

Biyani's Think Tank

A concept based exclusive material

Business Statistics

Sapna Gupta

M.Com. (ABHD)

Lecturer

Deptt. of Commerce & Management

Biyani Girls College, Jaipur



Biyani's
Group of Girls' Colleges

Published by :

Think Tanks

Biyani Group of Colleges

Concept & Copyright :

©**Biyani Shikshan Samiti**

Sector-3, Vidhyadhar Nagar,

Jaipur-302 023 (Rajasthan)

Ph : 0141-2338371, 2338591-95 • Fax : 0141-2338007

E-mail : acad@biyanicolleges.org

Website :www.gurukpo.com; www.biyanicolleges.org

First Edition : 2009

While every effort is taken to avoid errors or omissions in this Publication, any mistake or omission that may have crept in is not intentional. It may be taken note of that neither the publisher nor the author will be responsible for any damage or loss of any kind arising to anyone in any manner on account of such errors and omissions.

Laser Type Setted by :

Biyani College Printing Department

Preface

I am glad to present this book, especially designed to serve the needs of the students. The book has been written keeping in mind the general weakness in understanding the fundamental concepts of the topics. The book is self-explanatory and adopts the “Teach Yourself” style. It is based on question-answer pattern. The language of book is quite easy and understandable based on scientific approach.

This book covers basic concepts related to the microbial understandings about diversity, structure, economic aspects, bacterial and viral reproduction etc.

Any further improvement in the contents of the book by making corrections, omission and inclusion is keen to be achieved based on suggestions from the readers for which the author shall be obliged.

I acknowledge special thanks to Mr. Rajeev Biyani, *Chairman* & Dr. Sanjay Biyani, *Director (Acad.)* Biyani Group of Colleges, who are the backbones and main concept provider and also have been constant source of motivation throughout this Endeavour. They played an active role in coordinating the various stages of this Endeavour and spearheaded the publishing work.

I look forward to receiving valuable suggestions from professors of various educational institutions, other faculty members and students for improvement of the quality of the book. The reader may feel free to send in their comments and suggestions to the under mentioned address.

Author

Syllabus

B.Com. Part-I

BUSINESS STATISTICS

SECTION-A

1. **Introduction of Statistics** : Growth of Statistics, Definition, Scope, Uses, Misuses and Limitation of Statistics, Collection of Primary & Secondary Data, Approximation and Accuracy, Statistical Errors.
2. **Classification and Tabulation of Data** : Meaning and Characteristics, Frequency Distribution, Simple and Manifold Tabulation.
3. **Measuring of Central Tendency** : Arithmetic Mean (Simple and Weighted), Median (including quartiles, deciles and percentiles), Mode, Geometric and Harmonic Mean – Simple and Weighted, Uses and Limitations of Measures of Central Tendency.

SECTION-B

4. **Measures of Dispersion** : Absolute and Relative Measures of Dispersion; Quartile Deviation, Mean Deviation, Standard Deviation and their Coefficients; Uses and Interpretation of Measures of Dispersion.
5. **Skewness** : Measures of Skewness.
6. **Index Numbers** : Meaning and Uses of Index Numbers, Simple and Weighted Price Index Numbers, Methods of Construction of Index Numbers, Average of Relatives and Aggregative Methods, Problems in Construction of Index Numbers, Selection of Variables, Base, Weight, Average and Formula; Fishers Ideal Index Number, Base Shifting, Splicing and Deflating.

SECTION-C

7. **Correlation** : Meaning and Significance, Scatter Diagram, Correlation Graph, Karl Pearson's Coefficient of Linear Correlation between two Variables in Grouped and Ungrouped Data by Direct and Shortcut Methods, Coefficient of Correlation by Spearman's, Rank Differences Method and Concurrent Deviation Method.
8. **Linear Regression.**
9. **Presentation of Data**, Diagrammes/Graphs of Frequency Distribution and Histograms Ratio Scale Graph.

□ □ □

Content

S.No.	Name of Topic	Page No.
	Examination Paper 2008	7-10
1.	Introduction of Statistics	11-14
2.	Collection and Editing of Data	15-20
3.	Classification and Tabulation of Data	21-25
4.	Measures of Central Tendency	26-32
5.	Measures of Dispersion	33-39
6.	Measures of Skewness	40-43
7.	Index Numbers	44-52
8.	Correlation	53-58
9.	Linear Regression	59-62
10.	Diagrammatic and Graphic Presentation	63-72
11.	Practical Exercise	73-78

B.Com.(Part I) EXAMINATION, 2008
(Faculty of Commerce)
(Three – Year Scheme of 10+2+3 Pattern)

PAPER 57
ACCOUNTANCY AND BUSINESS STATISTICS
Second Paper
(Business Statistics)

OBJECTIVE PART-I
Time Allowed : One Hour
Maximum Marks – 40

Attempt all questions

-
- 1 Answer the following questions in not more than 20 words :
- (i) “Ram has secured 90 percent marks in Statistics, whereas his height is 1.9 metres and weight 58 kg.” Would you call these as data, state with reasons.
 - (ii) Write two causes of distrust of statistics.
 - (iii) Name any four sources of collection of secondary data.
 - (iv) What is bivariate frequency distribution?
 - (v) The mean wages of 100 labourers working in a factory running two shifts of 60 and 40 labourers respectively is Rs.38, the mean of 60 labourers working in the morning shift is Rs.40. Find the mean wages of 140 labourers working in the evening shift.
 - (vi) If the variance of a series is 144 and the mean is 50, find out the coefficient of variation.
 - (vii) What is the formula given by Bowley for measuring skewness?
 - (viii) What is the meaning of deflating of index numbers?
 - (ix) Why are there two regression lines?
 - (x) What is cumulative frequency curve or ogive curve?

- 2 Answer the following questions in not more than 50 words each :
- (i) What does the law of statistical regularity show?
 - (ii) Draw a blank table to present information regarding the college students according to :
 - (a) Faculty – Arts, Commerce and Science
 - (b) Sex – Male and Female
 - (c) For 2 years – 2005 and 2006
 - (iii) An airplane travels distance of 2500, 1200 and 500 kms. at speeds 500, 500 and 250 kms. per hour respectively. Find the average speed for the entire trip.
 - (iv) When is the false base line used?
 - (v) What is the difference between histogram and histogram?

PAPER 56
DESCRIPTIVE PART-II
Time Allowed : Two Hours
Maximum Marks – 60

*Attempt any three questions, selecting one question from each Section.
 Each question carries 20 marks*

SECTION-A

- 3 What is 'Statistical Sampling'? Describe the various methods of statistical sampling.
- 4 Find the missing frequency in the following distribution, if $N = 100$ and median is 30 :

Marks	N o. of Students
0 – 10	10
10 – 20	?
20 – 30	25
30 – 40	30
40 – 50	?
50 – 60	10

SECTION-B

- 5 Calculate the standard deviation and its coefficient from the following data :

Wages in Rs.	No. of Persons
48 and above	5
40 and above	15
32 – 40	20
16 – 32	45
8 – 24	32
Less than 16	20
Less than 8	8

- 6 Construct Fisher's ideal index no. from the following table :

Articles	Base year		Current Year	
	Price	Quantity	Price	Quality
Food	16	50	24	45
Clothing	18	30	24	25
Rent	20	5	15	8
Fuel	10	6	12	6
Misc.	10	10	14	12

SECTION-C

- 7 The results of B.Com. Examination is given below form the following data. Calculate coefficient of correlation between and success in the Examination :

Age in years	% of failure
18	38
19	40
20	35
21	32
22	34
23	37
24	42
25	46

26

52

27

56

- 8 From the data given below, find the expected value of X when value of Y is 15, and find the expected value of Y when the value of X is 30.

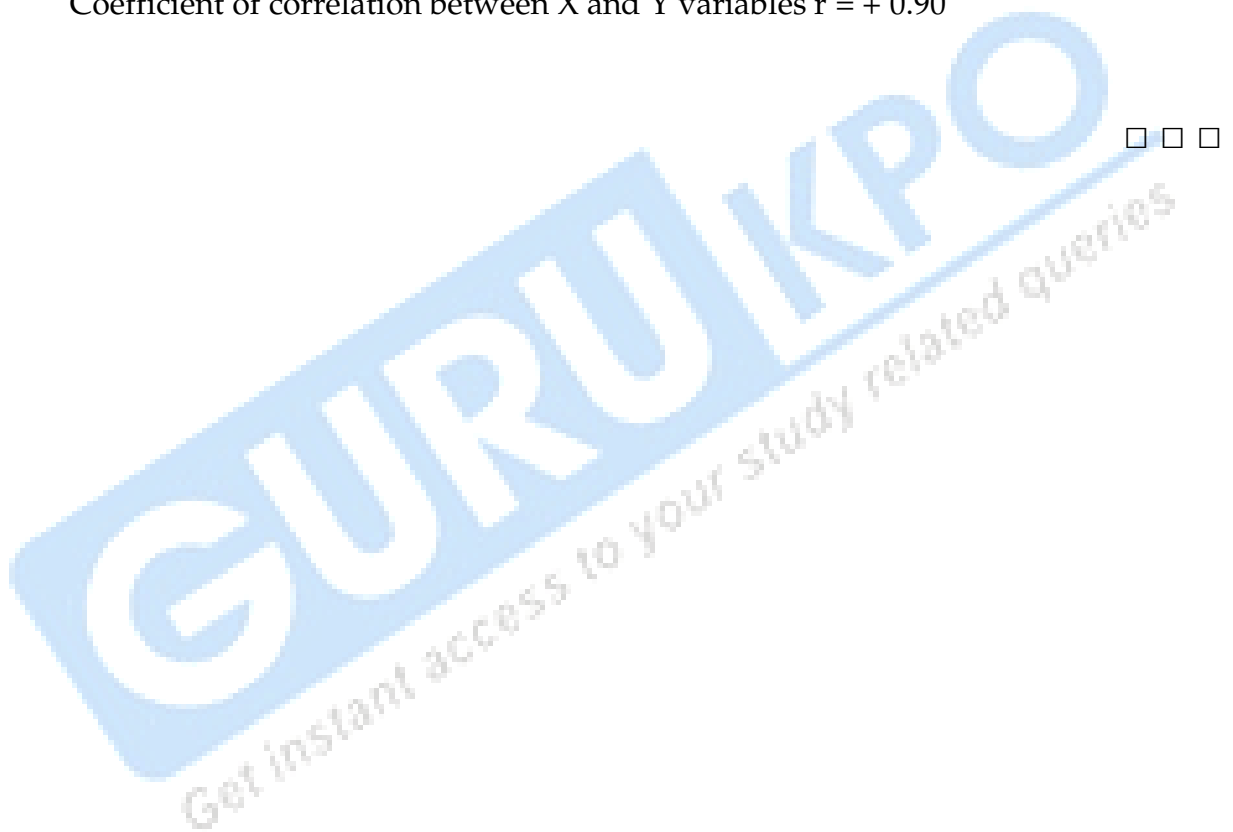
Arithmetic mean of X variable = 40

Arithmetic mean of Y variable = 20

S.D. of X variable (σ_x) = 5

S.D. of Y variable (σ_y) = 7

Coefficient of correlation between X and Y variables $r = + 0.90$



Introduction of Statistics

Q.1 Define 'Statistics' and give characteristics of 'Statistics'.

Ans.: 'Statistics' means numerical presentation of facts. Its meaning is divided into two forms - in plural form and in singular form. In plural form, 'Statistics' means a collection of numerical facts or data example price statistics, agricultural statistics, production statistics, etc. In singular form, the word means the statistical methods with the help of which collection, analysis and interpretation of data are accomplished.

Characteristics of Statistics -

- a) Aggregate of facts/data
- b) Numerically expressed
- c) Affected by different factors
- d) Collected or estimated
- e) Reasonable standard of accuracy
- f) Predetermined purpose
- g) Comparable
- h) Systematic collection.

Therefore, the process of collecting, classifying, presenting, analyzing and interpreting the numerical facts, comparable for some predetermined purpose are collectively known as "Statistics".

Q.2 What is meant by 'Data'?

Ans.: Data refers to any group of measurements that happen to interest us. These measurements provide information the decision maker uses. Data are the foundation of any statistical investigation and the job of collecting data is the same for a statistician as collecting stone, mortar, cement, bricks etc. is for a builder.

Q.3 Discuss the Scope of Statistics.

Ans.: The scope of statistics is much extensive. It can be divided into two parts –

- (i) **Statistical Methods** such as Collection, Classification, Tabulation, Presentation, Analysis, Interpretation and Forecasting.
- (ii) **Applied Statistics** – It is further divided into three parts:
 - a) **Descriptive Applied Statistics** : Purpose of this analysis is to provide descriptive information.
 - b) **Scientific Applied Statistics** : Data are collected with the purpose of some scientific research and with the help of these data some particular theory or principle is propounded.
 - c) **Business Applied Statistics** : Under this branch statistical methods are used for the study, analysis and solution of various problems in the field of business.

Q.4 Give reasons for distrust in Statistics.

Ans.: By distrust of statistics we mean lack of confidence in statistical statements and statistical methods. It is often commented by people

“Statistics can prove anything.”

“There are three type of lies – lies, damned lies and statistics – wicked in the order of their naming.”

The **main reasons** for such views are -

- a) Figures are convincing, and therefore people are easily led to believe them.
- b) Ignorance of limitation of statistics.
- c) Lack of test of accuracy.
- d) Contradiction of data from actual circumstances.
- e) Lack of specific ability to arrive at correct and appropriate results.
- f) Can easily be manipulated.

Q.5 Discuss the functions and importance/utility of Statistics.

Ans.: Statistical methods are used not only in the social, economic and political fields but in every field of science and knowledge. Statistical analysis has become more significant in global relations and in the age of fast developing information technology.

According to **Prof. Bowley**, *“The proper function of statistics is to enlarge individual experiences”*.

Following are some of the **important functions of Statistics** :

- a) To provide numerical facts.
- b) To simplify complex facts.
- c) To enlarge human knowledge and experience.
- d) Helps in formulation of policies.
- e) To provide comparison.
- f) To establish mutual relations.
- g) Helps in forecasting.
- h) Test the accuracy of scientific theories.
- i) To study extensively and intensively.

The use of statistics has become almost essential in order to clearly understand and solve a problem. Statistics proves to be much useful in **unfamiliar fields of application and complex situations** such as :-

- a) Planning
- b) Administration
- c) Economics
- d) Trade & Commerce
- e) Production management
- f) Quality control
- g) Helpful in inspection
- h) Insurance business
- i) Railways & transport Co
- a) Banking Institutions
- b) Speculation and Gambling
- c) Underwriters and Investors
- d) Politicians & social workers.

□ □ □

CHAPTER-2

Collection and Editing of Data

Q.1 What do you mean by Collection of Data? Differentiate between Primary and Secondary Data.

Ans.: Collection of data is the basic activity of statistical science. It means collection of facts and figures relating to particular phenomenon under the study of any problem whether it is in business economics, social or natural sciences.

Such material can be obtained directly from the individual units, called primary sources or from the material published earlier elsewhere known as the secondary sources.

Difference between Primary & Secondary Data

	Primary Data	Secondary Data
Basis nature	Primary data are original and are collected for the first time.	Data which are collected earlier by someone else, and which are now in published or unpublished state.
Collecting Agency	These data are collected by the investigator himself	Secondary data were collected earlier by some other person.
Post collection alterations	These data do not need alteration as they are according to the requirement of the investigation	These have to be analyzed and necessary changes have to be made to make them useful as per the requirements of investigation.
Time & Money	More time, energy and money has to be spent in collection of these data.	Comparatively less time and money is to be spent.

Q.2 What do you mean by Questionnaire? Give merits of a good Questionnaire.

Ans.: Questionnaire is a document containing questions related to the specific requirement of a statistical investigation for collection of information which is filled by the informants personally.

Requirements of a good **questionnaire** :-

- Questions should be simple, clear and short.
- Simple alternative or multiple choice questions.
- Unambiguous and precise.
- Questions should be in sequence.
- Directly relative questions.
- Test of accuracy.
- No restricted questions affecting personal whims.
- Assurance of secrecy to the informants.
- Probability of a perfect answer.

Q.3 What is Law of 'Statistical Regularity' and the Law of 'Inertia of Large Numbers'?

Ans.: Based on the mathematical theory of probability, Law of Statistical Regularity states that if a sample is taken at random from a population, it is likely to possess the characteristics as that of the population. A sample selected in this manner would be representative of the population. If this condition is satisfied, it is possible for one to depict fairly accurately the characteristics of the population by studying only a part of it.

The Law of 'Inertia of Large Numbers' is a corollary of the law of statistical regularity. It states that, other things being equal, larger the size of sample, more accurate the results are likely to be. This is because when large numbers are considered, the variations in the component parts tend to balance each other and, therefore, the variation in the aggregate is insignificant.

Q.4 What is Random Sampling?

OR

Define Random Sampling.

Ans.: Random Sampling is one in which selection of items is done in such a way that every item of the universe has an equal chance of being selected.

Random Sampling is based on probability and it is free from bias.

The different **methods of Random Sampling** are :-

- a) Lottery method.
- b) Rotating the drum.
- c) By systematic arrangement.
- d) Routed wheel method
- e) By random numbers

Q.5 What do you mean by Statistical Error? Give the difference between Mistake and Error?

Ans.: The difference between the actual value of the figure and its approximated value is called statistical error. For example if number of students of a college is 1,255 but in round figures it is written as 1,300, then the difference is called 'Statistical error'. In the words of Prof. Connor, *"In statistical sense, 'error' means the difference between the approximate value and the true or ideal value, accurate determination of which is not possible"*.

Difference between Mistake and Error

	Basis	Mistake	Error
a)	Nature	Committed deliberately	Not deliberate
b)	Source	Due to use of wrong method	Difference between actual and approximate value.
c)	Estimation	Difficult to estimate	Can be estimated
d)	Prevention	Can be avoided by carefulness	Cannot be avoided easily.

e)	State Occurrence	of	Can be committed at any stage	Creeps in only at the state of collecting, analysis and interpretation.
----	---------------------	----	----------------------------------	--

Q.6 Write a note on the Editing of Primary Data and Secondary Data for the purpose of Analysis and Interpretation.

Ans.: Before analysis and interpretation, it is necessary to edit the data to detect possible errors and inaccuracies, so that accurate and impartial results may be obtained. Thus editing means the process of checking for the errors and omissions and making corrections, if necessary.

The task of editing is a highly specialized one and requires high level of skill and carefulness to attain the proper degree of accuracy.

Editing of Primary Data : While editing primary data, the following points should be considered :-

- Editing for consistency
- Editing for completeness
- Editing for accuracy
- Editing for uniformity or homogeneity

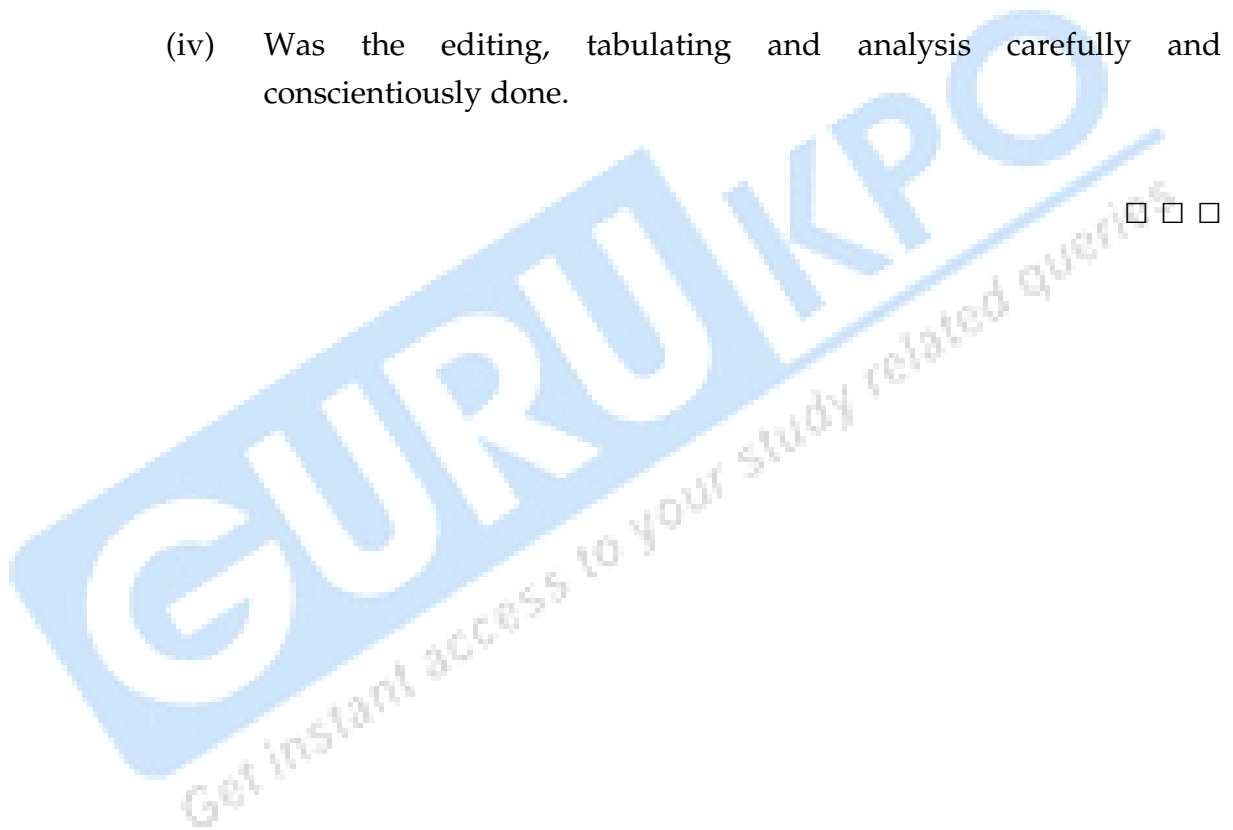
Editing of Secondary Data : Since secondary data have already been obtained, it is highly desirable that a proper scrutiny of such data is made before they are used by the investigator. **Bowley** rightly points out that “*secondary data should not be accepted at their face value*”.

Hence before using secondary data, the investigators should consider the following aspects :-

- **Whether the Data are Suitable for the Purpose of Investigation in View :** Quite often secondary data do not satisfy immediate needs because they have been compiled for other purpose. The variation can be in units of measurement, variation in the date/period to which data is related etc.
- **Whether the Data is Adequate for the Investigation :** Adequacy of data is to be judged in the light of the requirements of the survey and the geographical area covered by the available data. For example if our object

is to study ;the wage rates of the workers in sugar industry in India and the available data cover only the state of Rajasthan, it would not serve the purpose.

- **Whether the Data are Reliable :** The following points should be checked to find out the reliability of secondary data -
 - (i) The collecting agency was unbiased.
 - (ii) The enumerators are properly trained.
 - (iii) A proper check on the accuracy of the field work.
 - (iv) Was the editing, tabulating and analysis carefully and conscientiously done.



Classification and Tabulation of Data

Q.1 What is the meaning of Classification? Give objectives of Classification and essentials of an ideal classification.

Ans.: Classification is the process of arranging data into various groups, classes and sub-classes according to some common characteristics of separating them into different but related parts.

Main objectives of Classification :-

- (i) To make the data easy and precise
- (ii) To facilitate comparison
- (iii) Classified facts expose the cause-effect relationship.
- (iv) To arrange the data in proper and systematic way
- (v) The data can be presented in a proper tabular form only.

Essentials of an Ideal Classification :-

- (i) Classification should be so exhaustive and complete that every individual unit is included in one or the other class.
- (ii) Classification should be suitable according to the objectives of investigation.
- (iii) There should be stability in the basis of classification so that comparison can be made.
- (iv) The facts should be arranged in proper and systematic way.
- (v) Data should be classified according to homogeneity.
- (vi) It should be arithmetically accurate.

Q.2 What is Manifold Classification?

Ans.: When the data are classified into more than two classes according to more than one attribute, it is called manifold classification.

Q.3 What are Class Limits?

Ans.: The two values which determine a class are known as class limits. First or the smaller one is known as lower limit (L_1) and the greater one is known as upper limit (L_2)

Q.4 How many types of Series are there on the basis of Quantitative Classification? Give the difference between Exclusive and Inclusive Series.

Ans.: There are three types of frequency distributions -

- (i) **Individual Series :** In individual series, the frequency of each item or value is only one for example ;marks scored by 10 students of a class are written individually.
- (ii) **Discrete Series :** A discrete series is that in which the individual values are different from each other by a different amount.

For example:

Daily wages	5	10	15	20
No. of workers	6	9	8	5

- (iii) **Continuous Series :** When the number of items are placed within the limits of the class, the series obtained by classification of such data is known as continuous series.

For example

Marks obtained	0-10	10-20	20-30	30-40
No.of students	10	18	22	25

Difference between Exclusive and Inclusive Series

	Exclusive Series	Inclusive Series
Limits	Upper limit of one class is equal to the lower limit of next class.	The two limits are not equal.
Inclusion	The value equal to the upper limit is included in the next class.	Both upper & lower limits are included in the same class.
Conversion	It does not require any conversion.	Inclusive series is converted into exclusive series for calculation purpose.
Suitability	It is suitable in all situations.	It is suitable only when the values are in integers.

Q.5 What is Bivariate Frequency Distribution?

Ans.: A frequency distribution obtained by the simultaneous classification of data according to two characteristics is known as a bivariate frequency distribution.

Q.6 Give Sturges Formula for determining Magnitude of Classes?

Ans.: According to Prof. A. H. Sturges, class interval can be found using the following formula.

$$I = \frac{L - S}{1 + 3.322 \log N}$$

Where -

I = class interval N = No. of observations
L = Largest value S = Smallest value

Q.7 Define Tabulation. State the objectives of Tabulation and kinds of Tables.

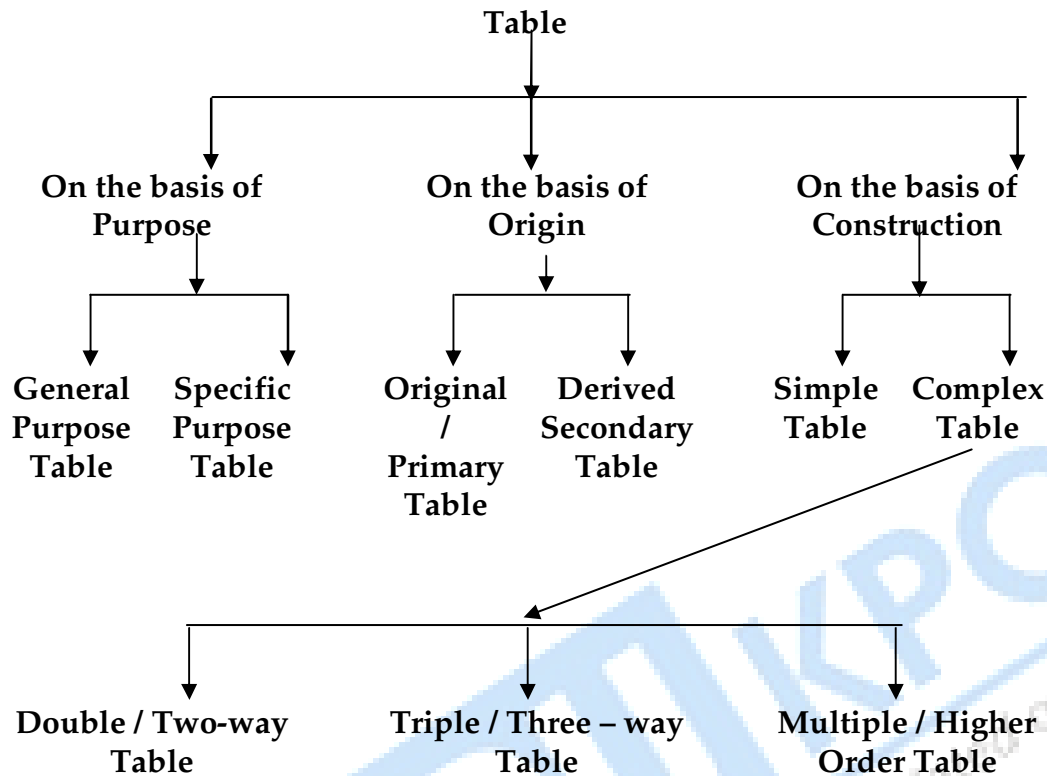
Ans.: According to Blair, *“Tabulation in its broad sense is an orderly arrangement of data in columns and rows.”*

Tabulation is a process of presenting the collected and classified data in proper order and systematic way in columns and rows so that it can be easily compared and its characteristics can be elucidated.

Objects of Tabulation :

- Orderly and systematic presentation of data.
- Making data precise and stable.
- To facilitate comparison.
- To make the problem clear and self evident.
- To facilitate analysis & interpretation of data.

Kinds of Table : The different kinds of tables are shown in the following chart -



Q.8 What are the main parts of a good Table?

Ans.: The number of parts depends mostly on the nature of the data. However, a table should have the following parts.

- (i) **Table No. :** Each table should be numbered so that the table may be referred with that number.
- (ii) **Title :** Every table must be given a suitable title which should be short, clear and complete.
- (iii) **Captions :** Caption refers to the column heading which explains what the column represents.
- (iv) **Stubs :** Stubs are the designations of the rows or row headings.
- (v) **Body :** It is the heart of the table. The body of the table contains the numerical information.
- (vi) **Ruling and Spacing :** Ruling and leaving the space depends on the needs of the topic and makes the table attractive and beautiful.
- (vii) **Footnotes :** In order to explain the figures shown in the table, explanatory notes may be given at the end of the table.
- (viii) **Source :** At the end of the table, the source or origin of given data is mentioned.

□ □ □

Measures of Central Tendency

Q.1 What do you mean by Measures of Central Tendency? Define Arithmetic Mean, Median and Mode.

Ans. : The central tendency of a variable means a typical value around which other values tend to concentrate; hence this value representing the central tendency of the series is called measures of central tendency or average.

According to Clark, *“Average is an attempt to find one single figure to describe whole of figures.”*

Arithmetic Mean (X) : The most popular and widely used measure of representing the entire data by one value is known as arithmetic mean. Its value is obtained by adding together all the items and by dividing this total by the number of items.

Arithmetic mean may be of **two types :**

- Simple Arithmetic mean
- Weighted arithmetic mean

Calculation of Arithmetic Mean :

Individual Series	Discrete & continuous series
<p>Direct Method :</p> $\bar{X} = \frac{\sum X}{N}$ <p>Where –</p> <p style="padding-left: 20px;">X → Values</p> <p style="padding-left: 20px;">N → No. of Items</p> <p>Shortcut Method :</p> $\bar{X} = A + \frac{\sum dx}{N}$ <p>Where –</p>	<p>Direct Method :</p> $\bar{X} = \frac{\sum fX}{N}$ <p>Where –</p> <p style="padding-left: 20px;">fX → Values X frequencies</p> <p style="padding-left: 20px;">N → Total of frequencies</p> <p>Shortcut Method :</p> $\bar{X} = A + \frac{\sum f dx}{N}$

A → Assumed Mean

dx → X - A

Step-Deviation Method :

Not applicable

Step-Deviation Method :

$$\bar{X} = A + \frac{\sum f dx' \times i}{N}$$

Where –

$$dx' \rightarrow \frac{X - A}{i}$$

i → Class interval

Median (M) : Median is that value of the variable which divides the group into two equal parts, one part comprising all values greater than, and the other all values less than the median

Calculation of Median :

a) Individual Series :

- Arrange the variables in ascending or descending order.
- $M = \text{size of } \left(\frac{N+1}{2} \right)^{\text{th}} \text{ item}$

where N → No. of items

b) Discrete Series :

- Arrange the variables in ascending or descending order.
- Calculate cumulative frequencies.
- Apply formula $M = \left(\frac{N+1}{2} \right)^{\text{th}} \text{ item}$; N = Total of frequency.

- The value (X) corresponding to this in the cumulative frequency will be the median.

c) Continuous Series :

- Arrange the variables in ascending or descending order.
- Calculate cumulative frequencies.
- Determine median class by using (N/2).
- Apply formula –

$$M = l_1 + \frac{i(m - c)}{f - c}$$

f

$l_1 \rightarrow$ lower limit of median group

$i \rightarrow$ class interval

$m \rightarrow N/2$

$c \rightarrow$ cumulative frequency preceding the median group.

$f \rightarrow$ frequency of median group

Mode (Z) : Mode is the value that appears most frequently in a series i.e. it is the value of the item around which frequencies are most densely concentrated.

Calculation of Mode :

a) Individual Series :

- (i) By inspection – value repeated most.
- (ii) By converting individual series into discrete series.
- (iii) By empirical relationship between the averages -

$$Z = 3M - 2X$$

b) Discrete Series :

- (i) By inspection – value having highest frequency.
- (ii) By grouping.
- (iii) By empirical relationship.

c) Continuous Series :

- (i) First calculate modal class by inspection or by grouping.
- (ii) Then apply the following formula -

$$l_1 + \frac{\Delta_1}{\Delta_1 + \Delta_2}$$

where $l_1 \rightarrow$ lower limit of modal class

$$\Delta_1 \rightarrow f_1 - f_0$$

$$\Delta_2 \rightarrow f_1 - f_2$$

$f_1 \rightarrow$ frequency of modal class

$f_0 \rightarrow$ frequency of preceding class.

$f_2 \rightarrow$ frequency of succeeding class

$i \rightarrow$ class interval

Q.2 What are the essentials of an Ideal Average.

- Ans.:** (i) Should be easy to understand.
(ii) Clearly and rigidly defined.
(iii) Based on all the observations.
(iv) Simple to compute.
(v) Least affected by fluctuations.
(vi) Capable of further Algebraic treatment.
(vii) Sampling stability.

Q.3 What is Geometric Man? Give Algebraic Characteristics of Geometric Mean and state when Geometric Mean is useful.

Ans.: Geometric mean is the n^{th} root of the product of N items or values.

Calculation of Geometric Mean (G) :

Individual Series	Discrete & Continuous Series
$G = \text{Antilog} \left(\frac{\sum \log X}{N} \right)$	$G = \text{Antilog} \left(\frac{\sum f \log X}{N} \right)$

Algebraic Characteristics of Geometric Mean :

- (i) The product of the items remain unchanged if each item is replaced by geometric mean.
(ii) Geometric mean cannot be found if the value of some item in the series is negative or zero.
(iii) The product of corresponding ratios on either side of the geometric mean is always equal.
(iv) Not affected by changing the sequence of items.

Geometric mean is **appropriate or useful** :-

- When ratios or percentages are to be found.
- In determining rates of increase or decrease.
- When the different values are at vast difference.

Q.4 What is Harmonic Mean? In which circumstances Harmonic Mean is used.

Ans.: Harmonic mean of a series is the reciprocal of the arithmetic mean of the reciprocal of the values of its items.

Calculation of Harmonic Mean (H.M.) :

<p>Individual Series</p> $\text{H.M.} = \text{Reciprocal} \left(\frac{\sum \text{Reci.X}}{N} \right)$	<p>Discrete & Continuous Series</p> $\text{H.M.} = \text{Reciprocal} \left(\frac{\sum (\text{Reci.X} \cdot f)}{N} \right)$
---	--

Harmonic Mean is **used in the following cases :-**

- For determining average speed or velocity.
- To find out average price.
- If the item given in the question which is variable is to be kept as constant in the answer, or vice versa, then harmonic mean will be calculated.

Q.5 What is Partition Value. Give formula for calculating different Partition Values?

Ans.: Values of the items that divide the series into many parts are known as partition values. A variable may be divided into four, five, eight, ten and hundred equal parts known as Quartiles, Quintiles, Octiles, Deciles and Percentiles. The aforesaid partition values gives an idea of the formation of the series which are used in the calculation of dispersion and skewness.

Formula used for Calculating Partition Values :

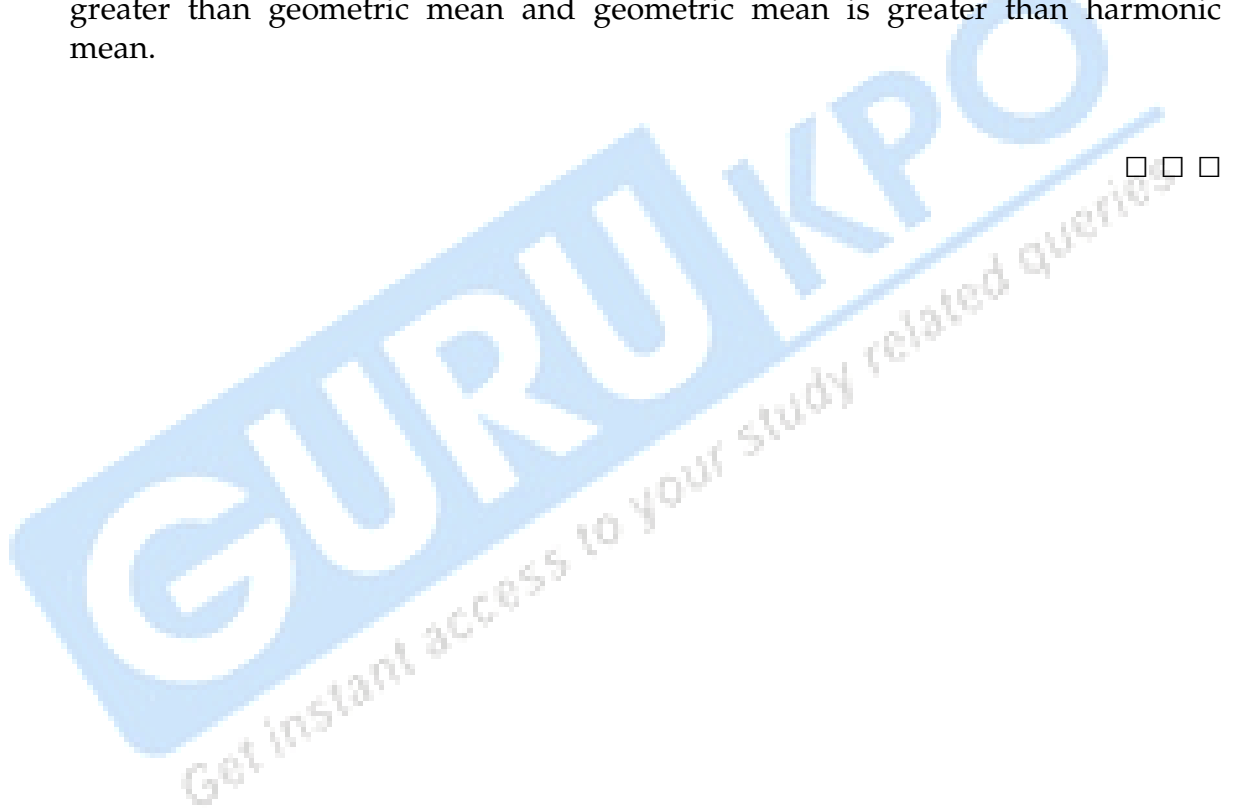
Measures	Individual & Discrete Series	Continuous Series
Quartiles :		
Q_1	Size of $\frac{N+1}{4}$ th item	$q_1 = (N/4)$ th item & $Q_1 = l_1 + \frac{i}{f}(q_1 - c)$
Q_3	Size of $3 \frac{N+1}{4}$ th item	$q_3 = 3 (N/4)$ th item & $Q_3 = l_1 + \frac{i}{f}(q_3 - c)$
Quintiles :		
Q_{n4}	Size of $4 \frac{N+1}{5}$ th item	$qn_4 = 4 (N/5)$ th item & $Q_3 = l_1 + \frac{i}{f}(qn_4 - c)$
Octiles :		
O_2	Size of $2 \frac{N+1}{8}$ th item	$o_2 = 2 (N/8)$ th item & $O_2 = l_1 + \frac{i}{f}(o_2 - c)$
Deciles :		
D_7	Size of $7 \frac{N+1}{10}$ th item	$d_3 = 7 (N/10)$ th item & $D_3 = l_1 + \frac{i}{f}(d_7 - c)$
Percentiles:		
P_{15}	Size of $75 \frac{N+1}{100}$ th item	$p_{15} = 75(N/100)$ th item & $P_{75} = l_1 + \frac{i}{f}(p_{75} - c)$

Q.6 How are Arithmetic Mean, Geometric Mean and Harmonic Mean related to each other? Why is Arithmetic Mean greater than the Geometric Mean and Geometric Mean is greater than the Harmonic Mean.

Ans.: Relation between arithmetic mean, geometric mean and harmonic mean can be given by -

$$(i) \quad \bar{X} > g > h \quad \text{and} \quad (ii) \quad g^2 = \bar{X}.h.$$

While calculating arithmetic mean, the bigger values are given more weightage than the small values, whereas in harmonic mean the smaller values are given much more weightage than to the larger values. Therefore, arithmetic mean is greater than geometric mean and geometric mean is greater than harmonic mean.



CHAPTER-5

Measures of Dispersion

Q.1 What do you mean by Dispersion. Give the meaning of Absolute Measure and Relative Measure with example.

Ans. : Dispersion is a measure of the extent to which the individual item vary from a central value Dispersion is used in two senses, (i) difference between the extreme items of the series and (ii) average of deviation of items from the mean.

Absolute Measure : The figure showing the limit or magnitude of dispersion is known as absolute measure and it is shown in the same unit as those of the original data, example measures of dispersion in the age of students, their height, weight etc.

Relative Measure : For comparative study the concerning absolute measure is divided by the corresponding mean or some other characteristic value to obtain a ratio or percentage, which is known as the relative measure.

Q.2 Explain the various methods for measuring Dispersion. Also give their merits and demerits?

Ans.: Following are the important methods of studying dispersion -

(1) **Numerical Methods :**

- | | |
|--------------------------|----------------------------------|
| a) Methods of limits | b) Methods of average deviation: |
| i) Range | i) Quartile Deviation |
| ii) Inter-quartile range | ii) Mean Deviation |
| iii) Percentile Range | iii) Standard Deviation |

(2) **Graphic Method - Lorenz Curve :**

- i) **Range :** The difference between the value of the smallest item and the value of the largest item of the series is called range.

Range = Largest item – Smallest item

Co-efficient of Range = $\frac{L - S}{L + S}$

Merits of Range :

- Simple and easy to be computed.

Fore more detail:- <http://www.gurukpo.com>

- It takes minimum time to calculate.
- Not necessary to know all the values, only smallest and largest value is required.
- Helpful in quality control of products.

Demerits of Range :

- Not based on all the items.
- Subject to fluctuation/uncertain measure.
- Cannot be computed in case of open-end distributions.
- As it is not based on all the values it is not considered as a good or appropriate measure.

ii) **Inter-Quartile Range :** Inter-quartile range represents the difference between the third-quartile and the first quartile. It is also known as the range of middle 50% values.

$$\text{Inter-quartile range} = Q_3 - Q_1$$

Merits :

- It is easy to calculate.
- Can be measured in open end distributions.
- It is least affected by the uncertainty of the extreme values.

Demerits :

- It does not represent all the values.
- It is an uncertain measure.
- It is very much affected by sampling fluctuations.

iii) **Percentile Range :** It is the difference between the values of the 90th and 10th percentile. It is based on the middle 80% items of the series.

$$\text{Percentile Range} = P_{90} - P_{10}$$

Merits and Demerits : Its use is limited. Percentile range has almost the same merits and demerits as those of inter-quartile range.

iv) **Quartile Deviation / Semi – Interquartile Range :**

Quartile Deviation gives the average amount by which the two quartiles differ from the median. Quartile deviation is an absolute measure of dispersion.

It is calculated by using the formula -

Fore more detail:- <http://www.gurukpo.com>

$$Q.D. = \frac{Q_3 - Q_1}{2}, \text{ coefficient of Q.D.} = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

Merits :

- It is easy to calculate and understand.
- It has a special utility in measuring variation in open end distributions.
- QD is not affected by the presence of extreme values.

Demerits :

- It is very much affected by sampling fluctuations.
- It does not give an idea of the formulation of the series.
- It is not capable of further algebraic treatment.

v) **Mean Deviation (δ) :** Mean Deviation is also known as average deviation or first measure of dispersion. It is the average difference between the items in a distribution and the median mean or mode of that series.

Computation of Mean Deviation :

Individual Series	Discrete & continuous series
<p>a) Deviation from Mean :</p> $\frac{\delta\bar{X} = \sum d\bar{x} }{N}, d\bar{x} = X - \bar{X} $ <p>Coefficient of $\delta\bar{X} = \frac{\delta\bar{X}}{\bar{X}}$</p>	<p>a) Deviation from Mean :</p> $\frac{\delta\bar{X} = \sum f d\bar{x} }{N}$
<p>b) Mean Deviation from Median :</p> $\frac{\delta M = \sum dM }{N}, dM = X - M $ <p>Coefficient of $\delta M = \frac{\delta M}{M}$</p>	<p>b) Mean Deviation from Median :</p> $\frac{\delta M = \sum f dM }{N}$
<p>c) Mean Deviation from Mode :</p> $\frac{\delta Z = \sum dZ }{N}, dZ = X - Z $ <p>Coefficient of $\delta Z = \frac{\delta Z}{M}$</p>	<p>b) Mean Deviation from Mode :</p> $\frac{\delta M = \sum f dZ }{N}$

Where –

$$dx^2 \rightarrow (X - A)^2$$

A \rightarrow Assumed Mean

Computation of Standard Deviation :

$$\text{Coefficient of S.D.} = \sigma / \bar{X}$$

$$\text{Variance} = \sigma^2$$

Coefficient of variation : Coefficient of variation is used for the comparative study of stability or homogeneity in more than two or more series.

$$\text{C.V} = \frac{\sigma}{\bar{X}} \times 100$$

Merits :

- Based on all the items.
- Well-defined and definite measure of dispersion.
- Least affected by sample fluctuations.
- Suitable for algebraic treatment.

Demerits :

- Standard Deviation is comparatively difficult to calculate.
- Much importance is given to the extreme values.

vii) Graphical Method - Lorenz Curve :

This curve was devised by Dr. Max O'Lorenz. He used this technique to show inequality of wealth and income of a group of people. It is simple, attractive and effective measure of dispersion, yet it is not scientific since it does not provide a figure to measure dispersion.

Merits :

- Easy to understand from the graph.
- Comparison can be done in two or more series.
- Attractive & effective.
- Concentration or density of frequency can be known from the curve.

Demerits :

- It is not a numerical measure; therefore it is not definite as a mathematical measure.
- Drawing of curve requires more time and labour.

Q.3 What is the best method of measuring Dispersion. Write the formula for calculating combined S.D.

Ans.: Standard Deviation is the best method of measuring dispersion as deviations taken from mean and algebraic signs are not ignored and it is algebraically correct.

Formula for calculating combined Standard Deviation :

$$\sigma_{12} = \sqrt{\frac{N_1 (\sigma_1^2 + D_1^2) + N_2 (\sigma_2^2 + D_2^2) + \dots}{N_1 + N_2}}$$

Where –

N_1 & $N_2 \rightarrow$ No. of items of two series

σ_1 & $\sigma_2 \rightarrow$ S.D. of the two series

$D_1 \rightarrow \bar{X}_1 - \bar{X}_{12}$

$D_2 \rightarrow \bar{X}_2 - \bar{X}_{12}$

$\bar{X}_{12} \rightarrow$ Combined arithmetic mean

□ □ □

Send your requisition at
info@biyanicolleges.org