NEERAJ®

STATISTICAL TECHNIQUES

By:

Ranveer, M.C.A.

Reference Book Cum Question Bank

New Edition

NEERAJ PUBLICATIONS

(Publishers of Educational Books)

(An ISO 9001 : 2008 Certified Company)

1507, Ist Floor, NAI SARAK, DELHI - 110 006

Ph. 011-23260329, 69414105 Fax 011-23285501

E-mail: info@neerajbooks.com

Website: www.neerajbooks.com

Price

₹ 160/-

Published by:

NEERAJ PUBLICATIONS

Admn. Office : **Delhi-110 007**
Sales Office : **1507, 1st Floor, Nai Sarak, Delhi-110 006**
E-mail: **info@neerajbooks.com** *Website*: **www.neerajbooks.com**

	Typesetting by: ***Competent Computers***	*Printed at*: ***Novelty Printers***

Notes:

1. *For the best & upto-date study & results, please prefer the recommended textbooks/study material only.*
2. *This book is just a Guide Book/Reference Book published by NEERAJ PUBLICATIONS based on the suggested syllabus by a particular Board /University.*
3. *The information and data etc. given in this Book are from the best of the data arranged by the Author, but for the complete and upto-date information and data etc. see the Govt. of India Publications/textbooks recommended by the Board/University.*
4. *Publisher is not responsible for any omission or error though every care has been taken while preparing, printing, composing and proof reading of the Book. As all the Composing, Printing, Publishing and Proof Reading etc. are done by Human only and chances of Human Error could not be denied. If any reader is not satisfied, then he is requested not to buy this book.*
5. *In case of any dispute whatsoever the maximum anybody can claim against NEERAJ PUBLICATIONS is just for the price of the Book.*
6. *If anyone finds any mistake or error in this Book, he is requested to inform the Publisher, so that the same could be rectified and he would be provided the rectified Book free of cost.*
7. *The number of questions in NEERAJ study materials are indicative of general scope and design of the question paper.*
8. *Subject to Delhi Jurisdiction only.*

How to get Books by V.P.P.

If you want to buy **NEERAJ® BOOKS** through V.P.P. then you can order the same at the Online Order Form of **www.neerajbooks.com** or send us an E-mail at **info@neerajbooks.com** or you may send us a Fax at **011-23285501** or write us a letter alongwith Rs. 100/- as advance by Bank Draft mentioning your complete requirement, class, course, subject code, medium, name, address and telephone no. which will help us in sending your required **NEERAJ® BOOKS** to you at the earliest.

23260329
69414105
Fax: 011-23285501

NEERAJ PUBLICATIONS

(Publishers of Educational Books)
(An ISO 9001 : 2008 Certified Company)
1507, 1st Floor, NAI SARAK, DELHI - 110 006
E-mail: **info@neerajbooks.com** *Website:* **www.neerajbooks.com**

CONTENTS

STATISTICAL TECHNIQUES

■■

Sample Preview of the Solved Sample Question Papers

Published by:

NEERAJ PUBLICATIONS

www.neerajbooks.com

QUESTION PAPER

(June – 2016)

(Solved)

STATISTICAL TECHNIQUES

Time: 2 hours] *[Maximum Marks: 50*

Note: *(i) Attempt both Sections, i.e. Section A and Section B.*
*(ii) Attempt any **four** questions from Section A.*
*(iii) Attempt any **three** questions from Section B.*
(iv) Non-scientific calculator is allowed.

SECTION-A

Q. 1. The mean and standard deviation of 20 items is found to be 10 and 2, respectively. At the time of checking it was found that one noted item with value 8 was incorrect. Calculate the mean and standard deviation, if the wrong item is deleted.

Sol. Let the variable x denote items. Then we are given

$$\bar{X} = \frac{\Sigma X}{20} = 10$$

$$\Rightarrow \quad \Sigma x = 200$$

$$\text{Corrected } \Sigma x = 200 - (\text{Wrong item value})$$

$$= 200 - 8$$

$$= 192$$

$$\therefore \text{ Corrected mean item} = \frac{192}{19} = 10.10$$

$$\text{Standard deviation on } \sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2 - (\bar{x})^2}$$

$$\Rightarrow \quad 2 = \sqrt{\frac{1}{20}\text{Incorrect}\sum_{i=1}^{n} x_i^2 - (10)^2}$$

$$\Rightarrow \quad 2 = \frac{1}{20}\text{Incorrect}\sum_{i=1}^{n} x_i^2 - 100$$

$$\Rightarrow \text{ Incorrect } \sum_{i=1}^{n} x_i^2 = 2080$$

$$\therefore \text{ Correct } \sum_{i=1}^{n} x_i^2 = \text{ Incorrect } \sum_{i=1}^{n} x_i^2 - (8)^2$$

$$= 2080 - 64$$

$$= 2016$$

$$\therefore \text{ Correct Standard deviation} = \sqrt{\frac{\text{correct}\sum x_i^2}{n} - (\text{correct Mean})^2}$$

$$= \sqrt{\frac{2016}{19} - (10.1)^2}$$

$$= \sqrt{106.1 - 102.1}$$

$$= \sqrt{4.09}$$

$$= 202$$

Q. 2. Let x_1 and x_2 be two independent random variables with variances Var $(x_1) = k$, Var $(x_2) = 2$. If the variance of $y = 3x_2 - x_1$ is 25, then find k.

Sol. Var $(y) = 3x_2 - x_1 = 25$

$\Rightarrow$ 3 var (x_2) – Var $(x_1) = 25$

$\Rightarrow \quad 3k - 2 = 25$

$\Rightarrow \quad 3\text{ k} = 27$

$$\boxed{k = 9}$$

Q. 3. *(a)* State and prove the Addition theorem of probability.

Sol. Prove the formula for general additional rule of three events

$$P(A \cup B \cup C) = P(A) + P(B) + P(A \cap B) - P(A \cap C) - P(A \cap B) + P(A \cap B \cap C)$$

Now

$$P(A \cup B \cup C)$$

$= P(A \cup (B \cup C))$

$= P(A) + P(B \cup C) - P(A \cap (B \cup C))$

$= P(A) + P(B) + P(C) - P(B \cap C) - P(A \cap (B \cup C))$

$= P(A) + P(B) + P(C) - P(B \cap C) - P((A \cap B) \cup (A \cup C))$

$= P(A) + P(B) + P(C) - P(B \cap C) - P(A \cap B) - P(A \cap C) + P((A \cap B) \cap (A \cap C))$

$= P(A) + P(B) + P(C) - P(B \cap C) - P(A \cap B) - P(A \cap C) + P(A \cap B \cap C)$

***(b)* Suppose that *A* and *B* are two independent events, associated with a random experiment. The probability of occurrence of event *A* or *B* is 0.8, while the probability of occurrence of event A is 0.5. Determine the occurrence of probability of event *B*.**

Sol. If events *A* and *B* are associated

$$P(A \text{ or } B) = P(A) + P(B)$$

$$\cdot 8 = \cdot 5 + P(B)$$

$$P(B) = \cdot 8 - \cdot 5$$

$$\boxed{P(B = \cdot 3)}$$

Q. 4. *(a)* What do you understand by a random variable? Define the types of random variables.

Ans. Ref.: See Chapter 3, Page No. 40, "Random variable", Page No. 42, "Discrete Random Variable" Page No. 43, "Continuous Random Variable".

***(b)* A bag contains 10 white and 3 black balls. Balls are drawn one by one without replacement till all the black balls are drawn. Find the probability that all black balls are drawn by the 6th draw.**

Sol. 1st Draw to select 1 balck ball then

$$= 13\, c_1 = 13$$

2nd Draw to select 2 black ball then

$$12\, c_1 = 12$$

3rd Draw to select 3rd black ball then

$$= 11 c_1 = 11$$

Next 4th, 5th and 6th Draw back ball are finished. So

Total black ball probability $= 13c_1 \times 12c_1 \times 11c_1$

$$= \frac{13.12!}{12!} \times \frac{12.11!}{11!\cdot} \times \frac{11.10!}{10!}$$

$$= 13 \times 12 \times 11$$

$$= 1716$$

Q. 5. A survey of 64 medical labs revealed that the mean price charged for a certain test was ₹ 120, with a standard deviation of ₹ 60. Test whether the data indicates that the mean price of this test is more than ₹ 100 at 5% level of significance.

Sol. $H_0 : \mu = 120$ *i.e.* average mean price charged for a certain test.

$H_1 : \mu < 120$

Now

$$Z_0 = \frac{\sqrt{n}(\bar{x} - \mu)}{\sigma}$$

Where

$$n = 64$$

$$\bar{X} = 100$$

$$\mu = 120$$

$$\sigma = 60$$

The critical region at 5% level of significance for this medical test is:

$W : Z_0 < -1.96$

$$\therefore \quad Z_0 = \frac{\sqrt{64}(100 - 120)}{60}$$

$$= \frac{8 \times -20}{60}$$

$$= -2.66.$$

At this does not tall under the critical region, H_o is accepted.

SECTION B

Q. 6. Describe the following tests in detail:

***(a)* Paired t-test**

Sol. Paired sample t-test is a statistical technique that is used to compare two population means in the case of two samples that are correlated.

Sample Preview of The Chapter

Published by:

NEERAJ PUBLICATIONS

www.neerajbooks.com

STATISTICAL TECHNIQUES

STATISTICS AND PROBABILITY

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	𝍸	3	3/80 = 0.0375
2	𝍸 |||	8	8/10 = 0.1000
3	𝍸 𝍸	10	10/80 = 0.1250
4	𝍸 𝍸 𝍸	15	15/80 = 0.1875
5	𝍸 𝍸 𝍸 𝍸	20	20/80 = 0.2500
6	𝍸 𝍸 |	11	11/80 = 0.1375
7	𝍸 ||	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