Statistical Techniques

By: Ranveer

This reference book can be useful for

BBA, MBA, B.Com, BMS, M.Com, BCA, MCA

and many more courses for Various Universities

NEERAJ
PUBLICATIONS
www.neerajbooks.com

Published by:

NEERAJ
PUBLICATIONS
(Publishers of Educational Books)

Sales Office : 1507, 1st Floor,
Nai Sarak, Delhi-110 006
E-mail: info@neerajbooks.com
Website: www.neerajbooks.com

Typesetting by: Competent Computers

Terms & Conditions for Buying E-Book

- The User must Read & Accept the Terms and Conditions (T&C) carefully before clicking on the accept option for Buying the Online Soft Copy of E-books. Under this Particular Facility you may buy only the Online Soft Copy of E-books, no Hard Copy or Printed Copy shall be provided under this facility.

- These E-Books are valid for 365 days online reading only (From the Date of Purchase) and no kind of Downloading, Printing, Copying, etc. are allowed in this facility as these products are just for Online Reading in your Mobile / Tablet / Computers.

- All the online soft copy E-books given in this website shall contain a diffused watermark on nearly every page to protect the material from being pirated / copy / misused, etc.

- This is a Chargeable Facility / Provision to Buy the Online Soft Copy of E-books available online through our Website Which a Subscriber / Buyer may Read Online on his or her Mobile / Tablet / Computer. The E-books content and their answer given in these Soft Copy provides you just the approximate pattern of the actual Answer. However, the actual Content / Study Material / Assignments / Question Papers might somewhat vary in its contents, distribution of marks and their level of difficulty.

- These E-Books are prepared by the author for the help, guidance and reference of the student to get an idea of how he/she can study easily in a short time duration. Content matter & Sample answers given in this E-Book may be Seen as the Guide/Reference Material only. Neither the publisher nor the author or seller will be responsible for any damage or loss due to any mistake, error or discrepancy as we do not claim the Accuracy of these solution / Answers. Any Omission or Error is highly regretted though every care has been taken while preparing these E-Books. Any mistake, error or discrepancy noted may be brought to the publishers notice which shall be taken care of in the next edition. Please consult your Teacher/Tutor or refer to the prescribed & recommended study material of the university / board / institute / Govt. of India Publication or notification if you have any doubts or confusions before you appear in the exam or Prepare your Assignments before submitting to the University/Board/Institute.

- Publisher / Study Badshah / shall remain the custodian of the Contents right / Copy Right of the Content of these reference E-books given / being offered at the website www.studybadshah.com.

- The User agrees Not to reproduce, duplicate, copy, sell, resell or exploit for any commercial purposes, any portion of these Services / Facilities, use of the Service / Facility, or access to the Service / Facility.

- The Price of these E-books may be Revised / Changed without any Prior Notice.

- The time duration of providing this online reading facility of 365 days may be alter or change by studybadshah.com without any Prior Notice.

- The Right to accept the order or reject the order of any E-books made by any customer is reserved with www.studybadshah.com only.

- All material prewritten or custom written is intended for the sole purpose of research and exemplary purposes only. We encourage you to use our material as a research and study aid only. Plagiarism is a crime, and we condone such behaviour. Please use our material responsibly.

- In any Dispute What so ever Maximum Anyone can Claim is the Cost of a particular E-book which he had paid to Study Badshah company / website.

- If In case any Reader/Student has paid for any E-Book and is unable to Access the same at our Website for Online Reading Due to any Technical Error/ Web Admin Issue / Server Blockage at our Website www.studybadshah.com then He will be send a New Link for that Particular E-Book to Access the same and if Still the Issue is Not Resolved Because of Technical Error/ Web Admin Issue / Server Blockage at our website then His Amount for that Particular Purchase will be refunded by our website via PayTM.

- All the Terms, Matters & Disputes are Subjected to "Delhi" Jurisdiction Only.

CONTENTS

STATISTICAL TECHNIQUES

Descriptive Statistics

1

INTRODUCTION

In this chapter, we will discuss the basics and terms related to statistics. Although, most of the learners have been acquainted with statistics in earlier classes, but really it is necessary here. In this chapter, we shall discuss the methods of how to collect the data as well as to organize these data (i.e. concept of frequency distribution). We will also study the various models of frequency distribution not only in tabular form, but also in diagrammatic representation. Thus, our objective of this chapter study is:

- To define qualitative and quantitative character and differences between the two.
- To define a discrete and a continuous variable and differences between the two.
- To draw the frequency table with their relative frequencies, cumulative frequencies and frequency densities.
- To explain the diagrammatical presentation of various frequency distributions.

CHAPTER AT A GLANCE

COLLECTING DATA AND KINDS OF DATA

Let's start with some examples which provide you a general idea about situations where we need to handle a large amount of data and where statistics can play a significant role. In these examples, we try to raise issues, which can be handled adequately by the various statistical tools:

Consider a system where 'customers' arrive at a 'counter' for 'service'. 'Customers' may be patients coming to a clinic for medical attention or may be aircrafts waiting for clearance from the air traffic control

to take-off or even broken down machines in a factory waiting for the attention of an operator and so on. Our objective is to prescribe a policy so that congestions can be avoided. However, neither the number of arrivals is fixed on all occasions nor is the service time the same for all customers - these are usually uncertain and thus subject to chance factors. How do we then propose to proceed?

Suppose a new brand of pain reliever has been marketed recently. The manufacturer claims that it relieves pain 25% time faster than any of the comparable brands already available in the market. How do we propose to verify this claim?

Obviously, we have to administer these drugs to a sample of individuals. But then how should the sample be chosen? Also, individuals may react differently to the same drug.

How do we take this into account? How do we process the sample data? Finally, the good old question – to what extent can we generalise our sample findings so as to be able to come up with a conclusion pertaining to the entirety? How reliable is the sample finding in this case?

Many such examples may be cited to illustrate the areas of application of statistics.

However, it will be gradually apparent that the basic issues involved in these illustrations are similar and can be discussed within a broad framework and this framework is provided by statistics.

Observation before it is arranged and analysed is called raw data. For data to be useful, our observations need to be organised so that we can pick out trends and come to logical conclusions.

Raw Materials of Statistics

In our daily life we have seen the various terms related to statistics, so we have to learn firstly, the meaning of statistics. Statistics is defined in two defferent contexts: **numerical data** and **discipline.** We will understand it by some examples like the statistics of run sconed by the **Chennai Superkings** in IPL-2 matches, statistics of marks obtained by the students in Economics in an examination, etc., these are termed as numerical data.

On the contrary, a player of the Chennai Superkings or a student of Economics are called the discipline. Thus, we can say that the numerical data arises in the ambid of life whereas the disciplines itself relates to the collection, analysis and interpretation of data. Hence, the combination of numerical data and disciplines is known as **Statistics.**

Two Basic Concepts Regarding the Statistics

Let us consider the two basic concepts regarding to the statistical study: One is **character** and another is **individual.** Again, we will understand these two concepts through an example:

● Suppose a teacher has awarded to his students on their performance in an examination by grades (A, B, C, D and E). Here, students are individuals and the grades are the characters.

Thus, through this example, we can say that one attribute who gets the benefits is called **individual** and types of attributes/benefits are called **characters.**

● **Sources of Data:** In any study/research, we collect the data from two sources: **Primary** and **Secondary.**

(i) Primary Data Sources: In any study/research, when we collect the data on the relevant groups of individuals by survey method, it is called the **primary source.**

(ii) Secondary Data Sources: In any study/ research when we take data from which is already published by the Government or any other agency, is called **secondary source.**

Note: In using the secondary data sources, we shall be careful that data should be reliable and relevant to their study.

● **Methods of Collecting Primary Data**

There are various methods to collect the primary data, one of the most popular method is direct observation, which is as under:

Direct Observation: Through counting or measurement or by inspection, when we collect the required information, then this type of observation is called **direct observation**.

> **One who provides the information is called informant.**

In direct observation we collect the information directly by informants or through enumerators by following methods:

(1) Questionnaire Method: In this method, the enumerators collect the data by filling out the questionnaire forms. This method is very useful for the educated informants.

(2) Interview Method: This method is generally useful for illiterate or uneducated informants. In this method, enumerators collect data/schedule by a thorough and logical questioning of each informant.

● **Classification of Characters**

The character is the types of attributes/benefits, which are classified into two broad categories: one is qualitative and another is quantitative character.

(1) Qualitative Character: Such a character that can't be counted or expressed numerically, but it has various forms for different individuals are called qualitative characters. As for example, the brand name of the motorbikes in Delhi is a character: it may be Herohonda, Yamaha, Bajaj, etc. whose possible forms can be differentiated orally but not numerically, is called qualitative character.

(2) Quantitative Character: Such a character that can be counted or expressed numerically for different individuals is called quantitative character. As for example, when we would ask a question how many motorbikes of Herohonda in Delhi, it can be counted and such type of character is called quantitative character.

● **Distinction between Qualitative Character and Quantitative Character**

Qualitative Character	Quantitative Character
● It can't be expressed numerically, but observed orally.	● It is expressed only numerically.
● It is generally known as attributes.	● It is known as variable.
● It is observable through the ranking of the preferences.	● It has no need for making the rank.

Now, let us consider the quantitative character which is classified into two parts: one is **discrete** and another is **continuous variable.**

(i) **Discrete Variable:** It is the variable which is not observable in internal, but which can be conceivable only some actual or isolated variables. As for example, the size of the family takes values like 1, 2, 3...... etc. height of the children like. 2 ft, 4 ft, 3.5 ft.....etc. are called discrete variables.

(ii) **Continuous Variable:** It can take any values in some interval, say the ages of the teachers between 25 years to 55 years, number of students of primary school between 5 years to 12 years, etc. are called the continuous variable. Similarly, suppose the lower level of an interval is 'β' and upper level of that interval is 'α', then the continuous variable is defined as [α, β] of the given data.

FREQUENCY DISTRIBUTION OF A VARIABLE
Frequency Distributions

Earlier, we have studied more about the collection of data. Now, we will be acquainted with the organization of data through frequency distribution. For the comfortable study of the frequency distribution, we, therefore, categorize the frequency distribution into two parts: ungrouped frequency distribution and grouped frequency distribution.

Now, let us start with the ungrouped frequency distribution.

Ungrouped Frequency Distribution: Ungrouped frequency distribution might have the data with qualitative in nature or the variable with discrete. So, we shall first discuss the ungrouped frequency distribution with qualitative character and then the ungrouped frequency distribution with discrete variable.

(1) Ungrouped Frequency Distribution of Qualitative Character: This concept can be easily understood by an illustration. Let us consider a college conducts a graduation examination which consists of four subjects like, Statistics, Economics, Mathematics and English. There are 100 students who have passed in all four subjects, which is shown in a frequency distribution table as:

Table 1: Frequency Distribution of the Passed Students

Subjects	No. of Students	Relative Frequency
Statistics	30	30/100 = 0.3
Economics	20	20/100 = 0.2
Mathematics	40	40/100 = 0.4
English	10	10/100 = 0.1
Total	100	1.0

Table 1 shows the frequency distribution of 100 students who have passed in four subjects like Statistics, Economics, Mathematics and English.

● The data of the second column are called frequencies of the four subjects.

● Column 1 and 2 show the frequency distribution among 100 students in four subjects.

● Column 3 shows the relative frequency of that four subjects by this formula:

$$\text{Relative frequency of a subject} = \frac{\text{Frequency of that subject}}{\text{Total frequency}}$$

As for example:

Relative frequency of statistics

$$= \frac{30}{100} = 0.3$$

Similary we can find all relative frequencies, which are shown in Table 1.

Note: ● A frequency must be non-negative.
● A relative frequency must be a rational number in the interval [0, 1].

If qualitative character is classified just in two classes known as Dichotomy.

(2) Ungrouped Frequency Distribution of a Discrete Variable: This concept is also to be understand through an illustration. Let us consider an economist collects the data on household size from so households of rural locality, which is shown in Table 2.

Table 2: Data of Household Size of 80 Rural Households

8	4	4	3	7	8	3
3	2	4	9	6	1	2
5	3	5	4	5	7	1
5	2	4	4	5	4	4
3	4	5	5	6	5	5
4	4	2	4	5	2	5
4	3	5	5	6	6	6
5	3	7	2	7	6	2
8	1	6	5	6	6	9
7	9	5	4	5	5	3

Now we create the frequency table of these discrete values of the households, which is shown in Table 3.

Table 3: Frequency Distribution for the Households Size of 80 Rural Households

Household Size	Tally Marks	Frequency	Relative Frequency
1	|||	3	3/80 = 0.0375
2	卌 |||	8	8/10 = 0.1000
3	卌 卌	10	10/80 = 0.1250
4	卌 卌卌	15	15/80 = 0.1875
5	卌 卌 卌卌	20	20/80 = 0.2500
6	卌卌		11	11/80 = 0.1375
7	卌		6	6/80 = 0.0750
8	||||	4	4/80 = 0.0500
9	|||	3	3/80 = 0.0375
Total		80	1.0000

● **Cumulative Frequency Distribution for the Discrete Variable:** There are two other ways to represent the frequency distribution of the discrete variable. These are: **Less than type** and **More than type** cumulative frequency distribution.

Table 4: Cumulative frequency distribution of the 'less than type' and 'more than type' of household size of 80 rural households

Less than type of cumulative frequency distribution			More than type of cumulative frequency distribution		
Household size	Frequency	Cumulative frequency	Household size	Frequency	Cumulaties frequency
Less than 1	3	3	More than 1	3	80
Less than 2	8	11	More than 2	8	77
Less than 3	10	21	More than 3	10	69
Less than 4	15	36	More than 4	15	59
Less than 5	20	56	More than 5	20	44
Less than 6	11	67	More than 6	11	24
Less than 7	6	73	More than 7	6	13
Less than 8	4	77	More than 8	4	7
Less than 9	3	80	More than 9	3	3
80			Any value more than 9		0

Procedure to Create Less Than Type of Cumulative Frequency: If we have to find the less than type of cumulative frequency of a particular row, we add the previous frequency with its frequency/ cumulative frequency. Like as, we have to find the cumulative frequency we add as, 3 + 8 = 11, similarly for the 3rd row, we add 11 + 10 = 21 and so on.

Note: Cumulative frequency of the first row is the same as the frequency of that row and cumulative frequency of the last row equals to the total frequency.

Procedure to Create More Than Type of Cumulative Frequency: If we have to find the more than type of cumulative frequency of a particular row, we substract the total frequency/remaining cumulative