Contents

[1. Measures of Central Tendency and Dispersion 2](#_Toc432432497)

[1.1 Frequency & Frequency Distribution 2](#_Toc432432498)

[1.1.1 Frequency 2](#_Toc432432499)

[1.1.2 Frequency Distribution 2](#_Toc432432500)

[1.2 Arithmetic Mean (A.M.) 3](#_Toc432432501)

[1.2.1 Simple Arithmetic Average 3](#_Toc432432502)

[1.2.2 Weighted Arithmetic Average 3](#_Toc432432503)

[1.2.3 Combined Mean 4](#_Toc432432504)

[1.3 Geometric Mean (G.M.) 5](#_Toc432432505)

[1.3.1 In case of discrete series without frequency 5](#_Toc432432506)

[1.3.2 In case of discrete series with frequency 5](#_Toc432432507)

[1.3.3 In case of continuous series 5](#_Toc432432508)

[1.3.4 This formula can be used for all three cases above 5](#_Toc432432509)

[1.4 Harmonic Mean (H.M.) 6](#_Toc432432510)

[1.4.1 For discrete series without frequency 6](#_Toc432432511)

[1.4.2 For discrete series with frequency 6](#_Toc432432512)

[1.5 Median 7](#_Toc432432513)

[1.5.1 For discrete series (with odd number) 7](#_Toc432432514)

[1.5.2 For discrete series (with even number) 7](#_Toc432432515)

[1.5.3 For continuous series 7](#_Toc432432516)

[1.6 Mode 8](#_Toc432432517)

[1.6.1 For discrete data with or without frequency 8](#_Toc432432518)

[1.6.2 For continuous series, mode is given by 8](#_Toc432432519)

[1.7 Quartiles 9](#_Toc432432520)

[1.8 Deciles 9](#_Toc432432521)

[1.9 Percentiles 9](#_Toc432432522)

[1.10 Range and Range Deviation 9](#_Toc432432523)

[1.11 Quartile Deviation 9](#_Toc432432524)

[1.12 Mean Deviation 9](#_Toc432432525)

[1.13 Standard Deviation 9](#_Toc432432526)

[2 Probability Theory 10](#_Toc432432527)

[2.1 THEORY 10](#_Toc432432528)

[2.1.1 Deterministic Experiment 10](#_Toc432432529)

[2.1.2 Random Experiment 10](#_Toc432432530)

[2.1.3 Random Variable or Stochastic Variable 10](#_Toc432432531)

[2.1.4 Sample Space or Event Space or Possibility Set 10](#_Toc432432532)

[2.1.5 Event 10](#_Toc432432533)

[2.1.6 Subset 10](#_Toc432432534)

[2.1.7 Elementary or Simple events 10](#_Toc432432535)

[2.1.8 Compound or Composite events 10](#_Toc432432536)

[2.1.9 Equally likely or Equiprobable events 10](#_Toc432432537)

[2.1.10 Mutually Exclusive events or disjoint events 10](#_Toc432432538)

[2.1.11 Mutually Exhaustive Event 10](#_Toc432432539)

[2.1.12 Union of events Denoted by 10](#_Toc432432540)

[2.1.13 Intersection of events Denoted by 10](#_Toc432432541)

[3 Probability Distributions 12](#_Toc432432542)

[4 Sampling and Sampling Distributions 13](#_Toc432432543)

[5 Estimations 14](#_Toc432432544)

[6 Test of Hypotheses 15](#_Toc432432545)

[6.1 Test of Hypotheses and Confidence Interval for Mean of single population 15](#_Toc432432546)

[6.1.1 Mean test, Single population, σ known 15](#_Toc432432547)

[6.1.2 Mean test, Single population, σ unknown 15](#_Toc432432548)

[6.2 Test of Hypotheses and Confidence Interval for Means of two population 15](#_Toc432432549)

[6.2.1 Mean test, Two independent population, σ1 & σ2 known 15](#_Toc432432550)

[6.2.2 Mean test, Two independent population, σ1 & σ2 unknown but σ1= σ2 15](#_Toc432432551)

[7 Analysis of Variance 16](#_Toc432432552)

[8 Correlation and Regression 17](#_Toc432432553)

[10 Quality and Quality Control 18](#_Toc432432554)

[11 Nonparametric Methods 19](#_Toc432432555)

[12 Time Series and Forecasting 19](#_Toc432432556)

[13 Index Numbers 19](#_Toc432432557)

[14 Decision Theory 19](#_Toc432432558)

# Measures of Central Tendency and Dispersion

## Frequency & Frequency Distribution

### Frequency

The number of units associated with each value of the variable is called frequency of that value.

For Example if a variable takes the value 15 and the value 15 occurs 3 times, then 3 is called the frequency of the value. Frequency is denoted by ‘f’ and variable by ‘x’

### Frequency Distribution

“A frequency distribution is a statistical table which shows, the set of all distinct values of the variable arranged in order of magnitude, either individually or in groups with their corresponding frequencies.” – Croxton & Cowden

A frequency distribution can be classified into:

1. Series of Individual Observations
2. Discrete frequency Distribution
3. Continuous Frequency distribution

#### Series of Individual observations

Series of individual observation is a series where items are listed one after each observation. This is called as arrays.

Eg.

#### Discrete Frequency Distribution

If the data series is presented indicating its exact measurement of units, then it is called as discrete frequency distribution. Eg.

#### Continuous Frequency Distribution

Continuous data series is one where the measurements are only approximations and are expressed in class intervals within certain limits. In continuous frequency distribution the classes are theoretically continuous from the starting of frequency distribution till the end without break. Eg.

Class Limit – The smallest and the largest value in the class.

Class Interval/Width – Difference between the upper and the lower limit in a class.

Methods of forming Class interval:

* Exclusive or Overlapping Method – The upper limit of one class is the lower limit of the next class. This method makes continuity of data. Eg.
* Inclusive or Non-Overlapping Method – Eg.

#### Other Terms:

Relative Frequency Distribution – Eg.

Percentage Frequency Distribution – Eg.

Frequency density Distribution – Eg.

Cumulative Frequency Distribution – Eg.

## Arithmetic Mean (A.M.)

Arithmetic mean is defined as the sum of all values divided by number of values and is represented by. Arithmetic mean is also called arithmetic average. It is the most commonly used measure of central tendency. There are two types of arithmetic average :

1. Simple arithmetic average
2. Weighted arithmetic average

### Simple Arithmetic Average

Arithmetic mean is simply sometimes referred as mean. Let xiis the variable which takes values over ‘n’ times, then the arithmetic mean is denoted by.

Which is equal to

Where ∑ is the Greek symbol for sigma which denotes summation of all xivalues.

#### For discrete data, the arithmetic mean is given by :

Where

n = number of observations

#### For discrete data with frequency, the arithmetic mean is given by :

Where is the total frequency

#### For continuous data with frequency, the arithmetic mean is given by :

Where d = (X – Assumed mean) / width of class interval

C.I. = the width of class interval

X = the mid value of class interval

A = the assumed mean

### Weighted Arithmetic Average

The weighted mean is computed by considering the relative importance of each of the values of the total value. The arithmetic mean gives equal importance to all the items of distribution. In certain cases relative importance of items is not the same. To give relative importance, weightage may be given to variables depending on cases. Thus weightage represents the relative importance of items.

The weighted arithmetic mean is computed by following equation :

Let are the variables and are the respective weights assigned. Then

### Combined Mean

Combined arithmetic mean can be computed if we know the mean and number of items in each group of data. The following equation is used to compute the combined mean.

Let be the mean of first and second group of data containin*g* items respectively then combined mean is equal to

## Geometric Mean (G.M.)

GM is nth root of product of quantities of the series. It is observed by multiplying the values of the items together and extracting the root of the product corresponding to the number of items. GM is never larger than AM, if there are zeros and negative numbers in the series the GM cannot be used.

In the field of business management various problems often arise relating to average percentage rate of change over a period of time. In such case the AM is not a appropriate average to employ. GM is highly useful in the construction of Index numbers.

The GM of series of “n” positive numbers is given by :

### In case of discrete series without frequency

### In case of discrete series with frequency

Where

### In case of continuous series

Where and are the mid points of class intervals.

### This formula can be used for all three cases above

## Harmonic Mean (H.M.)

It is the total number of items of a value divided by the sum of reciprocal of values of variables. It is a specified average which solves problems involving variables expressed in within ‘time rates’ that vary according to time. E.g. – Speed in Km/hr, price/unit.

The HM of series of “n” numbers is given by :

### For discrete series without frequency

### For discrete series with frequency

Where are the corresponding frequencies for values of *i* equal to 1 to *n*.

Relation between A.M., G.M. & H.M.

1. If all the items in a variable are same then
2. If the size vary then

## Median

Median of distribution is that value of the variate, which divides it into two equal parts. In terms of frequency curve, the ordinate drawn at median divides the area under the curve into two equal parts. Median is a positional average because its value depends upon the position of an item and not on its magnitude. Median of a set of values is the values is the value which is the middle most value when arranged in the ascending order of magnitude. Denoted by ‘M’.

### For discrete series (with odd number)

Step 1 : Arrange the numbers according to ascending order or create a cumulative frequency table.

Step 2 : Calculate n

Step 3 : value becomes the median.

Step 4 : Value comes in decimal, hence two values of the observation is divided by 2 and median comes out.

### For discrete series (with even number)

Step 1 : Arrange the numbers according to ascending order or create a cumulative frequency table.

Step 2 : Calculate n

Step 3 : value is found out.

Step 4 : Value comes as a whole number

### For continuous series

Step 1 : Calculate n

Step 2 : Calculate the cumulative frequency.

Step 3 : For determining the size we should take nth/2 item and median class is located accordingly with reference to the cumulative frequency.

Step 4 :

Where where, is left end point of n/2 class and , is right end point of previous class

h = Class width, *f =* frequency of median class,

C = cumulative frequency of class preceding the median class.

## Mode

Mode is the value which occurs with the maximum frequency. The modal class of a frequency distribution is the class with the highest frequency. Mode is the value of variable which is repeated greatest number of times in a series. It lies at the position of greatest density. Terms such as ‘bi-modal’ and ‘multi-modal’ are associated with two or more modes with the same frequency. Denoted by ‘z’.

It represents fashion and often it is used in business. For example shoe and readymade garment manufacturers will like to know the modal size of the of the people to plan their operations.

### For discrete data with or without frequency

Mode is the value with the highest frequency

### For continuous series, mode is given by

Where *L.L.* = Lower limit of modal class

= frequency of modal class

= frequency of previous class

= frequency of succeeding class

*C.I.*= width of class interval

Relation between mean median and mode

## Quartiles

A measure that divides an array into four equal parts is known as quartile. Each portion contains equal number of items. The first, second and third point are termed as first quartile (Q1), second quartile (Q2) known as the median and the third quartile (Q3).

Diagrammatic Representation –

## Deciles

## Percentiles

## Range and Range Deviation

## Quartile Deviation

## Mean Deviation

## Standard Deviation

# Probability Theory

## THEORY

### Deterministic Experiment

### Random Experiment

### Random Variable or Stochastic Variable

### Sample Space or Event Space or Possibility Set

### Event

### Subset

### Elementary or Simple events

### Compound or Composite events

### Equally likely or Equiprobable events

### Mutually Exclusive events or disjoint events

A set of events are said to be mutually exclusive is the occurrence of one of event prevents the occurrence of the others. Thus for a set of mutually exclusive events no two events can occur simultaneously. The occurrence of any one of the event totally excludes the occurrence of the other events. Mutually exclusive events cannot occur together.

### Mutually Exhaustive Event

### Union of events Denoted by

Union of two or more events is the event of occurrence of atleast one of events.

### Intersection of events Denoted by

Intersection of two or more events is the event of simultaneous occurrences of all events.

Case 1 .

If a sales manager would like to know the probability by which they will exceed the target for product A or B. This is answered by addition rule.

The Additional Theory of probability states that :

1. Theorem 1 – If A and B are two non-mutually exclusive events then the probability of the occurrence of atleast one of events A or B is given by:
2. Theorem 2 – If A and B are two mutually exclusive events then the probability of the occurrence of either A or B is given by:

Case 2 .

If a sales manager would like to know the probability by which they will exceed the target for product A and B. This is answered by multiplication rule.

The Multiplication Theory of probability states that :

1. Theorem 3 - If A and B are two independent events then the probability of the occurrence of events A or B is given by:

Case 3 .

Sometimes we wish to know the probability that the price of a particular petroleum product will rise, given that the Finance minister has increased the petrol price. Such answers can be found out by conditional probabilities. This is known as Conditional Probability.

The probability of simultaneous occurrence of two or more events is known as compound or joint probability. The usual notation of compound probability of two events A and B is or .

Conditional Probability of occurrence of an event A given that the event B has already occurred is given by . Here, A and B are dependent events.

The probability of joint occurrence of two events A and B

# Probability Distributions

# Sampling and Sampling Distributions

# Estimations

# Test of Hypotheses

## Test of Hypotheses and Confidence Interval for Mean of single population

### Mean test, Single population, σ known

### Mean test, Single population, σ unknown

## Test of Hypotheses and Confidence Interval for Means of two population

### Mean test, Two independent population, σ1 & σ2 known

### Mean test, Two independent population, σ1 & σ2 unknown but σ1= σ2

This is also known as Fisher’s Tstatistics

# Analysis of Variance

# Correlation and Regression

## Correlation

Correlation analysis determines the strength of a relation between two variables. It determines the degree to which two variables are related.

## Regression

Regression analysis determines the effect of one variable on the other i.e. the effect of an independent variable on a dependent variable. We develop an estimating equation that relates the known variable to the unknown. Regression analysis attempts to predict the average for a variablefor a given variable or variables, , etc. The dependent variable is plotted on the y-axis and independent on x-axis.

### Types of Regression

1. Simple Linear Regression
2. Multiple Linear Regression
3. Non Linear Regression
4. Polynomial Regression
5. Curvilinear Regression
6. Logistic Regression
7. Probit Repression – Same as logistic but follows Normal Distribution
8. Ordinal Logistic Regression – Known as ordered logistic regression used when dependent variable is ordinal in nature, eg. Very good, Good, Average, Below Average. It is similar to logistic regression except for the fact that it has more than two values for the dependent variable.
9. Ordinal Probit Regression – Known as Ordered probit Model it is an extension to probit model.

Regression types depend on the a. No. of dependent variable b. Shape of the regression line and c. Type of dependent variable.

# Quality and Quality Control

# Nonparametric Methods

# Time Series and Forecasting

# Index Numbers

# Decision Theory