fall below a given raw score. This may take values between 0 and 100. When the distribution of scores is reasonably normal, Z scores can be directly converted into percentiles. Figure 1.3 illustrates this relationship between Z scores and percentiles for a normally distributed variable. In figure 1.3, it can be seen that a Z score of 1.0 shows better score than 84% of the people and Z score of 0 indicates better performance than half of the people. This interpretation of Z scores is true only when scores are normally distributed.

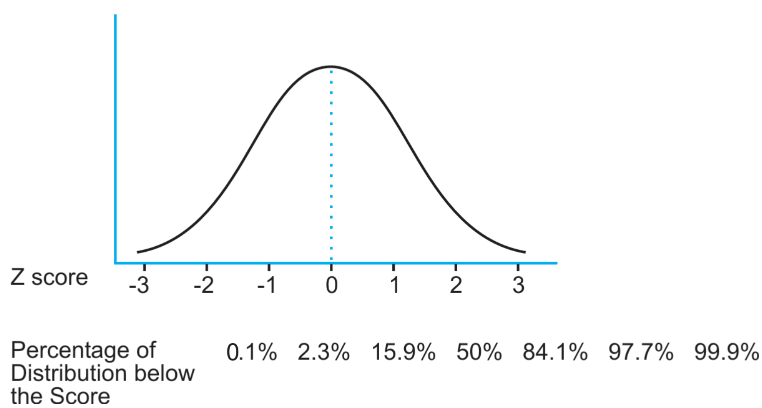


Fig. 1.3: *Relationship Between Z scores and Percentiles for Normally Distributed Variables*

Percentile scores have several advantages:

1. They have the property of directness of meaning.
2. These are easy to compute and are universally applicable.
3. They are independent of the shape of the original distributions and thus, may be used for comparison purposes even though the distributions differ in shape.

However, in spite of several advantages, percentile scores also have certain disadvantages. The primary disadvantage of the percentile scores is that changes in raw scores are ordinarily not reflected by proportional changes in percentile ranks. When one percentile rank is higher than another, the corresponding raw score of the one is also higher than that of the other, but we do not know by how much.

1.3.2.2 *Normalised Standard Scores*

As has been said above, one condition necessary for standard scores to be comparable is that their reference groups are comparable and they have roughly the same shape.

Unless this condition is met, the equal standard scores may not have equal rank in their respective distributions. Quite often one needs to compare distributions that are dissimilar in shape. In order to achieve comparability of scores from dissimilarly shaped distributions, normalised standard scores may be employed. The most commonly used normalised standard scores are: T scores and stanines. Normalised standard scores are like standard scores in the sense that their means and standard deviations are specified and have constant values. However, they differ in that, the process of transformation alters the shape of the original distribution of raw scores so that and have, the new distribution follows a normal curve and the statistical procedures based on the assumption of normally distributed data can be applied. Figure 4.4 shows how a skewed distribution of raw scores results in normal distribution after transformation.

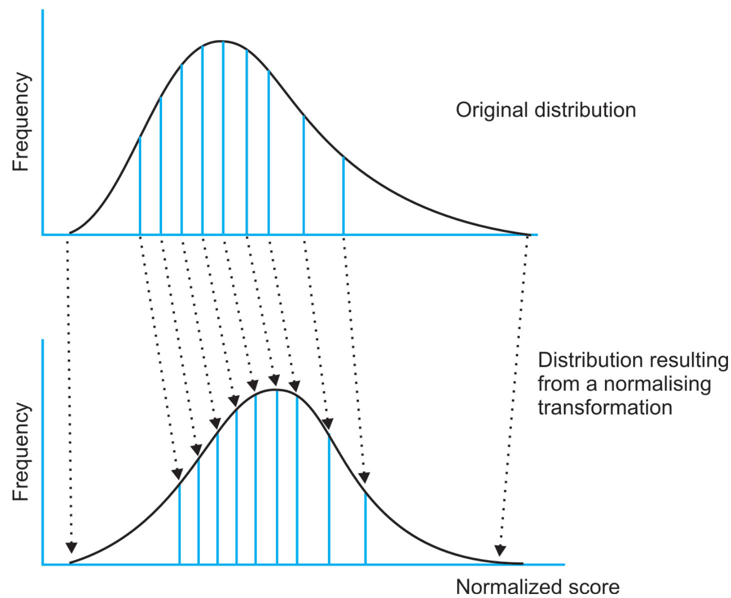


Fig. 1.4 : *Effect of Using a Normalised Standard Score Transformation*

There are several reasons for wanting to normalize a distribution:

- Most raw score distributions approximate the normal curve more closely than they do any other type of curve.
- Many physical and psychological measures generally yield normal distributions.
- Some advanced statistical procedures may be based on the assumption that the data are normally distributed, therefore, normalising transformations might be helpful here.

T score

One of the most important types of normalised standard scores is the T score. These were devised by McCall and were named so after an early leader in educational measurement, E.L. Thorndike. The T score is a normalised standard score based on the distribution with a mean of 50, a standard deviation of 10 and the shape of the normal curve, irrespective of the shape of the original scores. A distribution of Z scores can be altered to T scores by the equation

$T = 10Z + 50$

T scores have the advantage of general applicability. T scores are a convenient unit.

Stanines

Stanines are a type of normalised standard scores that are widely used in education for various purposes. The term stanine is a contraction for "standard nine" and the United States Air Force research workers devised these during World War II. These have whole number values ranging from 1 to 9, with a mean of 5 and a standard deviation of 1.96. The stanine is a relatively coarse unit of measurement, since the difference between successive stainines is one-half of a standard deviation. Its advantage is that of simplicity; only a single digit is required to express its value. In many situations, this amount of differentiation in performance is particularly poor at the extremes of the distribution.

Normalised standard scores as described above are the most satisfactory type of scores for a majority of purposes, but there are certain technical objections to normalising all distributions routinely. Such a transformation should be carried out only when:

(a) The sample size is large and representative.

(b) There is a good reason to force raw scores into the shape of a distribution different from that which they naturally exhibit.

(c) The original distribution of raw scores approximates normality so that the linearly derived standard scores and the normalised standard scores are very similar. Obviously, normalising a distribution that is already normal will produce little or no change.

The various types of derived scores described above, whether of the linear function type or the non-linear type, are not the only ones. These are the most common types used in psychological and educational measurements. In addition to these, a wide variety of other derived scores like age norms, sex norms, grade norms, etc. can be used depending upon their specific utility.

1.4 Summary

The goal of this unit was to discuss the most widely used methods of transforming raw scores into more interpretable, and therefore, more useful, derived scores. A raw score, by itself, is not interpretable. The solution to the problem of making scores meaningful involves providing an adequate frame of reference to them.

There are two major types of derived scores:

a. Those that preserve the proportional relation of inter-score distances. These are linear transformations where the shape of the original distribution is preserved. Z scores fall in this category. Z score states by how many standard deviations the corresponding raw score lies above or below the mean of the distribution.

b. The non-linear transformations of raw scores, unlike Z scores, do not preserve the inter-score distance and they don't preserve the shape of the original distribution. Percentile scores and normalised standard scores such as T scores and stanines are the commonly used derived scores that fall into this category.

Self-evaluation Exercises

1. What is the significance of score transformations in psychological measurement?
2. How are linear transformations different from non-linear transformations?
3. Why are Z scores are called 'standard scores?' List three important properties of Z scores.
4. Convert the following set of raw scores into Z scores in a distribution with a mean of 50 and a standard deviation of 8:
   (a) 58, (b) 46, (c) 66
5. For distribution A, mean is 20 and standard deviation is 7. Distribution B has a mean of 23 and standard deviation of 2. In which distribution will a raw score of 27 have a higher standing?
6. For a distribution with a mean of 50, a raw score of 43 corresponds to a Z score of –1.00. What is the standard deviation of this distribution?
7. List three uses of normalized standard scores. Write their special features.

Answers Key to Self-evaluation Exercises

1. In psychological measurement, it is necessary to describe an individual score or specifically where a score is located in the distribution. A raw score, by itself, is not interpretable and to give meaning to it we need a frame of reference. The raw score originally without meaning becomes interpretable by relating it to a measure of central tendency or a measure of variability. Such problems of unequivocal meaning can be dealt with by using derived scores or score transformations.
2. Linear transformations are defined by rules that include only a combination of multiplication, division, addition, and subtraction to set up one-to-one correspondence between raw and derived scores. Being linear transformations of scores these also preserve the proportional relation of inter-score distances in the distribution and, thus, do not change the shape of the distribution. On the other hand, in order to achieve comparability of score from dissimilarly shaped distributions, non-linear transformations may be used.
3. Z scores are standard scores that state the position of the raw scores in relation to the mean of the distribution, using standard deviation as the unit of measurement. They provide information about location of a score in a distribution.

Three important properties of Z scores are:

(i) The mean of any distribution of scores converted to Z scores is always zero: mx = 0

(ii) The standard deviation of any distribution expressed in Z scores is always one: sx = 1

(iii) Transforming raw scores to Z scores changes the mean to zero and standard deviation to one, but it does not change the shape of the distribution. This means that being a linear transformation; the proportional relation that exists among the distances between scores remains the same.

4. $$\text{Z score} = \frac{X - \bar{X}}{S_x}$$

(a) $$\text{Z score} = \frac{58 - 50}{8} = \frac{8}{8} = +1.00$$

(b) $$\text{Z score} = \frac{46 - 50}{8} = \frac{-4}{8} = -0.5$$

(c) $$\text{Z score} = \frac{66 - 50}{8} = \frac{16}{8} = +2.00$$

5. For distribution A, a raw score of 27 has a Z score of + 1.0 and for B a Z score of + 2.0. Therefore, a raw score of 27 has a higher standing in distribution B.
6. Standard deviation = 7
7. In order to achieve comparability of scores from dissimilarly shaped distributions, normalized standard scores may be employed. Normalised standard scores are like standard scores in the sense that their means and standard deviations are specified constant values. There are several reasons for wanting to normalise a distribution:
   - Most raw score distributions approximate the normal curve more closely than they do any other type of curve.
   - Many physical and psychological measures generally yield normal distributions.
   - Some advanced statistical procedures may be based on the assumption that the data are normally distributed, therefore, normalizing transformations might be helpful here.

Answer Key to Self-check Exercises

Self-check Exercise 1

1. Standard scores are the most satisfactory type of derived scores. They provide information about location of a score in a distribution using mean and standard deviation as frames of reference. They convey how far above or below a raw score is from the mean of the distribution in relation to standard deviation.

2. Z score $= \dfrac{X - \bar{X}}{S_x}$

(a) Z score $= \dfrac{108 - 100}{16} = \dfrac{8}{16} = +0.5$

(b) Z score $= \dfrac{100 - 100}{16} = \dfrac{0}{16} = 0.0$

(c) Z score $= \dfrac{120 - 100}{16} = \dfrac{20}{16} = +1.25$

Self-check Exercise 2

1. Z score of History $= \dfrac{X - \bar{X}}{S_x} = \dfrac{80 - 65}{15} = \dfrac{15}{15} = +1.00$

   Z score of Arithmetic $= \dfrac{X - \bar{X}}{S_x} = \dfrac{50 - 65}{7.5} = \dfrac{-15}{7.5} = -2.00$

   Z score of History is +1.00 and Z score of Arithmetic is -2.00. Therefore, performance in History is better than Arithmetic.

2. Z score of first student $= \dfrac{X - \bar{X}}{S_x} = \dfrac{37 - 28}{6} = \dfrac{9}{6} = +1.5$

   Z score of second student $= \dfrac{X - \bar{X}}{S_x} = \dfrac{46 - 35}{10} = \dfrac{11}{10} = +1.1$

   Z score of first student is +1.5 and Z score of second student is +1.1. Therefore, performance of first student is better.

Suggested Readings

Garrett, H.E. and Woodworth, R.S. 1981. *Statistics in Psychology and Education*. Vakils, Feffer and Simons Ltd., Bombay.

Minium, E.W. and Clarke, R.B. 1982. *Elements of Statistical Reasoning*. John Wiley and Sons, Inc., New York.

Minium, E.W., King, B.M. and Bear, G. 1993. *Statistical Reasoning in Psychology and Education*. John Wiley and Sons, Inc., New York.

Pagano, R.R. 1998. *Understanding Statistics in the Behavioural Sciences*. Brooks/Cole Publishing Company, Minnesota, USA.

2

Parametric and Non-parametric Statistics

Parametric and Non-parametric Statistics 2

2.0 Introduction

In a previous unit, you saw that frequency distribution could take a variety of shapes or forms. You were also introduced to some useful methods for describing and comparing raw data, which are commonly termed as 'Descriptive Statistics'. In educational and psychological measurement, we often need to move from description to decision-making. Decision-making means to draw a conclusion about large numbers of individuals. This stage of statistical treatment of data is known as statistical inference.

In dealing with educational guidance and counselling issues, like in a social investigation, we operate with limited time, energy and economic resources and hence, instead of studying every member of the large group (i.e. a population), we select a smaller number of members from that group and call it a sample. Most often, in educational studies, it is assumed that the large group of scores/measurements follows a normal distribution. This assumption, the process of generalising or making some statement, after studying the sample (a small group) about the entire population (large group) from which it was taken, is known as 'Parametric Statistics'. Thus, for parametric statistics, it is assumed that the characteristic studied is normally distributed in the specified population and is quantified as interval-level measurement.

Contrary to this, if the characteristic under study is not normally distributed in the specified population and is not quantifiable as interval-level measurement, there are some non-parametric tests, viz. Chi-Square (χ^2) Test and Spearman's Rank-Order Correlation Coefficient Test which are described in this unit.

2.1 Objectives

After going through this unit, you will be able to

- ***understand*** the concepts of sample, population, estimator, parameter, estimates and hypothesis; the fundamental concepts in drawing some conclusion about large groups.

- *describe* the procedure for drawing a conclusion about a population of units/subjects.
- *use* t-Test for small samples.
- *understand* Analysis of Variance (ANOVA), a method for studying the variation among several means.
- *use* the most commonly used Non-parametric Statistics, viz. Chi-Square Test and Spearman's Rank-Order Correlation Coefficient.

2.2 Sampling for Drawing Conclusion

In educational research, we operate with limited time, energy and economic resources, and we rarely are left with the option to study each and every member of the large group about which we want to draw some conclusion. In these situations, we study only a 'sample' – a smaller number of individuals or subjects taken from the larger group about which we want to make some comment or conclusion.

2.2.1 Some Key Terms Used in Sampling

Before using the method of sampling for drawing a conclusion about a group of individuals/objects/subjects, we must be aware of a few key terms as given below.

Unit of analysis: The type of entity/object/subject for which the characteristics under study is measured is called unit of analysis. In knowing the average achievement of students in an area in grade V, the students of grade V constitute the unit of analysis. In knowing the average achievement of schools in grade V, the schools would be the unit of analysis.

Population: The whole collection of units of a given type or condition, about which we wish to make conclusions, is called population. In the previous sections what we have been referring to as large groups are in fact examples of populations. For example, if we want to say or comment statistically about the achievement level of class VIII students of a particular district or region then our population would be: all the students studying in class VIII in all the schools in that particular district or region.

Let us consider another situation. If for the same purpose of commenting about the achievement level of class VIII students, we want to restrict our study or findings to students from Government schools, only and not from privately funded schools our population under study will comprise: all the students studying in class VIII, only in all the Government schools in that particular district or region.

Sampling frame: A listing of all the units belonging to the population under study, with their identification particulars, is called sampling frame. This listing or frame is used to draw the sample from that given population. In respect of the population just cited above as example, the identification particulars of all the students given in a listed manner will be the sampling frame. The testing may contain the columns Sl. No., Name, Age, School and School Address.

Sample: A sample is a subset of units from the sampling frame, which is selected for making the study and to give some statement or comment about the population. In the earlier example, instead of measuring the achievement of all the students studying in class VIII in all the schools of that particular district or region, we shall select some students in some fixed pre-decided number. These selected students will constitute our sample.

Sampling error: When we make some conclusion or statement in a measurable manner about the population on the basis of observations on a part (sample) of it, there may be some error (difference). The reason for the error is that, not all the units of the population have been covered under/the study/observation. This error is known as sampling error. There are statistical formulae to measure this error, which are beyond the scope of this unit. It is interesting to visualise and it is otherwise statistically observed that when the sample size, i.e., the number of units included in the sample, is increased, sampling error decreases.

Non-sampling error: Error in the conclusion/statement about a population caused by all the reasons other than the process of sampling is termed as non-sampling error. This error may be caused by error in measuring the variable under study, processing of data, etc. Sampling and non-sampling error are common to both complete enumeration and sample surveys. However, it is simple to understand that non-sampling error will occur more in the case of complete enumeration.

2.2.2 Simple Random Sampling (SRS)

A simple random sampling (SRS) scheme is the one in which every unit of population has an equal chance of being included in the sample. In practice, a simple random sample is drawn unit-by-unit and the desired measurements or study is done for these selected units. In SRS, a number is assigned to each unit of the population and a sample of pre-decided size is selected by drawing randomly from these numbers. Once the first unit of the sample is drawn, selection of the next unit can be done with or without replacement of the earlier selected unit back into the population.

When the unit selected at each draw is replaced back into the population before the next draw is made, the procedure of sampling is known as **Simple Random Sampling With Replacement (SRSWR)**. On the other hand, if the unit selected at a draw is not replaced back into the population before the next draw is made, the procedure is called **Simple Random Sampling Without Replacement (SRSWOR)**. In the first technique, there is always a possibility that a unit selected in a sample at a draw may reappear many times in the sample and then for the same sample size it may restrict the representation of the population. In practice, therefore, in educational research, SRSWOR is preferred over SRSWR.

There are different methods to select a sample under SRS scheme, viz. Lottery method or different the use of Random Number Tables. The Lottery method is the simplest method of selecting a simple random sample. In this method, if there are 'N' units in the population and we want to draw a sample of size 'n' then we will use a set of N tickets, with number 1 to N written on them. The tickets will be thoroughly mixed up and then 'n' tickets will be drawn one by one. The units, which have serial numbers occurring on those 'n' tickets, will be considered as selected and will be subjected to study or measurement.

Use of Random Number Tables is also a practical method of drawing a simple random sample. This method is a little tricky to understand but it is inevitable when the population size is huge, running in to many thousands. Use of the Random Number Table for creating sample sizes is more costly in terms of time and is generally used for state or national government wide studies where numbers are much greater. Nowadays, computer programmes can also be written to generate random samples for a particular sampling frame.

2.2.3 A Few Notations Used in Sampling

As you have understood, in any statistical study you typically try to obtain a sample which is representative of the larger population in which you have an interest. But, you should be careful to distinguish between characteristics of the sample you actually study and the population about which you make some conclusion or statement. In order to make this distinction in your statistical procedures, you can no longer use the same symbols to signify the mean and standard deviation of both sample and population. Instead, use different symbols depending on whether you are referring to sample or population characteristics. With reference to the **mean**, you can symbolise the mean of a *sample* as $\bar{x}$ and the mean of the *population* as $\bar{X}$, a notation you are well acquainted with all through previous modules under descriptive statistics. With reference to the **standard deviation**, which is a measure of variability in the population, you can denote/ write as,

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}\left(Xi - \bar{X}\right)^2 \text{ which is population variance.}$$

Square root of population variance is standard deviation of the population.

$$S^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(X_i - \bar{X}\right)^2 - \text{Population Mean Square}$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 - \text{Sample Mean Square}$$

Self-check Exercise 1

You have to comment on achievement of Class VIII girls, in subject Mathematics, studying in schools managed by Government (not by private bodies) in your district.

i. What will be your target population?

ii. What will be your sampling frame?

2.3 Hypothesis Testing and Estimation

In earlier Section, we saw that a population mean or proportion can be estimated from the information we get in a single sample. Although any such estimate has obvious importance, it does not constitute the complete decision making strength. It does not constitute the primary decision making goal or activity of social research. Researchers in education are preoccupied with the task of testing hypotheses about differences between two or more samples. In a slightly more technical language, we can understand that statistical inference (decision making) falls into two broad categories. One is 'Statistical Estimation' and another is 'Hypothesis Testing'.

By 'Statistical Estimation,' we mean the process of utilising information available from a random sample to estimate values of population parameters of interest. There are a variety of methods that can be used to estimate population means, proportions, variances, standard deviations, differences of two population means, differences of two population proportion, etc. In this Unit, these are not included as they are beyond the scope of the Unit.

2.3.1 Hypothesis Testing

Before understanding the idea of hypothesis testing, you should understand what exactly the word 'hypothesis' means in statistics. Hypothesis, in its word meaning, is simply an assertion, claim, or statement about something which is based on a few known facts but that has not yet been proved to be true or correct.

In statistics, a hypothesis is always a statement about the value of the parameter. It is a testable proposition offered by the researcher about the problem under investigation. The testable proposition should be liable to be converted into a quantitative statement. For example, a researcher may have a proposition that the level of achievement of students who have willingly joined a particular vocation is higher compared to those who have joined unwillingly. This proposition or research question involves making a comparison between two groups. Thus, a statistical hypothesis is a quantitative statement.

Below are a few examples of hypotheses:

i. $\mu_{rural} = 45$; where μ_{rural} denotes the average marks of students in rural area in a district in subject 'Mathematics' in Class X Board Examination.

ii. $\mu_{urban} = 55$; where μ_{urban} denotes the average marks of students in urban area in a district in subject 'Mathematics' in Class X Board Examination.

iii. μ_{rural} “ μ_{urban}; μ_{rural} and μ_{urban} denoting marks as explained above.

'Hypothesis Testing' is the operation of deciding whether or not data obtained for a random sample supports or fails to support a particular hypothesis. In practice, the result of testing a hypothesis is a declaration that the hypothesis is supported by the data or that it is not supported by the data collected under the sample.

Null Hypothesis: It is a measurable statement about the population parameter, which is tested for possible rejection under the assumption that it is true. By possible rejection we mean that, the null hypothesis is constructed in such a manner that it states that "there is no phenomenon" and the results in the sample could have arisen through chance. For example, if we want to compare the test scores of two random samples of boys and girls, a null hypothesis would be that the mean score of the boys is same as the mean score of the girls, and therefore there is no significant statistical difference between them.

$H_0: \mu_1 = \mu_2$

where, H_0 denotes the null hypothesis

μ_1 (mu 1) = the mean of scores of boys, and

μ_2 (mu 2) = the mean of scores of girl.

Alternative Hypothesis: This is complementary to null hypothesis. When a null hypothesis is formed, it is always in contrast to an implicit *alternative hypothesis*, which is accepted if the observed data values are sufficiently improbable under the null hypothesis. This is also considered as the research hypothesis. The precise formulation of the null hypothesis has a bearing on the alternative. If one wants to test the null hypothesis that the mean score of the boys is same as the mean score of the girls then alternative hypothesis could be–

A: Mean score of the boys is not the same as the mean score of the girls. This will lead to two tailed test i.e. $\mu_1 \neq \mu_2$.

B: Mean score of the boys is greater than the mean score of the girls. This will lead to right tailed test i.e. $\mu_1 > \mu_2$.

C: Mean score of the boys is less than the mean score of the girls. This will lead to left tailed test i.e. $\mu_1 < \mu_2$.

Level of Significance: In statistical studies, the level of significance is a number that expresses the probability that the result of a given experiment or study could have occurred purely by chance. This number can be a margin of error or it can indicate a confidence level as expressed by a statement such as the following: "If this experiment were repeated, there is a probability of ninety-five percent that our conclusions would be substantiated."

In other words, it is the percentage chance that null hypothesis is rejected even though it is true. If the null hypothesis is true, the significance level is the probability that it will be rejected in error. This chance of committing error arises due to fluctuations in sampling. Popular levels of significance are 5% and 1%.

5% Level of Significance: It means there is 5% chance that we reject the null-hypothesis even though it is true. Consider the above example. Say, we rejected null hypothesis that the mean score of the boys is the same as the mean score of the girls at a 5% level of significance on the basis of sample scores. It implies that there is a 5% chance that mean score of the boys is same as the mean score of the girls in the population and we concluded wrongly that they are not the same.

1% Level of Significance: There is 1% chance that we reject the null hypothesis, even though it is true. Consider the above example. Say we rejected the null hypothesis that the mean score of the boys is equal to the mean score of the girls at 1% level of significance on the basis of sample scores. It implies that there is a 1% chance that the mean score of the boys is equal to the mean score of the girls in the population and we concluded wrongly that they are not the same.

2.4 The t-Test

While testing differences between samples, educational researchers ask such questions as: Do Girls' performance differ from Boys, in the subject Mathematics? Do Rural students score less than Urban students in school leaving Board Examinations? Note that these research questions involve making a comparison between two groups: Boys versus Girls and Rural versus Urban.

In this section, we will learn how to use a statistical tool to evaluate hypotheses about group-level differences in outcomes or measurements. This tool is known as t-test. Specifically, we will learn to use two different applications of the t-test in evaluating two kinds of hypotheses.

- The one sample t-test, in which the level of outcome for a group is compared to a known standard or claimed value.
- The two sample t-test, where the outcome levels of two groups are compared to each other (This also has two situations which will be explained later).

One of the advantages of the t-test is that it can be applied to a relatively small number of cases. It is, as a convention, used to evaluate statistical differences for samples of 30 or less.

1. One Sample Test

This is also known as testing hypothesis about a single mean. One sample t-test compares the mean score of a sample to a known value, usually the population mean. The basic idea of one sample t-test is to compare the average of the sample (observed average) and the population (expected average), with an adjustment for the number of cases in the sample and the standard deviation of the average. The assumptions of one sample t-test are:

i. Population from which sample is drawn is normal.
ii. The sample is drawn at random.

Let us consider that 10 students are chosen at random from a class of Standard XII. Their heights are found to be, in inches, 62, 64, 66, 67, 68, 69, 70, 70, 71 and 71. With the observations of this selected sample, if we want to test that the mean height in the population of the students is more than 66 inches–

In the first place, let us note that from our knowledge of distribution of height, the population here is likely to be approximately normal and the units of the sample has been choosen randomly.

The null hypothesis will be

$$\left.\begin{array}{lll} H_O: \mu = \mu_0 & \text{or} & \mu = 66 \text{ inches} \\ H_A: \mu > \mu_0 & \text{or} & \mu > 66 \text{ inches} \end{array}\right\} \text{for this example}$$

t-test in the situation of one sample will be

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

with (n-1) degrees of freedom.

where, $s^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} (x_i - \bar{x})^2$

The decision rule is that if for (n-1) degrees of freedom at level of significance,

if $t_{calculated} \geq t_{tabulated}$, reject H_0

if $t_{calculated} \leq t_{tabulated}$, accept H_0.

For the given data on height of students in a sample of 10 students from a class of Standard XII–

$$\bar{x} = 67.8 \text{ inches and S. D.} = 3.048$$

and hence,

$$t = \frac{67.8 - 66}{3.048} \sqrt{10} = 1.87$$

We have,

$t_{calculated}$ = 1.87, and $t_{tabulated}$ at 9 degree of freedom and 5 per cent level of significance as equal to 2.262. Since, $t_{calculated} < t_{tabulated}$, we will accept and conclude that the mean height in the population of students is not more than 66 inches.

2. Two Sample Test

This is also known as testing the difference between means. Differences between groups often provide the rationale on which a social study or investigation is conducted. Here also the researchers confront two situations. One, where they have to test the hypothesis about the difference between two independent means, and two, where they have to test the hypothesis about the difference between two dependent means. We will discuss both the situations separately below.

a. Difference between two independent means: We often want to know whether the means of two independent samples differ statistically. For example, there are many questions in which we want to compare two categories of some categorical variable (e.g., compare boys and girls or rural and urban students as regard to achievement) or two set of subjects receiving different treatments in context of an experiment or intervention. The assumptions of this application of t-test are:

(i) Population from which samples are drawn is normal.
(ii) The samples are drawn independently and at random.
(iii) Population variances are equal but unknown.

Let us consider an example where we have to deal with scores obtained from two groups of students in achievement test conducted in the same school.

Student Number	1	2	3	4	5	6	7	8	9	10
Group I	44	45	50	55	60	66	57	49	56	68
Group II	35	38	40	44	48	50	54	48	56	64

$$\bar{x}_1 = \frac{550}{10} = 55.0, \quad \bar{x}_2 = \frac{477}{10} = 47.7$$

Here, the null hypothesis is–

$H_O: \mu_1 = \mu_2$, and

the alternative hypothesis $H_A: \mu_1 > \mu_2$

t-test in this situation of two samples is

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where, s_p is pooled S.D. of the two samples, i.e.

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

and,

$$s_1^2 = \frac{1}{(n_1 - 1)}\sum_{i=1}^{n_1}(x_i - \bar{x}_1)^2 \text{ ... First sample}$$

$$s_2^2 = \frac{1}{(n_2 - 1)}\sum_{i=1}^{n_2}(x_i - \bar{x}_2)^2 \text{ ... Second sample}$$

The decision rule is that if for $(n_1 + n_2 - 2)$ degree of freedom at α level of significance,

If $t_{calculated} > t_{tabulated}$, reject

If $t_{calculated} < t_{tabulated}$, accept H_0

For the given data on achievement score of two groups we have–

$\bar{x}_1 = 55.0, \quad \bar{x}_2 = 47.7$

$s_1^2 = 68.89, \quad s_2^2 = 78.68$

$$s_p^2 = 73.785$$

$$t = \frac{55.0 - 47.7}{8.59\sqrt{\frac{1}{10} + \frac{1}{10}}} = 1.90$$

We have,

$t_{calculated} = 1.90$, and

$t_{tabulated}$, at (10 + 10 – 2) i.e. 18 degree of freedom and 5 per cent level of significance as equal to 1.73.

Since, $t_{calculated} > t_{tabulated}$, we will reject H_O. It is concluded that Group I has statistically different achievement score and it is higher.

Remarks: In the present example, the size of two groups was same i.e. 10. In case of groups having different size, the test will also be applied in similar manner.

b. Difference between two dependent means (Paired t-test): We will now consider the application of t-test in the situation when the sample sizes are equal and two samples are not drawn independently. The sample observations are paired together. It means that the pair of observations correspond to the sample unit/subject. For example, if we want to test if there is any statistically significant difference in the performance of a group of children before and after the remedial classes, this is a case of a single sample measured at two different points in time (time 1 versus time 2). The assumptions of this application of t-test are:

i. Population from which sample is drawn is normal.
ii. The paired sample is drawn at random.
iii. Population variances are equal but unknown.

Let us now see an example where we want to test the effectiveness of three months of remedial classes on the numerical skills of students of Class V. We have a difference (increase or decrease) in marks obtained by a sample of 15 students in numerical skills as under:

4, 2, 8, 5, 3, –1, 0, 1, –2, 4, 6, 5, 0, 5, 3

Here, the null hypothesis is–

$H_O : \mu_1 = \mu_2$, that is, there is no significant difference in the numerical skills of the student after three months of remedial course, and

the alternative hypothesis is–

$H_A : \mu_1 < \mu_2$, that is, there is statistically significant effect of remedial course and the overall performance of students in numerical skills has improved.

t-test for application in this situation when we have pair of observations corresponding to the units of sample is given as,

$$t = \frac{\bar{d}}{\sqrt{\frac{s_d^2}{n}}} \text{ with (n – 1) degree of freedom}$$

$$\text{where, } \bar{d} = \frac{1}{n}\sum_{i=1}^{n} di \text{ and } di = x_{i1} - x_{i2}$$

where, x_{i1} is the i^{th} observation at time 1 (before remedial course) and x_{i2} is the i^{th} corresponding observation at time 2 (after remedial course).

The decision rule is that for (n – 1) degree of freedom at level of significance–

If $t_{calculated} \geq t_{tabulated}$, reject H_0

If $t_{calculated} \leq t_{tabulated}$, accept H_0

For the given pair of observations, we have,

$\bar{d} = 2.867$ and $s^2_d = 7.98$

hence

$$t = \frac{2.867}{\sqrt{\frac{7.98}{14}}} = \frac{2.867}{0.729}$$

$$= 3.933$$

We have $t_{calculated} = 3.933$, and $t_{tabulated}$, at (15 – 1), i.e. 14 degree of freedom for right tailed test at 5% level of significance as 2.145.

Since $t_{calcualted} > t_{tabulated}$, we will reject . We will conclude that the remedial course of three months duration has statistically significant effect on the performance of students under observation.

2.5 Analysis of Variance (ANOVA)

RURAL versus URBAN, BOYS versus GIRLS, PRIVATE versus GOVERNMENT represent the kind of two sample comparisons that we discussed earlier. But while dealing with educational research, the social reality cannot always be conveniently segregated into two groups; the subjects under study do not always divide themselves in so simple a fashion. As a result, we often seek to make comparisons among three, four, five, or more samples or groups. For example, we may study the influence of a management style of schools (Private, Government, Local Body, and Private Unaided) on achievement, or the degree of economic status of parents (Low income group, Middle income group, High income group) on performance in terminal exam results, and so on.

In such situations, instead of making use of a series of t-tests to make comparisons among all possible pairs of three or four samples (or groups), we conduct an Analysis of Variance (ANOVA). This technique not only saves us from making several comparisons but also makes the interpretations of result as regards to differences in means simpler. It provides us with a technique whether a significant difference is present among the three or four or more sample means we seek to compare.

An Example

Let us consider a research situation where we seek to determine the influence of school management on "achievement in English Language" in a district of India. We now want to make comparisons between several groups based on the management of schools. We are proceeding here with our research hypothesis that schools with different types of managements viz. Government, Local Body, Private Aided and Private Unaided, differ with respect to achievement in English Language in a particular class, say Class V.

Marks obtained by 20 students in a test, cross tabulated with the management of the school are:

Management of School	**Students**					**Mean**
	S-1	S-2	S-3	S-4	S-5	
Government	26	46	33	45	38	37.6
Local Body	32	33	28	35	39	33.4
Private Aided	45	52	46	49	52	48.8
Private Unaided	48	49	56	62	58	54.6

What we do in ANOVA

To conduct an analysis of variance, we treat the total variation in a set of scores as being divisible into two components.

i. The distance of raw scores from their group means known as **variation within groups**, and
ii. The distance of group means from one another referred to as **variation between groups**.

Once the total variation in the set of scores is divided into the two types of variation given above, the analysis of variance yields an F-ratio (or F-test) in which **variation between groups** and **variation within groups** are compared. The F-ratio can be regarded as indicating the size of between-groups variation in relation to the size of within-groups variation. Having obtained an F-ratio, we now determine whether it is large enough to reject the null hypothesis and accept the research hypothesis.

The larger our calculated F-ratio (which means larger between-groups variation and smaller within-groups variation), the more likely we will obtain a statistically significant result with regard to variation in the study variable on account of different groups.

In the present example we will proceed with Null hypothesis as:

***Null Hypothesis*:** Government, Local Body, Private Aided and Private Unaided Schools do not differ with respect to achievement.

This hypothesis will be tested for possible rejection against the Alternative or Research hypothesis as

***Alternative Hypothesis*:** Government, Local Body, Private-Aided, and Private-Unaided schools differ with respect to achievement.

The Analysis of variance is represented as:

ANOVA Table

Source of Variation	*Degree of Freedom*	*Sum of Squares (SS)*	*Mean of Sum of Squares (MSS)*	*F (Calculated)*
Between Groups	P-1	BGSS	BGMS = BGSS/(P-1)	F = BGMS/EMS
Error (Within Groups)	N-P	ESS = TSS - BGSS	EMS = ESS/(N-P)	
Total			N-1	TSS

where, TSS = (SS of each observation) – (CF) (*TSS is the sum of the squared deviations of every raw score from the total mean of the study. It is equal to a combination of its within-and between-group components*);

CF = Square of Grand Total/ Total No. of observations;

BGSS = Sum of {(Square of Group Total)/No. of observations in the group} - CF (*BGSS is the sum of the squared deviations of every sample mean from the total mean of the study.*);

P = Number of Groups; and

N = Total Number of Observations.

The Within Group Sum of Squares, also known as 'Error Term' is nothing but the sum of the squared deviations of every raw score from its group mean.

Remarks

It important to recall that as a predominant practice in statistical discussion, 'N' denotes the population size and 'n' denotes the sample size. But here we are denoting by 'N' the total number of observations in the sample. Notations make no difference as long as it is clear in the mind of user that what the notation is used for. The total number of observations in a study where the measurement is made on a large number of subjects may run into hundreds, thousands or many thousands.

Conclusion

If the value F (calculated) > F (tabulated) for (P-1), (N-P) d. f. at required level of significance, the null hypothesis is rejected. Otherwise, the null hypothesis is retained.

When the values of test scores enumerated for the present example as given below, we get the results of ANOVA:

Management of School	Students					Mean
	S-1	S-2	S-3	S-4	S-5	
Government	26	46	33	45	38	37.6
Local Body	32	33	28	35	39	33.4
Private Aided	45	52	46	49	52	48.8
Private Unaided	48	49	56	62	58	54.6

ANOVA Results

RAW SCORE_ENGLISH

Source of Variation	*Sum of Squares*	*d. f.*	*Mean Square*	*F*
Between Groups	1440.400	3	480.133	14.429
Within Groups	532.400	16	33.275	
Total	1972.800	19		

Conclusion: F (Calculated) value is 14.429. F (Tabulated) value for 3, 16 d .f. at 5% level of significance is 3.24. Since F (Calculated) is greater than F (Tabulated), the Null hypothesis is rejected. We accept the Alternative (Research) hypothesis that the achievement scores have significant difference over different management of a school. Or, we can understand that the management of schools is impacting on achievement.

2.5.1 F-Test to Test Equality of Variance

Here, the F-test used in ANOVA is nothing but the F-test statistic used for testing equality of variance.

F = Variance (X) / Variance (Y)

When employed in ANOVA it yields,

F = (Found variation in the group means)/(Expected variation in the group means)

Self-check Exercise 2

To investigate the influence of a branch of engineering on the starting salary of engineering graduates, researchers interviewed recent college graduates on their first jobs who studied one of the following branches in engineering, Computer Science, Mechanical and Electronics. The result obtained for 45 respondents are as follows:

Salary (in Rs.)		
Computer science	*Mechanical*	*Electronics*
10,500	7,000	7,500
12,300	9,500	9,000
14,000	10,000	8,000
9,500	11,000	9,300

9,000	8,500	10,500
8,500	7,500	10,000
7,500	7,000	7,000
15,750	10,500	11,250
18,450	14,250	13,500
21,000	15,000	12,000
14,250	16,500	13,950
13,500	12,750	15,750
12,750	11,250	15,000
11,250	10,500	10,500
9,500	9,500	9,500

Is there any significant difference between the starting salaries for different branches?

2.6 Non-parametric Statistics

You know that tests of significance, which require i) normality, and ii) an interval level measure, are referred to as parametric tests. There are situations in educational research when you cannot employ a parametric test for the reason that either you cannot honestly assume normality or when data are not amenable to an internal level measure. Suppose, for example, that you have to deal with a skewed distribution, such as annual income, or with data that have been categorised and counted (the nominal level) or ranked (the ordinal level). For such situations, statisticians have developed a number of non-parametric tests of significance. These are the tests whose list of requirements or assumptions does not include a normal distribution or the interval level measurement.

Two most widely used test statistics are Chi-Square (χ^2) Test and Spearman's Rank-Order Correlation Coefficient, which are described in the subsequent sections.

2.6.1 Chi-square Test

This is the most popular non-parametric test of significance in social research. This test is used to examine the difference between theoretical ratio (assumed ratio) and observed ratio (observed in a survey or experiment). It enables us to find if the deviation of the observed ratio from the assumed/theoretical ratio is just by chance or is statistically significant.

In fact, one of the most common and useful way to look at information in educational research is in the format of a table. Say, for example, we want to know whether boys or girls get into trouble more often in the subject Mathematics. There are many ways we might show information related to this question, but perhaps the most frequently and easily used method is to show this information in tabular form as shown.

Gender	Trouble in Mathematics		Total
	Yes	No	
Boys	92	142	234
Girls	74	166	240
Total	**166**	**308**	**474**

It can be quickly noticed that more boys than girls got into trouble in Mathematics. Calculating the percentage, we find that 39% of boys got trouble (92 boys out of 234 total boys = 39%), as compared with 31% of girls (74 girls out of 240 total girls = 31%). However, what if we wanted to test the hypothesis that boys get into trouble more often than girls in Mathematics, or in other words, there is some association of gender with the situation of getting trouble in Mathematics. These figures are a good start to examine that hypothesis. However, the figures in the table are only descriptive.

To examine the hypothesis, we need to employ a statistical test, which is called the chi-square test. Application of chi-square test in this situation is called the test of independence of attributes. It is the case of two variables.

In case of single variable, chi-square is applied to test the goodness of fit or the test of significance. By this, we mean that we will be testing whether the observed distribution of the variable fits the expected one or it does not.

1. ***One Variable Case***

Let us pose a **research question** like this: Do students have a preference for vocation?

Here, the **hypotheses** are

Null (H_O)	The observed distribution fits the expected or, in other words, there is no preference.
Alternative (H_A)	The observed distribution does not fit the expected or, in other words, there is a preference.

Here, the null hypothesis states that population of students does not differ with respect to the frequency of occurrence of a given characteristic — i.e., the liking for a vocation. It means that in a sample of, say, 100 students if preference is asked by each student about a vocation, then the frequency for each vocation should be 25. It means to understand that if we are asking students which of the four vocations they prefer and there is no preference, then we would expect 25% to prefer each vocation. Let us consider the following data:

	Vocation I	Vocation II	Vocation III	Vocation IV	
Expected	25	25	25	25	
Observed	35	30	20	15	n = 100

For computation of chi-square test, we define

E_j = the **Expected** frequency in the j-th column

O_j = the **Observed** frequency in the j-th column

and

$$\text{Chi-square } (\chi^2) = \sum_{j=1}^{c} \frac{(O_j - E_j)^2}{E_j} \qquad (1)$$

Here, c is the number of columns (in the present example, the number of vocations).

The, decision rule to apply Chi-square test will be that if c = the number of columns, then with degrees of freedom (df) = c-1, with α level of significance,

If χ^2calculated > χ^2 tab, reject H_0

If χ^2 calculated < χ^2 tab, retain H_0

For the given example, as per the formula given in (1) the calculated value of χ^2 will be

$$\chi^2 = \frac{(35-25)^2}{25} + \frac{(30-25)^2}{25} + \frac{(20-25)^2}{25} + \frac{(15-25)^2}{25}$$

$$= \frac{100}{25} + \frac{25}{25} + \frac{25}{25} + \frac{100}{25}$$

$$= 4 + 1 + 1 + 4$$

$$= 10$$

If we want to put forth our decision at 5 per cent level of significance, that is, for a value of = 0.05, then the decision will be like this,

We have

χ^2 Calculated = 10, and

χ^2 Tabulated = 7.82 at 5% level of significance and df = 3

Since

χ^2 calculated > χ^2 tabulated, we will reject the null hypothesis, and will accept the alternative hypothesis that there is preference about vocation and that is why the observed distribution does not fit the expected distribution.

2. *Two Variables Case*

Let us relook at the example given in Section 2.6.1 We have already discussed that this is the situation of two variable and in such situations where information is presented in cross-table, we apply Chi-square test to test the independence of attributes. Let us reproduce the information, given in Section 2.6.1, in the cross-table regarding gender-wise trouble in the subject 'Mathematics' again.

Gender	Trouble in Mathematics		Total
	Yes	No	
Boys	92	142	234
Girls	74	166	240
Total	**166**	**308**	**474**

Here the research question will be that whether getting in trouble in the subject 'Mathematics' is associated with the gender of the students.

The hypotheses are:

Null (H_O)	There is no association (or relation or contingency) between the two variables (attributes), that is, they are independent.
Alternative (H_A)	The two variables are associated or related or not independent.

For computation of Chi-square test, we will define

E_{jk} =the expected frequency of the cell defined by the j-th row and the k-th column

O_{jk} =the observed frequency of the cell defined by the j-th row and the k-th column

Here j and k are just to denote the number of rows and columns. In the present example, there are two rows and two columns. So, j will take values 1 and 2, and similarly k will also take the values 1 and 2. In total, we have here 4 cells having some frequency in them.

It is important to mention that there are certain assumptions with which we apply Chi-square test. They are:

i. The sample units are selected randomly.
ii. The scores are independent (i.e., each subject fits in only one cell of the table).
iii. For a 2×2 table, all expected cell frequencies should be at least equal to 10 (for larger cross-tables, this value is 5).

We can make sure that in the present example all these assumptions are fulfilled.

Now for computation, we have the observed frequencies given in the cells. For computing the expected cell frequencies, we use the formula

$$E_{jk} = \frac{(\text{Row Marginal} \times \text{Column Marginal})}{N}$$

A helpful check is that the sum of the expected cell frequencies is equal to N, that is,

$$\sum_{j=1}^{r}\sum_{k=1}^{c} E_{jk} = N$$

and,

$$\text{Chi-square } (\chi^2) = \sum_{j=1}^{r}\sum_{k=1}^{c} \frac{\left(O_{jk} - E_{jk}\right)^2}{E_{jk}} \qquad (2)$$

The decision rule to apply Chi-square test will be that if r is the number of rows (here 2–Boys, Girls) and c is the number of columns (here 2–Yes, No), then with degree of freedom (r–1) (c–1), with α level of significance

If calculated $> \chi^2$ tab, reject

If χ^2 calculated $< \chi^2$ tab, retain H_0

For the given example, as per the formula given in (2), the expected frequencies of the cell will be as under (given in brackets),

Gender	Trouble in Mathematics		Total
	Yes	No	
Boys	92 (81.9)	142 (152.1)	234
Girls	74 (84.1)	166 (155.9)	240
Total	**166**	**308**	**474**

According to the formula for χ^2, we must subtract each expected cell frequency from its corresponding observed cell frequency, square the difference, divide by the appropriate expected cell frequency, and add up these quotients to obtain the Chi-square value. Hence,

$$\chi^2 = \frac{(92-81.9)^2}{81.9} + \frac{(142-152.1)^2}{152.1} + \frac{(74-84.1)^2}{84.1} + \frac{(166-155.9)^2}{155.9}$$

$$= \frac{(10.1)^2}{81.9} + \frac{(-10.1)^2}{152.1} + \frac{(-10.1)^2}{84.1} + \frac{(10.1)^2}{155.9}$$

$$= \frac{102.01}{81.9} + \frac{102.1}{152.1} + \frac{102.1}{84.1} + \frac{102.01}{155.9}$$

$$= 1.25 + 0.67 + 1.21 + 0.65$$

$$= 3.78$$

If we want to put forth our decision at 5 per cent level of significance, that is, for a value of $\alpha = 0.05$, then the decision will be like this,

We have

χ^2 calculated = 3.78, and

χ^2 tabulated = 3.84 at $\alpha = 0.05$ and df = 1

Since

χ^2calculated $< \chi^2$tabulated, we will retain H_O, the null hypothesis, that there is no association between the attributes, that is, they are independent. In other words, we can conclude that boys are not significantly more likely to get in trouble in the subject 'Mathematics' than girls.

Self-check Exercise 3

Let us imagine that we are investigating the relationship between political orientation and child-rearing methods. We have a sample of 89 mothers having children between the ages of 5 to 10 years. Among these 89 mothers, 27 are liberals, 30 are moderates, and 32 are conservative. The mothers have also been categorised according to child-rearing methods, as permissive, moderate and authoritarian.

How can we test that child-rearing practice is independent of political orientation? The cross table cited gives the observed frequencies in different categories.

Child-Rearing Method	Political Orientation		
	Conservative (f observed)	Moderate (F observed)	Liberal (f observed)
Permissive	7	9	14
Moderate	10	10	8
Authoritarian	15	11	5

Using Chi-Square, test the relationship between political orientation and child-rearing method. State your conclusion at the 5 per cent level of significance.

2.6.2 Spearman's Rank-Order Correlation Coefficient Tests

To know the degree of relationship between two variables we use a technique in statistics called Correlation or Correlation Coefficient. It indicates the strength and direction of the linear relationship between two variables. For dealing with ordinal data, we have a technique known as Spearman's Rank-Order Correlation Coefficient Tests.

This test deals with the ordinal data. For example, if a group of n individual students is arranged in order of merit or proficiency in possession of two characteristics A and B, then we may like to know the degree of association between the two characteristics. Consider ranks of the same ten students in Mathematics and Physics, or, rank of their socio-economic status and time spent on watching television. In this situation we apply spearman's rank-order correlation coefficient (ρ). The formula is,

$$\rho = 1 - \frac{6\sum d^2}{N(N^2 - 1)}$$

where,

ρ = The rank-order correlation coefficient

D = rank difference between X and Y variable

N = the total number of cases

Dealing with Tied Ranks

In actual practices it is not always possible to rank or order our unit of observations avoiding ties at each and every position. To illustrate the procedure for obtaining a rank-order Correlation Coefficient in the case of tied ranks, let us say we are interested in determining the degree of association between position in Class XII board examination and IQ. Suppose we are able to rank a sample of ten such students with respect to their class position in board examination and their IQ measured by some test, and get the values as below:

Student	Class Position	I.Q. (X)	I.Q. Rank (Y)	Y corrercted
S –1	1	140	1 Tie	1.5
S –2	2	140	2 Tie	1.5
S –3	3	115	4	4
S –4	4	132	3	3
S –5	5	110	5 Tie	6
S –6	6	110	6 Tie	6
S –7	7	100	9	9
S –8	8	104	8	8
S -9	9	90	10	10
S –10	10	110	7 Tie	6

In situations as given above, in order to determine the exact position in the case of ties, we add the tied ranks and divide by the number of ties. Therefore the position of a 140 IQ, which has been ranked as 1 and 2, would be $(1 + 2)/2 = 1.5$, and in the same way the position of an IQ score of 110 is $(5 + 6 + 7)/3 = 6.0$.

2.7 Summary

The parametric statistics assumes that the characteristic under study is normally distributed in the specified population and is quantifiable as an interval-level measurement. When there is a large population of subjects and we have to make comments that are statistically genuine about the population, then we draw a random sample constituting some randomly drawn, independent subjects from that population. By making observations on these randomly drawn units of the sample, we try to make comments about the population. This unit briefly covers some initial concepts of sampling and some tests of significance, which helped us to draw some conclusions about the population under consideration, after studying a small part of population, i.e., the sample. The t-test is a very useful method for testing hypotheses about group level differences in outcomes or measurements.

We often seek to make comparisons among three, four, five, or more samples or groups with regard to some characteristics under study. In such situations, we make use of an analysis of variance (ANOVA), which saves us from making several comparisons and makes the interpretation of results with regard to differences in means simpler. This unit also covers some non-parametric tests, which are used when data have been categorised or ranked. The Chi-Square (χ^2) Test and Spearman's Rank-Order Correlation Coefficient Test have been given in the unit with examples.

Self-evaluation Exercises

1. Let us assume

   X: 5, 4, 3, 2, 1, 2, 6, 4, 4, 1, 3, 4, 5, 2, 2, 6, 4, 3, 4, 5

   and

   x: 4, 2, 5, 1, 3

   Represent set of scores representing population (N = 20) and sample (n = 5).

   Calculate:

   (i) Population Mean Square

   (ii) Sample Mean Square

2. 20 randomly selected students from a section of Class XII were told that the principal of the school wished to know their preference regarding a forthcoming visiting lecture series. In particular, they were asked to evaluate a professor who might be visiting the class. The professor's curriculum vitae was described to all 20 students in the same way with one exception: one-half of the students were told that the professor was 65 years old; one-half were told the professor was 25 years old. Their willingness to attend the lecture was measured in scores. The following results were obtained:

   X: 65, 38, 52, 71, 69, 72, 55, 78, 56, 80

   Y: 78, 42, 77, 50, 65, 70, 55, 51, 33, 59

   X is the score of students who were told professor is 25 years old. Y is the score of students who were told professor is 65 years old.

   Which statistical procedure would you apply to determine whether there is a significant difference between these groups of students with respect to their willingness to attend the lecture (Apply your test procedure at 5% level of significance).

3. In an experiment similar to that given in the self-evaluation exercise 2, 30 students were told about the professor in the same way with one exception: one-third of the students were told that the professor was 75 years old; one-third were told that the professor was 50

years old; and one-third were told that the professor was 25 years old. Their willingness to attend the lecture was measured in scores. The following results were obtained:

X: 65, 38, 52, 71, 69, 72, 55, 78, 56, 80

Y: 63, 42, 60, 55, 43, 36, 69, 57, 67, 79

Z: 67, 42, 77, 32, 52, 34, 45, 38, 39, 46

X is the score of students who were told professor is 25 years old.
Y is the score of students who were told professor is 50 years old.
Z is the score of students who were told professor is 75 years old.

Which statistical procedure would you apply to determine whether there is a significant difference between these groups of students with respect to their willingness to attend the lecture (Apply your test procedure at 5 per cent level of significance).

4. A random sample of students of a university was selected and asked their opinion about evaluation of teachers by students. The students were from 1st, 2nd and 3rd year of Graduate course and also from Postgraduate courses. The results are given below. Test the hypothesis at 5% level of significance that opinions are independent of class groupings.

Class	Number		Total
	Favouring Evaluation by Students	Opposed to Evaluation by Students	
1st Year	120	80	**200**
2nd Year	130	70	**200**
3rd Year	70	30	**100**
Postgraduates	80	20	**100**
Total	400	200	600

5. The ranks of the same 16 students in Mathematics and Physics are as follows. Two numbers within brackets denote the ranks of the students in Mathematics and Physics.
(1, 1); (2, 10); (3, 3); (4, 4); (5, 5); (6, 7); (7, 2); (8, 6); (9, 8); (10, 11); (11, 15); (12, 9); (13, 14); (14, 12); (15, 16); (16, 13)
Calculate the Spearman's rank correlation coefficient for proficiencies of this group in Mathematics and Physics.

Answer Key to Self-evaluation Exercises

1. i. Population Mean Square = 2.263
ii. Sample Mean Square = 2.00
2. $t_{calculated} = 0.896$
Since $t_{calculated}$ is greater than $t_{tabulated}$ at 5% level of significance for 18 degrees of

freedom, the null hypothesis that there is no significance difference will be rejected.

3. $F_{calculated} = 3.62$

Since $F_{calculated}$ is less than the $F_{tabulated}$ value i.e. 4.21 at 5% level of significance, it implies that there is no significant difference in willingness of students on account of knowledge about age of professor. Here null hypothesis is accepted.

4. χ^2 calculated = 12.743

χ^2 calculated value is greater than χ^2 tabulated value for 3 degree of freedom at 5% level of significance. Hence the null hypothesis that there is no influence of class-grouping is rejected and we conclude that the opinions about evaluation of teachers by students are dependent on class-grouping.

5. $\rho = 0.8$

Answer Key to Self-check Exercises

Self-check Exercise 1

i. Our target population will be all the girl students studying Mathematics in Class VIII of such schools in the district, which are being managed by Government.

ii. Sampling frame will be listing of all such girls falling in the target population. This will include identification particulars of all these girls.

Self-check Exercise 2

There is significant difference in the starting salary hence we will conclude that there is influence of branch of engineering on the starting salary of engineering graduates.

Self-check Exercise 3

χ^2 tabulated is 9.49

χ^2 calculated is 7.58

df = 4

p = 0.05

Since χ^2 calculated is less than χ^2 tabulated, we accept the null hypothesis that child-rearing method is independent of political orientation. There is no association between these two attributes.

Suggested Readings

Garrett, H. E. and Woodworth, R. S. 1969. *Statistics in Psychology and Education*. Vakils, Fetter and Simons Pvt. Ltd., Bombay.

Minium, E. W. and Clarke, R. B. 1982. *Elements of Statistical Reasoning*. John Wiley and Sons, New York.

Miniun, E. W., King, B. M. and Bear, G. 1993. *Statistical Reasoning in Psychology and education*. John Wiley and Sons, Inc., New York.

Pagano, R. R. 1998. *Understanding Statistics in the Behavioural Sciences*. Brooks/Cole Publishing Company, California.

NOTES

List of Course Material

1. **Course Guide**

   Major inputs include objectives, scope, rules, syllabi as well as procedures for admission, transaction and evaluation for all the three phases of the course.

2. **Course Modules***

   i. Module- I : Introduction to Guidance
   ii. Module-II : Counselling Process and Strategies
   iii. Module-III : Guidance for Human Development and Adjustment
   iv. Module-IV : Career Development-I
   v. Module V : Career Information in Guidance and Counselling-I
   vi. Module VI : Assessment and Appraisal in Guidance and Counselling-I
   vii. Module VII : Basic Statistics in Guidance and Counselling-I
   viii. Module VIII : Guidance in Action
   ix. Module IX : Special Concern in Counselling
   x. Module X : Developing Mental Health and Coping Skills
   xi. Module-XI : Career Development-II
   xii. Module XII : Career Information in Guidance and Counselling-II
   xiii. Module XIII : Assessment and Appraisal in Guidance and Counselling-II
   xiv. Module XIV : Basic Statistics in Guidance and Counselling-II

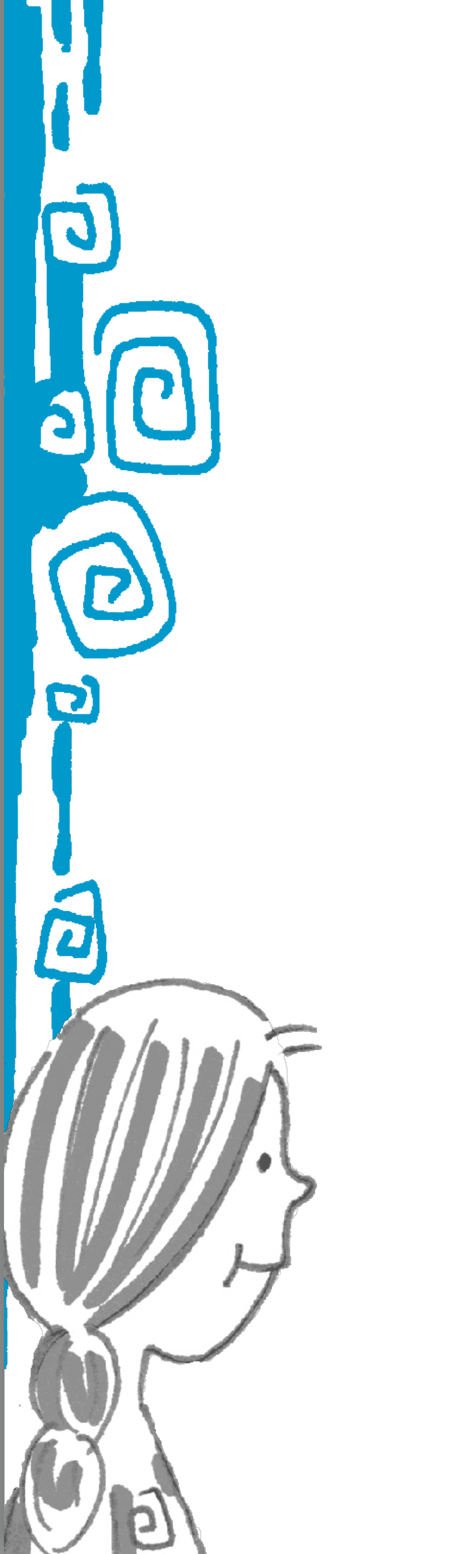
* *Each module consists of number of self-learning units.*

3. **Practical Handbook**

   Provides areas and strategies for conducting and undergoing practicum, field experience and internship.

4. **Tutor Guide**

   Lists guidelines for tutors, supervisors for course transaction and evaluation during all the three phases of the course.



2363

राष्ट्रीय शैक्षिक अनुसंधान और प्रशिक्षण परिषद्
NATIONAL COUNCIL OF EDUCATIONAL RESEARCH AND TRAINING

ISBN 978-81-7450-956-7