STATISTICS

A Text Book for I PUC

2015 - 2016

Department of Pre University Education,

Malleshwaram, Bengaluru -12

www.pue.kar.nic. in

Revised Edition: 2015

STATISTICS

Printed on 80 GSM paper

Director's Message

Dear Students,

We at the Department of Pre-university Education, Karnataka strive to empower each student to dream big and equip them with the tools that enable them to reach new heights and successfully deal with the challenges of life. As Swami Vivekananda said, "**Real education is that which enables one to stand on one's own legs".**

The course contents in this book are designed with the objective of equipping you well for the next level of study.

We wish you well on your journey and look forward to you becoming a responsible citizen of the nation and give back to the betterment of the society.

With best wishes,

Sd/-
C. Shikha, IAS
Director
Department of Pre University Education
Bengaluru

Preface

The revised syllabus in statistics for the first year P.U course is the outcome of an elaborate and truly democratic approach of teachers from various districts of the state.

Sincere efforts were made to prepare this book in a very lucid and unambiguous manner. A large number of solved problems and self practice problems are provided in each unit so that the students understand the basic concepts of the subject easily. The text also includes some standard problems which are likely to benefit students taking competitive examinations. Our endeavour is to provide a good text book of quality. We hope that this book would serve the purpose of making the study of statistics interesting as well as stimulating. The end of this book provides model question papers for the benefit of the students.

The committee expresses its gratitude to Dr. Ramegowda, I.A.S., Director, Dept of P.U education Karnataka for giving us an opportunity to prepare the text.

The committee is also grateful to Sri. V. K. Nagaraj, Joint director for her valuable suggestions.

The committee is thankful in general to all the members of the syllabus committee of P.U. Board for their valuable support .

The committee also thanks the principals and management of their respective colleges for their co-operation and encouragement given to them in this venture.

The committee is also indebted to Prof. Jayaram Bhat.K, Retired selection grade lecturer in statistics, St Joseph P.U college, Bajpe, Mangalore and Prof.B.C.Patil, Principal H.S.P.U college, Ranebennur, Haveri district for providing their expert views on the text.

In the preparation of this text, we were benefitted a lot from many books and publications. The committee expresses its sincere thanks to the authors and publishers of those books.

We hope that teachers and students will find this book useful.

Text book committee

1. **Sri. Kumaraswamy. N.** — **Chairman**
   Lecturer, National Pre-University College,
   Shivamogga.
2. **Sri. Channamallaiah. A. M.** — **Co-ordinator**
   Lecturer, B.E.S Pre-University College,
   Jayanagara, Bengaluru.
3. **Sri. V. H. Mirji.** — **Member**
   Lecturer, Jyoti Pre-University College,
   Belgaum.
4. **Sri. Devaraya Shanbhogue. H.** — **Member**
   Lecturer, St. Lawrence Pre-University College,
   Moodubelle, Udupi District.
5. **Smt. Kanakavalli M. B.** — **Member**
   Lecturer, NMKRV Pre-Univerity College,
   Jayanagar, Bengaluru
6. **Sri. H. B. Somaraddy.** — **Member**
   Lecturer, S.U.J.M Pre-University College,
   Harapanahalli, Davanagere District.
7. **Sri. S. R. Byakod.** — **Member**
   Lecturer, SECAB Pre-University College for Boys,
   Behind Taj Boudi, Bijapur.
8. **Sri. Sandesh. B.** — **Member**
   Lecturer, Acharya Pre-University College,
   Nandana Education Trust,
   Bye-pass Road, Shivamogga.

I P U C S t a t i s t i c s (31) Objectives

1. The syllabus framed for I PUC is very much useful to the students who join Engineering, B.A (Economics), B.Com, BBM, BCA, CA, ICWA, B.Sc(Mathematics), and post graduation courses in Economics, Sociology, MBA etc.
2. The syllabus is very much useful ti the students who are studying pharmacy, medicines, Bio-technology, nursing courses etc.
3. The syllabus is very much useful for competitive exams like banking, PGCET and other competitive examinations.
4. Also reading of this book enhances and enriches the general knowledge.
5. To create an aptitude for Statistics.
6. The syllabus enables to selection the most appropriate method among available alternative methods to solve a problem.
7. It healps in improving skills of drawing graphs, diagrams and preparing tables.
8. Finally this syllabus helps to make the subject student friendly and socially relevant.

Course Guidelines

1. To teach P.U Statistics, the teaching faculty must have master's degree in Statistics. i.e. M.Sc (Statistics)/M.A (Statistics) / M-Stat.
2. Every section will have 4 hours of theory classes per week and practical class of 2 hours per week per batch.
3. Each batch of practicals must have at least 20 and at the most 40 students.
4. For practical classes and for examinations "Scientific calculators" are allowed to use.
5. Statistical tables will be supplied by the department.

I PUC Statistics Syllabus (120 Hours)

	Pre – requisites A Student must have the knowledge of the following mathematical concepts. a) Laws of indices. b) Common logarithms and its applications. c) Solving simultaneous equations. d) Set theory. e) Permutations and combinations.	(8Hours)
I.	Introduction to Statistics and some basic concepts.	(7 Hours)
II.	Organization of data.	(7 Hours)
III.	Classification and Tabulation of data.	(10 Hours)
IV.	Diagrammatic and Graphical representation of data	(10 Hours)
V.	Analysis of univariate data. a) Measures of central tendency. b) Measures of position. c) Measures of dispersion. d) Moments, Skewness and Kurtosis.	(25 Hours)
VI.	Analysis of bivariate data. Correlation and Regression.	(15 Hours)
VII.	Association of attributes. Yule's coefficient of association	(5 Hours)
VIII.	Interpolation and Extrapolation (Only binomial method with one missing value)	(5 Hours)
IX.	Theory of Probability.	(14 Hours)
X.	Random Variable and Mathematical expectation of a discrete random variable	(14 Hours)

List of Practicals (40 Hours)

1.	Formation of Univariate and Bivariate frequency distributions	(4 Hours)
2.	Preparation of blank tables and tables with information.	(4 Hours)
3.	Diagrammatic representation of data.	(4 Hours)
4.	Graphical representation of frequency distribution.	(4 Hours)
5.	Measures of central tendency and positions.	(4 Hours)
6.	Measures of dispersion,	(4 Hours)
7.	Measures of Skewness.	(2 Hours)
8.	Correlation.	(4 Hours)
9.	Regression.	(2 Hours)
10.	Association of attributes, Interpolation and Extrapolation.	(2 Hours)
11.	Probability applications	(2 Hours)
12.	Mathematical Expectation.	(2 Hours)
13.	Covariance and Correlation of Random Variables.	(2 Hours)

Blow up syllabus

I. Introduction to Statistics and some basic concepts:

Meaning – Origin – Scope of Statistics. Definitions – in singular and plural sense. Characteristics, Branches, Functions and limitations of Statistics. Statistical applications in other subjects. Distrust of Statistics – causes and remedies. Some basic concepts – units, population, sample, qualitative characteristic, quantitative characteristic, attribute, variables (discrete and continuous), nominal scale and ordinal scale.

ಸಂಖ್ಯಾಶಾಸ್ತ್ರದ ಪರಿಚಯ ಮತ್ತು ಮೂಲ ಸಂಗತಿಗಳು : ಸಂಖ್ಯಾಶಾಸ್ತ್ರದ ಮೂಲ, ವ್ಯಾಪ್ತಿ, ವ್ಯಾಖ್ಯೆಗಳು- ಏಕ ವಚನ, ಬಹುವಚನ ರೂಪದಲ್ಲಿ. ಗುಣಲಕ್ಷಣಗಳು, ಶಾಖೆಗಳು, ಕಾರ್ಯಗಳು ಮತ್ತು ಇತಿಮಿತಿಗಳು. ಉಳಿದ ವಿಷಯಗಳಲ್ಲಿ ಸಂಖ್ಯಾಶಾಸ್ತ್ರದ ಪ್ರಯೋಜನ, ಸಂಖ್ಯಾಶಾಸ್ತ್ರದಲ್ಲಿ ಅಪನಂಬಿಕೆ - ಕಾರಣಗಳು ಮತ್ತು ಪರಿಹಾರ ಸೂತ್ರಗಳು. ಮೂಲ ಸಂಗತಿಗಳು - ಘಟಕಗಳು, ಸಮಷ್ಟಿ, ನಿದರ್ಶಕ (ಪ್ರತಿಚಯ), ಗುಣಾತ್ಮಕ ಲಕ್ಷಣ, ಪರಿಮಾಣಾತ್ಮಕ ಲಕ್ಷಣ, ಗುಣಧರ್ಮ, ಚಲಕ (ವಿಚ್ಛಿನ್ನ ಮತ್ತು ಸತತ), ಶಾಬ್ದಿಕ ಅಳತೆ ಮತ್ತು ಕ್ರಮಸೂಚಕ ಅಳತೆ.

II. Organization of data:

Statistical enquiry and its stages. Primary and Secondary data. Methods of collection of primary data, with merits and demerits. Essentials of a good questionnaire. Questionnaire and schedule with respect to their relative merits and demerits. Sources of secondary data. Census Enumeration and Sample Survey with respect to their relative merits and demerits. Pilot survey. Sampling – Methods of sampling - simple random sampling, stratified sampling and systematic sampling. Errors of sampling.

ದತ್ತಾಂಶ(ನ್ಯಾಸ)ದ ವ್ಯವಸ್ಥೆ : ಸಾಂಖ್ಯಿಕ ಪರಿಶೀಲನೆ ಮತ್ತು ಇದರ ಹಂತಗಳು. ಪ್ರಾಥಮಿಕ ಮತ್ತು ದ್ವಿತೀಯ ದತ್ತಾಂಶ- ಪ್ರಾಥಮಿಕ ದತ್ತಾಂಶದ ಸಂಗ್ರಹಣಾ ವಿಧಾನಗಳು (ಗುಣದೋಷಗಳೊಂದಿಗೆ). ಉತ್ತಮ ಪ್ರಶ್ನಾವಳಿಯ ಅವಶ್ಯಕತೆಗಳು, ಪ್ರಶ್ನಾವಳಿ ಮತ್ತು ತಪಶೀಲು ಪಟ್ಟಿಗಳ ಗುಣದೋಷಗಳು. ದ್ವಿತೀಯ ದತ್ತಾಂಶದ ಮೂಲಗಳು. ಗಣತಿ ಮತ್ತು ನಿದರ್ಶಕ ಸಮೀಕ್ಷೆ ಅವುಗಳ ಗುಣದೋಷಗಳು. ಮಾರ್ಗದರ್ಶಕ ಸಮೀಕ್ಷೆ, ಪ್ರತಿಚಯನಗಳು - ಸರಳ ಅಕಸ್ಮಿಕ, ಸ್ತರೀಕೃತ ಮತ್ತು ವ್ಯವಸ್ಥಿತ ಪ್ರತಿಚಯನಗಳು. ಪ್ರತಿಚಯನಗಳಲ್ಲಿನ ದೋಷಗಳು.

III. Classification and Tabulation:

Classification- Introduction, Meaning and objectives of classification, Types - (i) Chronological (Temporal) (ii) Geographical (Spatial)

(iii) Qualitative (Simple and manifold) (iv) Quantitative. Explanation with examples.

Formation of discrete and continuous frequency distributions: Explanation of range, class, class limits, class intervals: inclusive, exclusive, open-end classes, correction factor, conversion of inclusive classes to exclusive classes, class mid-point, width of the class, formation of less than and more than cumulative frequency distributions, frequency density, relative frequency. Rules of classification. Formation of uni-variate and bi-variate frequency distributions.

Tabulation: Meaning, difference between classification and tabulation. Parts of a table-Format of a table and brief explanation of parts of table, Rules of tabulation, Types of tables-(i) simple and complex (many-fold) tables (ii) General and special purpose tables; preparation of blank tables, tables with numerical information.

ವರ್ಗೀಕರಣ ಮತ್ತು ಸಾರಣೀಯನೆ(ಕೋಷ್ಟಕೀಕರಣ): ವರ್ಗೀಕರಣದ ಪರಿಚಯ, ಅರ್ಥ ಮತ್ತು ಉದ್ದೇಶಗಳು, ಪ್ರಕಾರಗಳು (i) ಕಾಲಾನುಕ್ರಮದ (ii) ಭೌಗೋಳಿಕ (iii) ಗುಣಾತ್ಮಕ ಹಾಗು (iv) ಪರಿಮಾಣಾತ್ಮಕ ವರ್ಗೀಕರಣಗಳ ವಿವರಣೆ (ಉದಾಹರಣೆಯೊಂದಿಗೆ).

ವಿಚ್ಛಿನ್ನ ಮತ್ತು ಸತತ ಆವೃತ್ತಿ ವಿತರಣಾ ಪಟ್ಟಿಯ ತಯಾರಿಕೆ - ವಿಸ್ತಾರ, ವರ್ಗ, ವರ್ಗಮಿತಿಗಳು, ಸಂವೃತ ಮತ್ತು ವಿಮುಕ್ತ ಹಾಗು ತೆರದ (ವಿವೃತ್ತಾಂತ) ವರ್ಗಾಂತರಗಳು, ಸಂವೃತ ವರ್ಗಾಂತರಗಳನ್ನು ವಿಮುಕ್ತ ವರ್ಗಾಂತರಗಳಾಗಿ ಪರಿವರ್ತಿಸುವುದು. ವರ್ಗದ ಮಧ್ಯಬಿಂದು, ವರ್ಗದ ಗಾತ್ರ, ವರ್ಗ ಆವೃತ್ತಿ, ಮೇಲ್ಮಿತಿಗಿಂತ ಕಡಿಮೆ ಮತ್ತು ಕೆಳಮಿತಿಗಿಂತ ಹೆಚ್ಚಿನ ಸಂಚಿತ ಆವೃತ್ತಿ ವಿತರಣೆಗಳ ರಚನೆ, ಆವೃತ್ತಿ ಸಾಂದ್ರತೆ, ಸಾಪೇಕ್ಷ ಆವೃತ್ತಿಗಳ ವ್ಯಾಖ್ಯೆ (ಉದಾಹರಣೆಗಳೊಂದಿಗೆ), ಆವೃತ್ತಿ ವಿತರಣಾ ಪಟ್ಟಿ ತಯಾರಿಕೆಯ ನಿಯಮಗಳು. ಏಕಚಲಕ ಮತ್ತು ದ್ವಿಚಲಕದ ಆವೃತ್ತಿ ವಿತರಣಾ ಪಟ್ಟಿಯ ತಯಾರಿಕೆಯ ಲೆಕ್ಕಗಳು.

ಸಾರಣೀಯನೆ : ಅರ್ಥ, ವರ್ಗೀಕರಣ ಮತ್ತು ಸಾರಣೀಯನೆಗಳ ವ್ಯತ್ಯಾಸಗಳು. ಸಾರಣಿಯ ಭಾಗಗಳು ಮತ್ತು ಅವುಗಳ ಸಂಕ್ಷಿಪ್ತ ವಿವರಣೆ ಸಾರಣೀಯನೆಯ ನಿಯಮಗಳು, ಸಾರಣಿಗಳ ಪ್ರಕಾರಗಳು (ವಿಧಗಳು) - (i) ಸರಳ ಮತ್ತು ಸಂಕೀರ್ಣ (ಬಹು-ಮುಖ) ಸಾರಣಿಗಳು (ii) ಸಾಮಾನ್ಯ ಮತ್ತು ವಿಶೇಷ ಉದ್ದೇಶದ ಸಾರಣಿಗಳು, ಖಾಲಿ ಸಾರಣಿ ಮತ್ತು ಸಾಂಖ್ಯಿಕ ಮಾಹಿತಿ ಒಳಗೊಂಡ ಸಾರಣಿಗಳ ರಚನೆ.

IV. Diagrammatic and Graphical representation of data:

Diagrams- Meaning, needs of diagrams, General rules of construction of diagrams – simple, multiple, component, percentage bars – problems and constructions.

Component (sub divided) Pie diagram.

Graphs- Explanation, types – Histogram (Equal and Unequal width of class intervals), examples for obtaining mode from Histogram. Frequency Polygon and frequency curve – meaning, method of construction with and without histogram. Ogives- Meaning, method of construction of both less than and more than Ogives. Examples. Obtaining the values of Median and Quartiles from less than Ogive. Comparison of tables and diagrams, difference between diagrams and graphs.

ದತ್ತಾಂಶದ ಚಿತ್ರ ಹಾಗೂ ಆಲೇಖ ನಿರೂಪಣೆ :

ಚಿತ್ರಗಳು - ಅರ್ಥ, ಉಪಯೋಗಗಳು, ಹೋಲಿಕೆ, ಚಿತ್ರ ಮತ್ತು ಆಲೇಖಗಳ ವ್ಯತ್ಯಾಸ. ಚಿತ್ರ ತಯಾರಿಸುವ ಸಾಮಾನ್ಯ ನಿಯಮಗಳು. ಸರಳ, ಬಹು, ವಿಭಜಿತ ಮತ್ತು ಶೇಖಡಾ ಸ್ತಂಭ ನಕ್ಷೆಗಳ ರಚನಾಕ್ರಮ ಮತ್ತು ಲಕ್ಷಣಗಳು. ಪೈ ನಕ್ಷೆಯ ರಚನಾ ಕ್ರಮ ಮತ್ತು ಲಕ್ಷಣಗಳು.

ಆಲೇಖಗಳು - ವಿವರಣೆ, ವಿಧಗಳು, ಆವೃತ್ತ್ಯಾಕೃತಿ ರಚನೆ (ಸಮ ಮತ್ತು ವಿಭಿನ್ನ ಗಾತ್ರವುಳ್ಳ ವರ್ಗಾಂತರಗಳಿಗೆ), ಆಲೇಖದ ಸಹಾಯದಿಂದ ಬಹುಲಕ ಕಂಡುಹಿಡಿಯುವ ಲಕ್ಷಣಗಳು, ಆವೃತ್ತಿ ಬಹುಭುಜಾಕೃತಿ ಮತ್ತು ಆವೃತ್ತಿ ವಕ್ರಗಳ ಅರ್ಥ, ರಚನಾ ಕ್ರಮ (ಆವೃತ್ತ್ಯಾಕೃತಿ ಮೂಲಕ ಅಥವಾ ನೇರವಾಗಿ). ಸಂಚಿತ ಆವೃತ್ತಿ ವಕ್ರಗಳು - ಮೇಲ್ಮಿತಿಗಿಂತ ಕಡಿಮೆ ಸಂಚಿತ ಆವೃತ್ತಿ ವಕ್ರ ಮತ್ತು ಕೆಳಮಿತಿಗಿಂತ ಹೆಚ್ಚಿನ ಸಂಚಿತ ವಕ್ರಗಳನ್ನು ಎಕೆಯುವ ಕ್ರಮ ಮತ್ತು ಲಕ್ಷಣಗಳು. ಮಧ್ಯಾಂಕ ಮತ್ತು ಚತುರ್ಥಾಂಕಗಳನ್ನು ಮೇಲ್ಮಿತಿಗಿಂತ ಕಡಿಮೆ ಸಂಚಿತ ಆವೃತ್ತಿ ವಕ್ರದಿಂದ ಪಡೆಯುವ ಲಕ್ಷಣಗಳು. ಸಾರಣಿ ಮತ್ತು ಚಿತ್ರಗಳ ಹೋಲಿಕೆ. ಚಿತ್ರ ಮತ್ತು ಆಲೇಖಗಳ ವ್ಯತ್ಯಾಸ.

V. Analysis of univariate data:

a) Measures of central tendency (Averages)- Meaning, objectives, definition of an average, definition of central tendency, essentials of a good average. Types of averages,

i) Arithmetic mean- Definition, problems for both ungrouped and grouped data, one missing frequency problems, properties of mean with numerical applications, corrected mean in case of one wrong observation, merits and demerits.

ii) Median- Definition, problems for ungrouped and grouped data, one missing frequency problems, graphical solution of median, merits and demerits.

iii) Mode- Definition, problems for ungrouped and grouped data

(excluding grouping method). Graphical solution of mode, merits and demerits, Empirical relationship among mean, median and mode.

iv) Geometric mean- Definition, problems for ungrouped and grouped data, problems on growth rates and interest rates, merits and demerits.

v) Harmonic mean- Definition, problems for ungrouped and grouped data, problems on average speed, average number of days required for the completion of a given work, merits and demerits. Relationship between AM, GM and HM.

b) Measures of position- Meaning, definitions of quartiles, deciles and percentiles. Problems on ungrouped and grouped data.

c) Measures of dispersion- Meaning, objectives, definition, essentials of a good measure of dispersion. Absolute and relative measures of dispersion.

Types – absolute and relative measures.

i) Range- Definition, absolute and relative measures formulae, problems on ungrouped data.

ii) Quartile deviation- Definition, absolute and relative measures formulae, problems on ungrouped and grouped data.

iii) Mean deviation- Definition, absolute and relative measures formulae, problems on ungrouped and grouped data based on mean, median and mode.

iv) Standard deviation- Definition, problems on ungrouped and grouped data, properties of standard deviation with numerical applications and variance. Coefficient of variation for ungrouped and grouped data.

d) Moments, Skewness and kurtosis-

i) Moments- Meaning, definition of central moments, description of first four central moments. Formulae for β_1, γ_1, β_2 and γ_2

ii) Skewness- Definition, types with diagram, measures of skewness – Karl Pearson's and Bowley's coefficient of skewness for grouped data. Skewness based on moments (for given values of μ_2 and μ_3 finding β_1).

iii) Kurtosis- Definition, explanation of kurtosis with neat diagram, measure of kurtosis based on moments (given the values of μ_4 and μ_2 finding β_2).

ಏಕಚಲಕದ ವಿಶ್ಲೇಷಣೆ:

ಎ) ಕೇಂದ್ರಿಯ ಮೌಲ್ಯಮಾಪನ ಅಳತೆಗಳು (ಸರಾಸರಿಗಳು) - ಅರ್ಥ, ಉದ್ದೇಶಗಳು, ಸರಾಸರಿಯ ವ್ಯಾಖ್ಯೆ, ಕೇಂದ್ರಿಯ ಮೌಲ್ಯಮಾಪನದ ವ್ಯಾಖ್ಯೆ, ಉತ್ತಮ ಸರಾಸರಿಗೆ ಅವಶ್ಯವಿರುವ ಅಂಶಗಳು. ಸರಾಸರಿಯ ಪ್ರಕಾರಗಳು (ವಿಧಗಳು)

(i) ಅಂಕಗಣಿತ ಸರಾಸರಿ : ವ್ಯಾಖ್ಯೆ, ಅವರ್ಗೀಕೃತ ಮತ್ತು ವರ್ಗೀಕೃತ ದತ್ತಾಂಶದ ಲೆಕ್ಕಗಳು. ಕಾಣೆಯಾದ (ಜಿಟ್ಟುಹೋದ) ಅವೃತ್ತಿಯ ಲೆಕ್ಕಗಳು, ಸರಾಸರಿಯ ಗುಣಧರ್ಮಗಳು, ಸಾಂಖ್ಯಿಕ ಅನ್ವಯಗಳು(ಲೆಕ್ಕಗಳು), ತಪ್ಪು ಪ್ರಾಪ್ತಾಂಕದಿಂದ ಸರಿಪಡಿಸಿದ ಸರಾಸರಿ, ಗುಣ ಮತ್ತು ದೋಷಗಳು.

(ii) ಮಧ್ಯಾಂಕ : ವ್ಯಾಖ್ಯೆ, ಅವರ್ಗೀಕೃತ ಮತ್ತು ವರ್ಗೀಕೃತ ದತ್ತಾಂಶದ ಲೆಕ್ಕಗಳು, ಕಾಣೆಯಾದ ಅವೃತ್ತಿಯ ಲೆಕ್ಕಗಳು, ಗುಣ ಮತ್ತು ದೋಷಗಳು. ಮಧ್ಯಾಂಕವನ್ನು ಆಲೇಖದಿಂದ ಪಡೆಯುವ ವಿಧಾನ.

(iii) ಬಹುಲಕ : ವ್ಯಾಖ್ಯೆ, ಅವರ್ಗೀಕೃತ ಮತ್ತು ವರ್ಗೀಕೃತ ದತ್ತಾಂಶದ ಲೆಕ್ಕಗಳು(ಗುಂಪುಗೊಡಿಸುವ ವಿಧಾನ ಹೊರತುಪಡಿಸಿ), ಬಹುಲಕವನ್ನು ಆಲೇಖದಿಂದ ಪಡೆಯುವ ವಿಧಾನ, ಗುಣ ಮತ್ತು ದೋಷಗಳು. ಸರಾಸರಿ, ಮಧ್ಯಾಂಕ ಹಾಗೂ ಬಹುಲಕಗಳ ನಡುವಿನ ಅನುಭವ ಜನ್ಯ ಸಂಬಂಧ ಮತ್ತು ಲೆಕ್ಕಗಳು.

(iv) ಗುಣೋತ್ತರ ಸರಾಸರಿ : ವ್ಯಾಖ್ಯೆ, ಅವರ್ಗೀಕೃತ ಮತ್ತು ವರ್ಗೀಕೃತ ದತ್ತಾಂಶದ ಲೆಕ್ಕಗಳು, ಬೆಳವಣಿಗೆಯ ದರ ಮತ್ತು ಬಡ್ಡಿದರಗಳಿಗೆ ಸಂಬಂಧಿಸಿದ ಲೆಕ್ಕಗಳು, ಗುಣ ಮತ್ತು ದೋಷಗಳು.

(v) ಹರಾತ್ಮಕ ಸರಾಸರಿ : ವ್ಯಾಖ್ಯೆ, ಅವರ್ಗೀಕೃತ ಮತ್ತು ವರ್ಗೀಕೃತ ದತ್ತಾಂಶದ ಲೆಕ್ಕಗಳು, ಸರಾಸರಿ ವೇಗ, ನೀಡಿದ ಕೆಲಸ ಮುಗಿಸಲು ಅವಶ್ಯವಿರುವ ಸರಾಸರಿ ದಿನಗಳ ಸಂಖ್ಯೆಯ ಲೆಕ್ಕಗಳು, ಗುಣ ಮತ್ತು ದೋಷಗಳು ಅಂಕಗಣಿತ ಸರಾಸರಿ (AM), ಗುಣೋತ್ತರ ಸರಾಸರಿ (GM) ಮತ್ತು ಹರಾತ್ಮಕ ಸರಾಸರಿ (HM) ಗಳಿಗೆ ಇರುವ ಸಂಬಂಧ ಮತ್ತು ಲೆಕ್ಕಗಳು.

ಐ) ಸ್ಥಾನೀಯ ಬೆಲೆಗಳು : ಅರ್ಥ, ಚತುರ್ಥಾಂಕ ದತಾಂಶ, ಶತಾಂಕಗಳು ವ್ಯಾಖ್ಯೆ, ಅವರ್ಗೀಕೃತ ಮತ್ತು ವರ್ಗೀಕೃತ ದತ್ತಾಂಶದ ಲೆಕ್ಕಗಳು.

ಸಿ) ಹರವಿನ ಮಾಪನ : ಅರ್ಥ, ಉದ್ದೇಶಗಳು, ವ್ಯಾಖ್ಯೆ, ಉತ್ತಮ ಹರವಿನ ಮಾಪನಕ್ಕೆ ಅವಶ್ಯವಿರುವ ಅಂಶಗಳು. ನಿರಪೇಕ್ಷ ಮತ್ತು ಸಾಕ್ಷೇಪ ಅಳತೆಗಳು. ಹರವಿನ ಮಾಪನದ ಪ್ರಕಾರಗಳು-ನಿರಪೇಕ್ಷ ಮತ್ತು ಸಾಕ್ಷೇಪ ಅಳತೆಗಳು.

(i) ವ್ಯಾಪ್ತಿ : ವ್ಯಾಖ್ಯೆ, ನಿರಪೇಕ್ಷ ಮತ್ತು ಸಾಪೇಕ್ಷ ಅಳತೆಯ ಸೂತ್ರಗಳು, ಅವರ್ಗೀಕೃತ ದತ್ತಾಂಶಕ್ಕೆ ಲೆಕ್ಕಗಳು.

(ii) ಚತುರ್ಥಕ ವಿಚಲನೆ : ವ್ಯಾಖ್ಯೆ, ನಿರಪೇಕ್ಷ ಮತ್ತು ಸಾಪೇಕ್ಷ ಅಚತೆಯ ಸೂತ್ರಗಳು, ಅವರ್ಗೀಕೃತ ದತ್ತಾಂಶಕ್ಕೆ ಲೆಕ್ಕಗಳು.

(iii) ಸರಾಸರಿ ವಿಚಲನೆ : ವ್ಯಾಖ್ಯೆ, ನಿರಪೇಕ್ಷ ಮತ್ತು ಸಾಪೇಕ್ಷ ಅಚತೆಯ ಸೂತ್ರಗಳು. ಸರಾಸರಿ, ಮಧ್ಯಾಂಕ ಮತ್ತು ಬಹುಲಕ ಅಧರಿಸಿದ ಅವರ್ಗೀಕೃತ ಮತ್ತು ವರ್ಗೀಕೃತ ದತ್ತಾಂಶದ ಲೆಕ್ಕಗಳು.

(iv) ನಿಯತ ವಿಚಲನೆ : ವ್ಯಾಖ್ಯೆ, ಅವರ್ಗೀಕೃತ ಮತ್ತು ವರ್ಗೀಕೃತ ದತ್ತಾಂಶದ ಲೆಕ್ಕಗಳು. ನಿಯತ ವಿಚಲನೆಯ ಗುಣಧರ್ಮಗಳು (ಲೆಕ್ಕಗಳೊಂದಿಗೆ) ಮತ್ತು ವಿಚಲನೆ. ಅವರ್ಗೀಕೃತ ಮತ್ತು ವರ್ಗೀಕೃತ ದತ್ತಾಂಶಕ್ಕೆ ವಿಚಲನೆಯ ಸಹಗುಣಕದ ಲೆಕ್ಕಗಳು.

ಞ) ಭ್ರಾಮಕ (ಘೂರ್ಣ)ಗಳು, ವಿಷಮತೆ ಮತ್ತು ಶಿಖರತೆ:

(i) ಭ್ರಾಮಕಗಳು : ಅರ್ಥ, ಕೇಂದ್ರಿಯ ಭ್ರಾಮಕಗಳ ವ್ಯಾಖ್ಯೆ. ಮೊದಲ ನಾಲ್ಕು ಕೇಂದ್ರಿಯ ಭ್ರಾಮಕಗಳ ವಿವರಣೆ. β_1 γ_1, β_2 ಮತ್ತು γ_2 ಗಳ ಸೂತ್ರ.

(ii) ವಿಷಮತೆ : ವ್ಯಾಖ್ಯೆ, ಚಿತ್ರದೊಂದಿಗೆ ಪ್ರಕಾರಗಳು, ವಿಷಮತೆಯ ಅಳತೆ- ಕಾರ್ಲಪಿಯಸನ್ನನ ಮತ್ತು ಬೌಲಿಯ ವಿಷಮತೆಯ ಸಹಗುಣಕಗಳ ಸೂತ್ರಗಳು ಮತ್ತು ವರ್ಗೀಕೃತ ದತ್ತಾಂಶಕ್ಕೆ ಲೆಕ್ಕಗಳು. ಭ್ರಾಮಕಗಳನ್ನು ಅಧರಿಸಿದ ವಿಷಮತೆ (μ_2 ಮತ್ತು μ_3 ಬೆಲೆ ನೀಡಿದಾಗ β_2 ಕಂಡುಹಿಡಿಯುವುದು).

(iii) ಶಿಖರತೆ : ವ್ಯಾಖ್ಯೆ , ಚಿತ್ರದೊಂದಿಗೆ ಶಿಖರತೆ ವಿವರಣೆ, ಭ್ರಾಮಕಗಳನ್ನು ಅಧರಿಸಿದ ಶಿಖರತೆಯ ಅಳತೆ (μ_4 ಮತ್ತು μ_3 ಬೆಲೆ ನೀಡಿದಾಗ β_2 ಕಂಡುಹಿಡಿಯುವುದು).

VI. Analysis of Bivariate data:

a) Correlation- Definition, Types – Simple, multiple, partial. Causation – Spurious, positive, negative, perfect and no correlation explanation with examples. Significance of study of correlation analysis.

Measurement of correlation- scatter diagram explanation with charts, merits and demerits. Problems regarding construction of scatter diagram.

Karl Pearson's coefficient of correlation – definition, formulae for ungrouped and grouped data. Properties of coefficient of correlation, interpretation. Problems – ungrouped and grouped data.

Spearman's coefficient of rank correlation – without ties and tie (one or two), interpretation. Problems – with ties (One or two

repeated ranks) and without ties.

b) Regression- Definition, regression lines/equations of x on y and y on x. Properties of regression coefficient and regression lines/ equations. Problems on ungrouped and grouped data, uses of regression analysis. Comparison between correlation and regression.

ದ್ವಿಚಲಕ ದತ್ತಾಂಶದ ವಿಶ್ಲೇಷಣೆ :

ಎ) ಸಹಸಂಬಂಧ : ವ್ಯಾಖ್ಯೆ , ಪ್ರಕಾರಗಳು- ಸರಳ, ಬಹು, ಭಾಗಶಃ (ಅಪೂರ್ಣ). ಕಾರಣ ವಿಧಾನ – ಕೃತಕ, ಧನಾತ್ಮಕ, ಋಣಾತ್ಮಕ, ಸಂಪೂರ್ಣ ಮತ್ತು ಸಹಸಂಬಂಧವಿಲ್ಲದ ಉದಾಹರಣೆಗಳೊಂದಿಗಿನ ವಿವರಣೆಯ ಸಹಸಂಬಂಧ ವಿಶ್ಲೇಷಣೆಯ ಅಭ್ಯಾಸದ ಅರ್ಥಗರ್ಭತತೆ (ಗಮನಾರ್ಹತೆ).

ಸಹಸಂಬಂಧದ ಅಳತೆ : ಚದುರಿಕೆಯ ಚಿತ್ರ ಮತ್ತು ವಿವರಣೆ, ಗುಣ ಮತ್ತು ದೋಷಗಳು, ಚದುರಿಕೆಯ ಚಿತ್ರ ತಯಾರಿಸುವ ಲೆಕ್ಕಗಳು.

ಕಾರ್ಲಪಿಯರ್‌ಸನ್ನನ ಸಹಸಂಬಂಧದ ಗುಣಕ : ವ್ಯಾಖ್ಯೆ, ಅವರ್ಗೀಕೃತ ಮತ್ತು ವರ್ಗೀಕೃತ ದತ್ತಾಂಶಗಳಿಗೆ ಸೂತ್ರಗಳು, ಸಹಸಂಬಂಧ ಗುಣಕದ ಗುಣಧರ್ಮಗಳು, ಗುಣ ಮತ್ತು ದೋಷಗಳು, ಅವರ್ಗೀಕೃತ ಮತ್ತು ವರ್ಗೀಕೃತ ದತ್ತಾಂಶಕ್ಕೆ ಸಂಬಂಧಿಸಿದ ಲೆಕ್ಕಗಳು.

ಸ್ಪಿಯರ್ ಮನ್ನನ ದರ್ಜಾ ಸಹ ಸಂಬಂಧ: ದರ್ಜೆಗಳು ಪುನರಾವರ್ತನೆ ಆಗದಿದ್ದಾಗ ಮತ್ತು ಆದಾಗ (ಒಂದು ಅಥವಾ ಎರಡು), ಗುಣ ಮತ್ತು ದೋಷಗಳು, ದರ್ಜೆಗಳು ಪುನರಾವರ್ತನೆಯಾದಾಗ ಆಗದಿದ್ದಾಗ ಮತ್ತು ಆಗದಿರುವಾಗಿನ (ಒಂದು ಅಥವಾ ಎರಡು) ಲೆಕ್ಕಗಳು.

ಬಿ) ಹಿಂಚಲನೆ (ಸಮಾಶ್ರಯಣ): ವ್ಯಾಖ್ಯೆ, x ನ ಹಿಂಚಲನೆ (ಸಮಾಶ್ರಯಣ) y ಮೇಲಿನ ಹಾಗೂ y ನ ಹಿಂಚಲನೆ x ಮೇಲಿನ ಸಮೀಕರಣೆಗಳು, ಹಿಂಚಲನಾ ಸಹಗುಣಕಗಳು ಮತ್ತು ಹಿಂಚಲನಾ ರೇಖೆಯ ಗುಣಧರ್ಮಗಳು. ಅವರ್ಗೀಕೃತ ಮತ್ತು ವರ್ಗೀಕೃತ ದತ್ತಾಂಶದ ಲೆಕ್ಕಗಳು, ಹಿಂಚಲನಾ ವಿಶ್ಲೇಷಣೆಯ ಉಪಯೋಗಗಳು, ಸಹಸಂಬಂಧ ಮತ್ತು ಹಿಂಚಲನೆಗಳ ಹೋಲಿಕೆ.

VII. Association of attributes:

Introduction, definition, notations- meaning and methods of association. Yule's coefficient of association and its applications.

ಗುಣಧರ್ಮಗಳ ಸಹಚರ್ಯ : ಪರಿಚಯ, ವ್ಯಾಖ್ಯೆ, ಸಂಕೇತಗಳ ಅರ್ಥ, ಯುಲೇನ ಸಹಚರ್ಯಗುಣಕದ ಸೂತ್ರ ಮತ್ತು ಲೆಕ್ಕಗಳು.

VIII. Interpolation and Extrapolation:

Meaning and utilities of interpolation and extrapolation. Binomial expansion method of interpolation (with one missing value) and Extrapolation for next successive value. Merits and demerits, applications.

ಅಂತರ್ವೇಶನ ಮತ್ತು ಬಹಿರ್ವೇಶನ : ಅಂತರ್ವೇಶನ ಮತ್ತು ಬಹಿರ್ವೇಶನಗಳ ಅರ್ಥ ಮತ್ತು ಉಪಯೋಗಗಳು. ದ್ವಿಪದ ವಿಸ್ತರಣೆ ಉಪಯೋಗಿಸಿಕೊಂಡು ಅಂತರ್ವೇಶನ (ಒಂದು ಬೆಲೆ ಕಾಣೆಯಾದಾಗ), ಬಹಿರ್ವೇಶನಗಳಿಗೆ (ಮುಂದಿನ ಒಂದು ಬೆಲೆಗೆ) ಸಂಬಂಧಿಸಿದ ಲೆಕ್ಕಗಳು.

IX. Theory of probability:

Introduction to probability, Definition of Experiment, Outcomes, Deterministic experiment, Random experiment, Sample space, Null event, Simple event, Compound event, Sure event with examples. Meaning of Favourable and Exhaustive outcomes, Equally likely events, Union and intersection of events,Mutually exclusive events, Complement of an event with examples. Meaning of Classical and Empirical/statistical methods of assigning probabilities. Classical/ Mathematical, Empirical/Statistical and axiomatic definitions of probability. Statement and proofs (on basis of Mathematical definition) of $P(\Phi)=0$ $P(S) = 1$, $0 \leq P(A) \leq 1$, $P(A) + P(A^1) = 1$, Statement and proofs (on basis of Mathematical definition) of addition theorem of probability- for two non-mutually and mutually exclusive events. Definition and examples of independent, dependent events and conditional probability. Statement and proofs (on basis of Mathematical definition) of multiplication theorem of probability- for dependent and independent events and applications.

ಸಂಭವನೀಯತೆ : ಸಂಭವನೀಯತೆ ಪರಿಚಯ, ಪ್ರಯೋಗ, ಪರಿಣಾಮಗಳು, ನಿಶ್ಚಿತ ಪ್ರಯೋಗ, ಆಕಸ್ಮಿಕ (ಯಾದೃಚ್ಛಿಕ) ಪ್ರಯೋಗ, ನಿದರ್ಶಕ ವ್ಯೂಮ (ಪ್ರತಿಚಯ ಆಕಾಶ), ಶೂನ್ಯ ಘಟನೆ, ಸರಳ ಘಟನೆ, ಸಂಯುಕ್ತ ಘಟನೆ ಮತ್ತು ನಿಶ್ಚಿತ ಘಟನೆಗಳ ಉದಾಹರಣೆಗಳೊಂದಿಗೆ ವ್ಯಾಖ್ಯೆ, ಒಲವುಳ್ಳ ಮತ್ತು ಶಕ್ಯ ಪರಿಣಾಮಗಳು, ಸಮ ಸಂಭವ ಘಟನೆಗಳು, ಘಟನೆಗಳ ಸಂಯೋಗ ಮತ್ತು ಛೇದನ, ಪರಸ್ಪರ ವಿಮುಕ್ತ ಘಟನೆಗಳು ಮತ್ತು ಪೂರಕ ಘಟನೆಗಳ ಉದಾಹರಣಗಳೊಂದಿಗೆ ಅರ್ಥ, ಶಾಸ್ತ್ರಿಯ ಮತ್ತು ಅನುಭವಜನ್ಯ ಸಂಭವನೀಯತೆಯ ಹಂಚಿಕೆಯ ವಿಧಾನಗಳ ಅರ್ಥ. ಶಾಸ್ತ್ರೀಯ, ಅನುಭವಜನ್ಯ ಮತ್ತು ಸ್ವಪ್ರಮಾಣಿಕೃತ ಸಂಭವನೀಯತೆಯ ವ್ಯಾಖ್ಯೆ. $P(\Phi)=0$, $\leq P(A) \leq 1$, $P(A) + P(A^1)=1$ ಗಳ ಹೇಳಿಕೆ ಮತ್ತು ಸಾಧನೆ. ಪರಸ್ಪರ ವಿಮುಕ್ತ ಮತ್ತು ವಿಮುಕ್ತವಲ್ಲದ ಘಟನೆಗಳ ಸಂಕಲನ ಸಂಭವತೆಯ ಪ್ರಮೇಯಗಳ ಹೇಳಿಕೆ ಮತ್ತು ಸಾಧನೆ (ಗಣಿತ ರೂಪಕ ವ್ಯಾಖ್ಯೆ ಅಧಲಿಸಿ). ಸ್ವತಂತ್ರ ಮತ್ತು ಅವಲಂಭಿತ ಘಟನೆಗಳ ವ್ಯಾಖ್ಯೆ ಮತ್ತು ಉದಾಹರಣೆಗಳು. ಅಧೀನ ಸಂಭವತೆಯ ವ್ಯಾಖ್ಯೆ ಅವಲಂಭಿತ ಮತ್ತು ಸ್ವತಂತ್ರ ಘಟನೆಗಳ ಗುಣಾಕಾರ ಸಂಭವನೀಯತೆಯ ಪ್ರಮೇಯಗಳ ಹೇಳಿಕೆ ಮತ್ತು ಸಾಧನೆ (ಗಣಿತ ವ್ಯಾಖ್ಯೆ ಆಧರಿಸಿ), ಇವುಗಳಿಗೆ ಸಂಬಂಧಿಸಿದ ಲೆಕ್ಕಗಳು.

X. Random Variable and Mathematical expectation of a discrete random variable:

Definition with examples of discrete and continuous random variables. Definition of probability mass function and probability density function. Bivariate and marginal probability distributions definitions with examples.

Definition of Expected value/mean, variance and standard deviation of a discrete random variable. Related functions defined on a discrete random variable, applications (solving numerical and verbal problems), including the case of missing probabilities. Expectation and variance of following functions with proofs - a, aX, aX±b where 'a' and 'b' are any two constants and related examples.

Statement and proofs of addition and multiplication theorem of Expectation. Covariance and correlation coefficient of bivariate random variables.

ಅಕಸ್ಮಿಕ ಚಲಕ ವಿಚ್ಛಿನ್ನ ಅಕಸ್ಮಿಕ ಚಲಕದ ಗಣಿತ ನಿರೀಕ್ಷೆ: ವಿಚ್ಛಿನ್ನ ಮತ್ತು ಸತತ ಅಕಸ್ಮಿಕ ಚಲಕಗಳ ವ್ಯಾಖ್ಯೆ ಮತ್ತು ಉದಾಹರಣೆಗಳು. ಸಂಭವ ರಾಶಿ ಫಲನೆ ಮತ್ತು ಸಂಭವ ಸಾಂದ್ರತಾ ಫಲನೆಗಳ ವ್ಯಾಖ್ಯೆ. ದ್ವಿಚಲಕ ಮತ್ತು ಅಂಚಿನ ಬಿಡಿಯ ಸಂಭವ ವಿವರಣೆಗಳ ವ್ಯಾಖ್ಯೆ, ಉದಾಹರಣೆಗಳು. ವಿಚ್ಛಿನ್ನ ಅಕಸ್ಮಿಕ ಚಲಕದ ಗಣಿತ ನಿರೀಕ್ಷೆ ಮತ್ತು ಅಕಸ್ಮಿಕ ಚಲಕದ ಇತರೆ ಫಲನೆಗಳ ವ್ಯಾಖ್ಯೆ, ಗಣಿತ ನಿರೀಕ್ಷೆಯ ಪದೋಕ್ತಿಗಳಲ್ಲಿ ವಿಚಲನೆ ಮತ್ತು ನಿಯತ ವಿಚಲನೆ, ಸಾಂಖ್ಯಿಕ ಮತ್ತು ಶಾಖ್ಯಿಕ ಲೆಕ್ಕಗಳು (ಕಾಣೆಯಾದ ಸಂಭವನೀಯತೆಗಳು ಸೇರಿ). a ಮತ್ತು b ಎರಡು ಸ್ಥಿರ ಸಂಖ್ಯೆಗಳಾದಾಗ a, ax ಮತ್ತು ax±b ಗಳ ಗಣಿತ ನಿರೀಕ್ಷೆ ಮತ್ತು ವಿಚಲನೆಗಳ ಸಾಧನೆ ಮತ್ತು ಸಂಬಂಧಿಸಿದ ಲೆಕ್ಕಗಳು. ಎರಡು ಅಕಸ್ಮಿಕ ಚಲಕಗಳಿಗೆ ಗಣಿತ ನಿರೀಕ್ಷೆಯ ಸಂಕಲನ ಮತ್ತು ಗುಣಾಕಾರ ಪ್ರಮೇಯಗಳ ಹೇಳಿಕೆ ಮತ್ತು ಸಾಧನೆ ಹಾಗೂ ಸಂಬಂಧಿತ ಲೆಕ್ಕಗಳು. ಅಕಸ್ಮಿಕ ದ್ವಿಚಲಕಗಳ ಸಹವಿಚಲನೆ ಮತ್ತು ಸಹಸಂಬಂಧಗಳ ಸೂತ್ರ ಮತ್ತು ಲೆಕ್ಕಗಳು.

*****

ಕರ್ನಾಟಕ ಸರ್ಕಾರ

ಸಂಖ್ಯೆ ಕೈಶಾ/1/97-98.

ನಿರ್ದೇಶಕರ ಕಛೇರಿ
ಪದವಿ ಪೂರ್ವ ಶಿಕ್ಷಣ ಇಲಾಖೆ
ಅರಮನೆ ರಸ್ತೆ ಬೆಂಗಳೂರು-1
ದಿನಾಂಕ: 18-08-1997.

ಸುತ್ತೋಲೆ

ವಿಷಯ:- ಸಂಖ್ಯಾಶಾಸ್ತ್ರ ವಿಷಯದಲ್ಲಿ ಪ್ರಾಯೋಗಿಕ ತರಗತಿಗಳ ಬೋಧನಾ ಅವಧಿಯ ಬಗ್ಗೆ ಮರು ಸ್ಪಷ್ಟೀಕರಣ.

1995-96ನೇ ಸಾಲಿನಿಂದ ಜಾರಿಗೆ ಬಂದ ಹೊಸ ಪಠ್ಯಕ್ರಮದ ಪ್ರಕಾರ ಸಂಖ್ಯಾಶಾಸ್ತ್ರ ವಿಷಯಕ್ಕೆ ವಾರಕ್ಕೆ ಮೂರು ಗಂಟೆಗಳ ತತ್ವ (ಬೋಧನೆ) ತರಗತಿ ಮತ್ತು ಎರಡು ಗಂಟೆಗಳ ಪ್ರಾಯೋಗಿಕ ಅವಧಿಯನ್ನು ಗೊತ್ತುಪಡಿಸಲಾಗಿದ್ದು, 20 ವಿದ್ಯಾರ್ಥಿಗಳ ತಂಡವನ್ನು ರಚಿಸಲು ತಿಳಿಸಲಾಗಿತ್ತು. ಕೆಲವು ಕಾಲೇಜುಗಳ ಪ್ರಾಚಾರ್ಯರು 20 ವಿದ್ಯಾರ್ಥಿಗಳ ತಂಡವನ್ನು ರಚಿಸುವುದರಿಂದ ಬೋಧನಾ ಅವಧಿ ಹೆಚ್ಚಾಗಿ ಹೆಚ್ಚಿನ ಉಪನ್ಯಾಸಕರು ಬೇಕಾಗುತ್ತದೆ ಮತ್ತು ಹೊಸ ಹುದ್ದೆಗಳನ್ನು ನಿರ್ಮಿಸಬೇಕಾಗುತ್ತದೆ ಎಂದು ತಿಳಿಸಿದ್ದರಿಂದ ಈ ಕಛೇರಿಯಿಂದ ಸುತ್ತೋಲೆಯನ್ನು ಹೊರಡಿಸಿ ತಂಡಗಳನ್ನು ರಚಿಸುವ ಅಗತ್ಯವಿಲ್ಲವೆಂದು ತಿಳಿಸಲಾಗಿದ್ದು, ಕೇವಲ ಐದು ಗಂಟೆಗಳ ಬೋಧನಾ ಅವಧಿಯನ್ನು ಗೊತ್ತುಪಡಿಸುವಂತೆ ಪ್ರಾಚಾರ್ಯರಿಗೆ ತಿಳಿಸಲಾಗಿತ್ತು. ಒಂದು ವರ್ಷ ಈ ರೀತಿ ಪಾಠ ನಡೆದ ನಂತರ ಅನೇಕ ಸಂಖ್ಯಾಶಾಸ್ತ್ರದ ಉಪನ್ಯಾಸಕರು, ಕರ್ನಾಟಕ ಸಂಖ್ಯಾಶಾಸ್ತ್ರ ಅಧ್ಯಾಪಕರ ಮಹಾಮಂಡಳ ಇಲಾಖೆಗೆ ಮನವಿಗಳನ್ನು ಸಲ್ಲಿಸಿ ತಂಡಗಳನ್ನು ರಚಿಸದೆ ಇರುವುದರಿಂದ ಸಂಖ್ಯಾಶಾಸ್ತ್ರದ ಶಿಕ್ಷಣ ವ್ಯವಸ್ಥೆಗೆ ತೊಂದರೆಯಾಗುತ್ತಿರುವುದಾಗಿ ತಿಳಿಸಿದ್ದಾರೆ. ಕೆಲವು ಅಧ್ಯಾಯಗಳನ್ನು ಪ್ರಾಯೋಗಿಕ ಪಠ್ಯದಂತೆ ನಡೆಸಿಕೊಡಲು ಸಾಧ್ಯವಿಲ್ಲವೆಂದು ತಿಳಿಸಿದ್ದಾರೆ. ತಂಡ ರಚಿಸದೆ ಇರುವುದರಿಂದ ಗಣಕ ಪರಿಮಾಣಾತ್ಮಕ ದತ್ತಾಂಶಗಳ ವಿಶ್ಲೇಷಣೆ, ಬೃಹತ್ ಸಂಖ್ಯೆಗಳ ವಿಸ್ತರಣೆಗಳ ಸಂಯೋಜನೆ, ನ್ಯಾಸ ಸೂಚನೆ, ನ್ಯಾಸ ಪ್ರದರ್ಶನ, ಇತ್ಯಾದಿ ಪ್ರಕ್ರಿಯೆಗಳನ್ನು ವಿದ್ಯಾರ್ಥಿಗಳಿಂದ ವೈಯಕ್ತಿಕ ಗಮನ ಕೊಟ್ಟು ಮಾಡಿಸಲು ಸಾಧ್ಯವಾಗುತ್ತಿಲ್ಲ ಎಂದು ತಿಳಿಸಿದ್ದಾರೆ. ಈ ಎಲ್ಲಾ ವಿಷಯಗಳನ್ನು ಗಮನದಲ್ಲಿಟ್ಟುಕೊಂಡು ಈ ವಿಷಯವನ್ನು ಮರು ಪರಿಶೀಲಿಸಿದ ನಂತರ ಈ ಮರು ಸ್ಪಷ್ಟೀಕರಣ ನೀಡಲಾಗಿದೆ.

ಪದವಿ ಪೂರ್ವ ಶಿಕ್ಷಣ ತರಗತಿಗಳ ಒಂದು ವಿಭಾಗಕ್ಕೆ ಗರಿಷ್ಠ ಸಂಖ್ಯೆ ಐವತ್ತು ನೂರು ಇರುವುದರಿಂದ ವಿದ್ಯಾರ್ಥಿಗಳನ್ನು ಮೂರು ತಂಡಗಳನ್ನಾಗಿ ರಚಿಸಿ ಪ್ರತಿ ತಂಡಕ್ಕೆ ಎರಡು ಅವಧಿಗಳ ಪ್ರಾಯೋಗಿಕ ತರಗತಿಗಳನ್ನು ನಡೆಸಲು ಸೂಚಿಸಲಾಗಿದೆ. ಇದರಿಂದ ಒಟ್ಟು ಬೋಧನಾ ಅವಧಿ (3+6) 9 ಒಂದು ವಿಭಾಗಕ್ಕೆ ಇರುತ್ತದೆ. ಪ್ರಥಮ ಮತ್ತು ದ್ವಿತೀಯ ಪಿ.ಯು. ತರಗತಿಗಳು ಸೇರಿ ಕೇವಲ 18 ಅವಧಿಗಳಾಗುತ್ತವೆ. ಇದರಿಂದ ಹೆಚ್ಚಿನ ಉಪನ್ಯಾಸಕರ ಅವಶ್ಯಕತೆ ಇರುವುದಿಲ್ಲ.

ಪದವಿ ಕಾಲೇಜುಗಳು, ಕಾಲೇಜು ಶಿಕ್ಷಣ ಆಯುಕ್ತರು ಹೊರಡಿಸುವ ಸುತ್ತೋಲೆಯಂತೆ ತಂಡಗಳನ್ನು ರಚಿಸಿ ಪ್ರಾಯೋಗಿಕ ತರಗತಿಗಳನ್ನು ನಡೆಸಲು ಅವಕಾಶ ನೀಡಲಾಗಿದೆ.

ಈ ಸೂಚನೆಯನ್ನು ಗಮನಿಸಿ ಪ್ರಾಚಾರ್ಯರು ವೇಳಾಪಟ್ಟಿಯನ್ನು ರಚಿಸಲು ಸೂಚಿಸಲಾಗಿದೆ.

ನಿರ್ದೇಶಕರು,
ಪದವಿ ಪೂರ್ವ ಶಿಕ್ಷಣ ಇಲಾಖೆ.

ಪ್ರತಿಯನ್ನು:
ಪದವಿ ಪೂರ್ವ ಶಿಕ್ಷಣವನ್ನು ಬೋಧಿಸುವ ಎಲ್ಲಾ ಕಾಲೇಜುಗಳ ಪ್ರಾಚಾರ್ಯರಿಗೆ ಸೂಕ್ತ ಕ್ರಮಕ್ಕಾಗಿ ಕಳುಹಿಸಲಾಗಿದೆ.

Contents

A

PRE–REQUISITES (Not for Examination)

Base : The factor which is repeatedly multiplied by it-self again and again is called base .

Indices : The number of times the factor appears is called the power or index. The plural form of index is called indices .

Ex : $3^2 = 3 \times 3 = 9$. Here, 3 is called base and 2 is called power .

Ex : $2^3 = 2 \times 2 \times 2 = 8$. Here, 2 is called base and 3 is called power.

Note : Any number to the power of '0' is '1'. That is $a^0 = 1$, where 'a' is any non-zero number .

Basic Laws of Indices :

(1) $a^m \times a^n = a^{m+n}$ (2) $\frac{a^m}{a^n} = a^{m-n}$ (3) $(a^m)^n = a^{mn}$

(4) $(ab)^m = a^m b^m$ (5) $\left(\frac{a}{b}\right)^m = \frac{a^m}{b^m}$ (6) $(a)^{\frac{1}{2}} = \sqrt{a}$

Ex: $3^3 \times 3^2 = 3^{3+2} = 3^5 = 243$

Ex: $4^3 \times 4^2 = 4^{3+2} = 4^5 = 1024$

Ex: $5^2 \times 5^1 = 5^{2+1} = 5^3 = 125$

Ex: $6^2 \times 6^0 = 6^{2+0} = 6^2 = 36$

Ex: $\frac{4^3}{4^2} = 4^{3-2} = 4^1 = 4$

Ex: $\frac{5^2}{5^1} = 5^{2-1} = 5^1 = 5$

Ex: $(2^3)^2 = 2^{3\times 2} = 2^6 = 64$

Ex: $(3^2)^2 = 3^{2\times 2} = 3^4 = 81$

Ex: $(3^2)^{\frac{1}{2}} = 3^{2\times\frac{1}{2}} = 3$

Ex: $(4^3)^{\frac{1}{3}} = 4^{3\times\frac{1}{3}} = 4$

Ex: $(5\times 2)^2 = 5^2 \times 2^2 = 25 \times 4 = 100$

Ex: $(3\times 4)^2 = (12)^2 = 144$

Ex: $(121 \times 144)^{\frac{1}{2}} = (121)^{\frac{1}{2}} \times (144)^{\frac{1}{2}}$
$= 11 \times 12 = 132$

Ex: $(25 \times 4)^{\frac{1}{2}} = (100)^{\frac{1}{2}} = 10$

Ex: $\left(\frac{20}{5}\right)^2 = \frac{20^2}{5^2} = \frac{400}{25} = 16$

Ex: $\left(\frac{15}{3}\right)^2 = (5)^2 = 25 = \frac{225}{9}$

Ex: $\left(\frac{36}{9}\right)^{\frac{1}{2}} = (4)^{\frac{1}{2}} = 2$

Ex: $\left(\frac{16}{4}\right)^{\frac{1}{2}} = \frac{(16)^{\frac{1}{2}}}{(4)^{\frac{1}{2}}} = \frac{\sqrt{16}}{\sqrt{4}} = \frac{4}{2} = 2$

Logarithms: The logarithm of a number to a given base is the index or the power to which the base must be raised to get the given number. Thus if $a^m = x$, then ' **m** ' is called ' **logarithm of x to the base 'a '** and usually written as $m = \log_a x$

Note: Logarithms with base 10 is called common logarithms. Here, the base is not to be mentioned explicitly . That is log x stands for $m = \log_a x$

The logarithm of any number has two parts viz. , integral part and decimal part. The integral part is called as **characteristic** and decimal part is called as **mantissa .** A characteristic may be positive or negative where as mantissa is always positive.

Characteristic of the given number :

The characteristic of numbers greater than 1 , is positive and is one less than number of digits before the decimal point . The characteristic of numbers less than 1 , is negative and numerically one more than the number of zeros between the decimal point and the first significant digit . The negative characteristic is written as bar . Thus, $\bar{1}$ means -1.

Number	Charact eristic
121.0	2
12.1	1
1.21	0
0.121	$\bar{1}$
0.0121	$\bar{2}$

Mantissa of the given number :

The mantissa of a number is obtained from the logarithm table. In the logarithm table mantissa are arranged in 90 rows (10 to 99) and 10 columns (0 - 9) together with 9 more mean differences columns (1 - 9). Mantissa of a number is not affected by the position of the decimal point in the number. Thus, mantissa of 12.1, 1.21 and 0.121 are same.

Number	Mantissa
12.10	0828
1.21	0828
0.121	0828

Logarithm Table

	0	1	2	3	4	5	6	7	8	9	Mean Differences								
											1	2	3	4	5	6	7	8	9
10	0000	0043	0086	0128	0170	0212	0253	0294	0334	0374	4	8	12	17	21	25	29	33	37
11	0414	0453	0492	0531	0569	0607	0645	0682	0719	0755	4	8	11	15	19	23	26	30	34
12	0792	**0828**	0864	0899	0934	0969	1004	1038	1072	1106	3	7	10	14	17	21	24	28	31
13	1139	1173	1206	1239	1271	1303	1335	1367	1399	1430	3	6	10	13	16	19	23	26	29
14	1461	1492	1523	1553	1584	1614	1644	1673	1703	1732	3	6	9	12	15	18	21	24	27

In logarithm table see row beginning with 12 under the column 1 we get 0828 . Thus , the mantissa is .0828 $\Rightarrow$ log(12.1) = 1.0828

Antilogarithm of the given number :

The antilogarithm table gives significant figures corresponding to first four digits of known mantissa . Here, the figures are arranged in 99 rows (00 - 99) and 10 columns (0 - 9) together with 9 more mean differences columns (1 - 9) . The position of the decimal point is decided by the characteristic .

Antilogarithm Table

	0	1	2	3	4	5	6	7	8	9	Mean Differences								
											1	2	3	4	5	6	7	**8**	9
00	1000	1002	1005	1007	1009	1012	1014	1016	1018	1021	0	0	1	1	1	1	2	2	2
01	1023	1026	1028	1030	1033	1035	1038	1040	1042	1045	0	0	1	1	1	1	2	2	2
02	1047	1050	1052	1054	1057	1059	1062	1064	1067	1069	0	0	1	1	1	1	2	2	2
03	1072	1074	1076	1079	1081	1084	1086	1089	1091	1094	0	0	1	1	1	1	2	2	2
04	1096	1099	1102	1104	1107	1109	1112	1114	1117	1119	0	1	1	1	1	2	2	2	2
05	1122	1125	1127	1130	1132	1135	1138	1140	1143	1146	0	1	1	1	1	2	2	2	2
06	1148	1151	1153	1156	1159	1161	1164	1167	1169	1172	0	1	1	1	1	2	2	2	2
07	1175	1178	1180	1183	1186	1189	1191	1194	1197	1199	0	1	1	1	1	2	2	2	2
08	1202	1205	**1208**	1211	1213	1216	1219	1222	1225	1227	0	1	1	1	1	2	2	**2**	3
09	1230	1233	1236	1239	1242	1245	1247	1250	1253	1256	0	1	1	1	1	2	2	2	3
10	1259	1262	1265	1268	1271	1274	1276	1279	1282	1285	0	1	1	1	1	2	2	2	3
11	1288	1291	1294	1297	1300	1303	1306	1309	1312	1315	0	1	1	1	2	2	2	2	3
12	1318	1321	1324	1327	1330	1334	1337	1340	1343	1346	0	1	1	1	2	2	2	2	3

In antilogarithm table see row beginning with 08 under the column 2 and the mean differences column 8 in same row . Thus , we get 1208 + 2 = 1210 $\Rightarrow$ antilog (1.0828) = 12.10

Basic Laws of Logarithms :

(i) $\log xy = \log x + \log y$ (ii) $\log\left(\frac{x}{y}\right) = \log x - \log y$ (iii) $\log x^k = k \log x$

Ex: $\sqrt{(121 \times 144)}$

$= (121 \times 144)^{\frac{1}{2}}$

$= \text{Antilog}\,[\log (121 \times 144)^{\frac{1}{2}}]$

$= \text{Antilog}\,[\frac{1}{2}\{\log (121 \times 144)\}]$

$= \text{Antilog}\,[\frac{1}{2}(\log 121 + \log 144)]$

$= \text{Antilog}\,[\frac{1}{2}(2.0828 + 2.1584)]$

$= \text{Antilog}\,[\frac{1}{2}(4.2412)]$

$= \text{Antilog}\,(2.1206)$

$= 132.0$

Ex: $\sqrt[3]{(2 \times 4 \times 8)}$

$= (2 \times 4 \times 8)^{\frac{1}{3}}$

$= \text{Antilog}\,[\log (2 \times 4 \times 8)^{\frac{1}{3}}]$

$= \text{Antilog}\,[\frac{1}{3}\{\log (2 \times 4 \times 8)\}]$

$= \text{Antilog}\,[\frac{1}{3}(\log 2 + \log 4 + \log 8)]$

$= \text{Antilog}\,[\frac{1}{3}(0.3010 + 0.6021 + 0.9031)]$

$= \text{Antilog}\,[\frac{1}{3}(1.8062)]$

$= \text{Antilog}\,(0.60206)$

$= \text{Antilog}\,(0.6021)$

$= 4.000$

Ex: $\sqrt{(12 \times 3)}$

$= (12 \times 3)^{\frac{1}{2}}$

$= \text{Antilog}\,[\log (12 \times 3)^{\frac{1}{2}}]$

$= \text{Antilog}\,[\frac{1}{2}\{\log (12 \times 3)\}]$

$= \text{Antilog}\,[\frac{1}{2}(\log 12 + \log 3)]$

$= \text{Antilog}\,[\frac{1}{2}(1.0792 + 0.4771)]$

$= \text{Antilog}\,[\frac{1}{2}(1.5563)]$

$= \text{Antilog}\,(0.77815)$

$= \text{Antilog}\,(0.7782) = 6.00$

Ex: $\sqrt[4]{(2 \times 4 \times 4 \times 8)}$

$= (2 \times 4 \times 4 \times 8)^{\frac{1}{4}}$

$= \text{Antilog}\,[\log (2 \times 4 \times 4 \times 8)^{\frac{1}{4}}]$

$= \text{Antilog}\,[\frac{1}{4}\{\log (2 \times 4 \times 4 \times 8)\}]$

$= \text{Antilog}\,[\frac{1}{4}(\log 2 + \log 4 + \log 4 + \log 8)]$

$= \text{Antilog}\,[\frac{1}{4}(0.3010 + 0.6021 + 0.6021 + 0.9031)]$

$= \text{Antilog}\,[\frac{1}{4}(2.4083)]$

$= \text{Antilog}\,(0.60207)$

$= \text{Antilog}\,(0.6021)$

$= 4.000$

Ex: $\left(\frac{15}{3}\right)^2$

$= \text{Antilog}\left[\log\left(\frac{15}{3}\right)^2\right]$

$= \text{Antilog}\left[2\log\left(\frac{15}{3}\right)\right]$

$= \text{Antilog}\left[2\left(\log 15 - \log 3\right)\right]$

$= \text{Antilog}\left[2\left(1.1761 - 0.4771\right)\right]$

$= \text{Antilog}\left[2\left(0.6990\right)\right]$

$= \text{Antilog}\left(1.3980\right)$

$= 25.00$

Ex: $\left(\frac{20}{4}\right)^3$

$= \text{Antilog}\left[\log\left(\frac{20}{4}\right)^3\right]$

$= \text{Antilog}\left[3\log\left(\frac{20}{4}\right)\right]$

$= \text{Antilog}\left[3\left(\log 20 - \log 4\right)\right]$

$= \text{Antilog}\left[3\left(1.3010 - 0.6021\right)\right]$

$= \text{Antilog}\left[3\left(0.6989\right)\right]$

$= \text{Antilog}\left(2.0967\right)$

$= 124.9$

Solving simultaneous linear equations :

A pair of values for x and y which satisfy both linear equations is called a solution of two simultaneous linear equations .

To solve the simultaneous linear equations following steps are to be performed :

Step 1)if the coefficients of any one of the unknowns are not equal , then multiply the equations by the suitable constants , to make the coefficients of any one of the unknowns are equal .

Step 2)if the signs of the equal coefficients are like, then subtract the resulting equations ; if the signs are unlike, add them .

Examples : Solve the following simultaneous linear equations:

1. $x + y = 10$ ———(1)

   $x - y = 2$ ———(2)

Here, the coefficients of y are equal and the signs are unlike, therefore by adding these two equations , we get $2x = 12 \Rightarrow x = 6$

By substituting this value of x in any one the above equations , we get y = 4.

2. $3x + 7y = 36$ ———(1)

$3x + 2y = 21$ ———(2)

Here, the coefficients of x are equal and the signs are like, therefore by subtracting 2nd equation from 1st equation, we get $5y = 15. \Rightarrow y = 3$

By substituting this value of y in any one the above equations, we get x =5.

3. $2x + 3y = 8$ ———(1)

$3x + 5y = 13$ ———(2)

Here, the coefficients of either x or y are not equal , therefore the coefficient of x is made equal, by multiplying the 1st equation by 3 (i.e., coefficient of x of the 2nd equation) and the 2nd equation by 2 (i.e., coefficient of x of the 1st equation), we get

$6x + 9y = 24$ ———(3)

$6x + 10y = 26$ ———(4)

The signs are like, therefore by subtracting 3rd equation from 4th equation, we get y = 2 .

By substituting this value of y in any one the above equations, we get x =1.

4. $x + 5y = 20$ ———(1)

$2x + 7y = 31$ ———(2)

Here, to make the coefficient of x equal, multiply the 1st equation by 2, we get

$2x + 10y = 40$ ———(3)

The signs are like, therefore by subtracting 2nd equation from 3rd equation,

we get $3y = 9 \Rightarrow y = 3$

By substituting this value of y in any one the above equations, we get $x = 5$.

5. $8x + 5y = 21$ ———(1)

$7x - 3y = 11$ ———(2)

To make the coefficient of y equal , multiply the 1st equation by 3 (i.e., coefficient of y of the 2nd equation) and the 2nd equation by 5 (i.e. , coefficient of y of the 1st equation) ,

we get $24x + 15y = 63$ ———(3)

$35x - 15y = 55$ ———(4)

The signs are unlike, therefore by adding these two equations, we get

$59x = 118 \Rightarrow x = 2$

By substituting this value of x in any one the above equations, we get $y = 1$.

Set :

A set is a collection of well defined and distinct objects. The objects in the set are called as elements or members. Usually, sets are denoted by capital letters A, B, C etc and their elements are denoted by small letters a, b, c etc. Under roster method all the elements in the sets are written in flower brackets and are being separated by commas.

A set without element is empty set or null set. It is denoted by Φ (Phi). A set with only one element is singleton set . A set with all possible elements is universal set. It is denoted by 'U'.

Ex: The universal set of vowels of English alphabets is U = { a, e, i, o, u } . Vowel of English alphabets before 'e' is a singleton set. That is {a} . Set of even prime number is also a singleton set. That is { 2 } . Vowel of English alphabets after 'u' is a null set. That is Φ

A set with definite number of elements is called finite set; otherwise, it is an infinite set.

Ex : The set of letters of English alphabets is finite set and the set of natural numbers is infinite set.

Union of sets :

A set formed by the elements belonging to at least one of the sets is union of those sets . If A and B are any two sets , then their union is denoted by (A∪B) or (A+B) or (A or B) .

Ex: If A = { 2, 4, 6 } , B = { 1, 3, 5 } then (A∪B) = { 1, 2, 3, 4, 5, 6 }

Ex: If A = { 2, 4, 6 } , C = { 5 } then (A∪C) = { 2, 4, 5, 6 }

Ex: If A = { 2, 4, 6 } , D = { 3, 6 } then (A∪D) = { 2, 3, 4, 6 }

Ex: If A = { 2, 4, 6 } , E = { 1, 2, 4, 5 } then (A∪E) = { 1, 2, 4, 5, 6 }

Intersection of sets :

A set formed by the elements belonging to all the sets (common elements) is intersection of those sets . If A and B are any two sets, then their intersection is denoted by (A∩B) or (AB) or (A and B).

Ex: If A = { 2, 4, 6 } , B = { 1, 3, 5 } (A∩B) = { } = Φ

Ex: If A = { 2, 4, 6 } , C = { 5 } then (A∩C) = ϕ

Ex: If A = { 2, 4, 6 } , D = { 3, 6 } then (A∩D) = { 6 }

Ex: If A = { 2, 4, 6 } , E = { 1, 2, 4, 5 } then (A∩E) = { 2, 4 }

Disjoint (Mutually exclusive) sets :

Two or more sets are said to be disjoint or mutually exclusive , if their intersection is null set.

Ex : If A = { 2, 4, 6 } , B = { 1, 3, 5 } then (A∩B) = { } = ϕ

Ex : If A = { 2, 4, 6 } , C = { 5 } then (A∩C) = ϕ

Complement of a set :

For any set A, the complement of A is a set formed by the elements of universal set by excluding the elements of set A. It is denoted by

A^C or $\overline{A}$ (A-bar) or A^1 (A-dash).

Note: Complementary sets are always mutually exclusive .

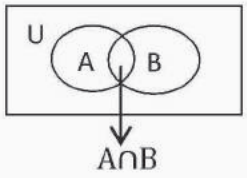
Ex: If U = {1, 2, 3, 4, 5, 6} and A = {2, 4, 6} then A^1 = U – A= {1, 3, 5}. Here, $(A \cap A^1) = \phi$ and $(A \cup A^1) = U$

Subset :

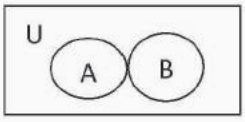
The set B is said to be a subset of A if every element of B is an element of A. That is $B \subset A$.

Venn diagram :

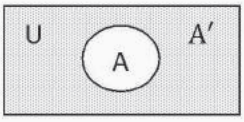
Diagrammatic representation of sets is called Venn diagram.



$A \cap B$



A and B are mutually exclusive



Complement of A

Relation between numbers of elements of the sets :

If n(A), n(B), $n(A \cup B)$, $n(A \cap B)$ are numbers of elements in the corresponding sets, then $n(A) + n(B) = n(A \cup B) + n(A \cap B)$.

Example : There are 60 students in a class . Every student learns at least one of the subjects Kannada or English . 45 students offer Kannada and 30 students offer English . How many students offer (i) both the subjects Kannada and English (ii) only Kannada (iii) only English ?

Solution:

Here, n(K) = 45, n(E) = 30 and $n(K \cup E) = 60$

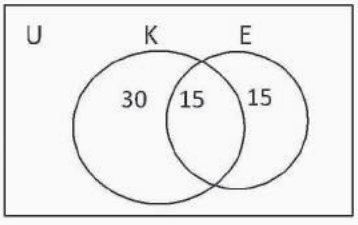
We know that, $n(K \cup E) = n(K) + n(E) - n(K \cap E)$

We know that,

(i) $n(K \cap E) = 45 + 30 - 60 = 15$

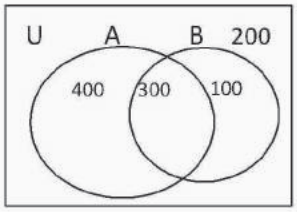
(ii) $n(\text{Only Kan}) = n(K) - n(K \cap E) = 45 - 15 = 30$

(iii) $n(\text{Only Eng}) = n(E) - n(K \cap E) = 30 - 15 = 15$



Example: In a survey of 1000 persons in Bangalore, it was found that 700 read news paper A, 400 read news paper B and 300 read both news papers A and B. How many persons (i) read at least one news paper (ii) do not read both ?

Solution:



We know that,

$n(A \cup B) = n(A) + n(B) - n(A \cap B)$

(i) $n(A \cup B) = 700 + 400 - 300 = 800$

(ii) $n(A \cup B)' = n(U) - n(A \cup B) = 1000 - 800 = 200$

De Morgan's Laws :

I. $(A \cup B)^1 = (A^1 \cap B^1)$ The complement of union of sets is intersection of their complements.

II. $(A \cap B)^1 = (A^1 \cup B^1)$ The complement of intersection of sets is union of their complements.

Example: Verify De Morgan's Laws If $U = \{1, 2, 3, 4, 5, 6\}$, $A = \{2, 4, 6\}$ and $B = \{3, 6\}$.

Solution: We know that, $A^1 = (U - A) = \{1, 3, 5\}$

(Set of elements of U except the elements of A)

Similarly, $B^1 = (U - B) = \{1, 2, 3, 4, 5, 6\} - \{3, 6\} = \{1, 2, 4, 5\}$

Here, $(A \cup B) = \{2, 3, 4, 6\}$, $(A \cap B) = \{6\}$, $(A' \cup B') = \{1, 2, 3, 4, 5\}$ and $(A' \cap B') = \{1, 5\}$.

Now, $(A \cup B)^1 = U - (A \cup B) = \{1, 5\} = (A' \cap B')$

and $(A \cap B)^1 = U - (A \cap B) = \{1, 2, 3, 4, 5\} = (A' \cup B')$

Order Pair :

A pair of values is written in the form (x, y) is called order pair . Here , x is the first value and y is the second value .

Factorials :

The product of first n natural numbers is n factorial denoted by 'n!' .

That is $n! = 1 \times 2 / (n-1) \times n = n \times (n-1) \ldots\ldots 2 \times 1 = n \times (n-1)!$

Example : Find the value of 3! .

Solution : $3! = 3 \times 2 \times 1 = 6$

Example : Find the value of 4! .

Solution : $4! = 4 \times 3! = 4 \times 6 = 24$

Example : Find the value of 6! .

Solution : $6! = 6 \times 5 \times 4! = 6 \times 5 \times 24 = 720$

Permutations :

The number of selections and arrangements (permutations) of 'r' items out of 'n' different items is ${}^{n}P_{r}$, given by ${}^{n}P_{r} = \frac{n!}{(n-r)!}$

Example: Evaluate ${}^{5}P_{2}$.

Solution: ${}^{5}P_{2} = \frac{5!}{(5-2)!} = \frac{5 \times 4 \times 3!}{3!} = 20$

Example: Evaluate ${}^{10}P_{3}$.

Solution: ${}^{10}P_{3} = \frac{10!}{(10-3)!} = \frac{10 \times 9 \times 8 \times 7!}{7!} = 720$

Example: Evaluate ${}^{100}P_{2}$.

Solution: ${}^{100}P_{2} = \frac{100!}{(100-2)!} = \frac{100 \times 99 \times 98!}{98!} = 9900$

Example: Evaluate ${}^{4}P_{1}$.

Solution: ${}^{4}P_{1} = \frac{4!}{(4-1)!} = \frac{4 \times 3!}{3!} = 4$

Example: How many photos are to be suited by taking 2 persons out of 3.

Solution: Here, the number of photos are to be suited is ${}^{3}P_{2} = \frac{3!}{(3-2)!} = \frac{3 \times 2 \times 1!}{1!} = 6$

Example: Evaluate ${}^{3}P_{3}$.

Solution: ${}^{3}P_{3} = \frac{3!}{(3-3)!} = \frac{3!}{0!} = \frac{3 \times 2 \times 1}{1} = 6$

Note: 0! = 1

Combinations :

The number of selections (combination) of 'r' items out of 'n' different items is ${}^{n}c_{r}$, given by ${}^{n}c_{r} = \frac{n!}{(n-r)!\, r!}$

Example: Evaluate ${}^{5}c_{2}$.

Solution: ${}^{5}c_{2} = \frac{5!}{(5-2)!2!} = \frac{5 \times 4 \times 3!}{3! \times 2 \times 1} = 10$

Example: Evaluate ${}^{10}c_{3}$.

Solution: ${}^{10}c_{3} = \frac{10!}{(10-3)!3!} = \frac{10 \times 9 \times 8 \times 7!}{7! \times 3 \times 2 \times 1} = 120$

Example: Evaluate ${}^{10}c_{7}$.

Solution: ${}^{10}c_{7} = \frac{10!}{(10-7)!7!} = \frac{10!}{3! \times 7!} = \frac{10 \times 9 \times 8 \times 7!}{3 \times 2 \times 1 \times 7!} = 120$

Note: ${}^{n}c_{r} = {}^{n}c_{n-r}$, ${}^{n}c_{0} = 1 = {}^{n}c_{n}$, ${}^{n}c_{1} = n = {}^{n}c_{n-1}$

Example: Find n, if ${}^{n}c_{7} = {}^{n}c_{3}$.

Solution: We know that, ${}^{n}c_{r} = {}^{n}c_{n-r}$. That is, n = r + n - r = 7 + 3 = 10

Example: A bag contains 4 white balls and 6 black balls. In how many ways 2 balls are to be selected?

Solution: Here, total number of balls is 10 and selection of 2 balls out of 10 balls is ${}^{10}c_{2}$ ways.

That is, ${}^{10}c_{2} = \frac{10!}{(10-2)!2!} = \frac{10 \times 9 \times 8!}{8! \times 2 \times 1} = 45$

Example: How many cricket teams are to be formed with 11 players out of 15 players?

Solution: Here, the number cricket teams are to be formed in ${}^{15}c_{11}$ ways.

That is, ${}^{15}c_{11} = \frac{15!}{(15-11)!11!} = \frac{15!}{4! \times 11!} = \frac{15 \times 14 \times 13 \times 12 \times 11!}{4 \times 3 \times 2 \times 1 \times 11!} = 1365.$

******

Unit-I

INTRODUCTION - MEANING AND SCOPE

Origin and Development of Statistics :

When man started counting, comparing and keeping records in one or the other way, the subject 'Statistics' seems to have originated. Statistics took birth when the people started expressing their ideas, such as members in a family, live stock, and quantities given or taken in the barter system etc., in numerical terms. As man started living in a society, keeping records of many social aspects became a necessity. Farmers needed the record of the climatic conditions during a year ; doctors needed the record of effect of different herbals on different diseases. Such information was passed on from one person to the other or one generation to the next. Gradually, when kings started ruling, figures regarding the army, population of their kingdom, value of tax, expenditures, volume of trades, lands etc., were to be maintained. Different departments were to be established to maintain the records of these aspects. A ruler could be successful only through proper management of these departments. Obviously, only an organized and a dynamic body of the state could venture into collection of statistics which in the past was mainly on population, its composition and wealth. In the history, we come across famous rulers, throughout the world, many of whom are associated with reforms in keeping such records. In this way Statistics was related to the **'State Craft'**. Hence, in the olden days Statistics was regarded as the Science of State Craft and was the by-product of the administrative activity of a state. But at that time this procedure was not called as Statistics. May be it amounted to only mathematics or any interpretation of it as logic (**Tarka Shasthra**).

The word Statistics seems to have been derived from the Latin word

'status' or the Italian word **'statista'** or the German word **'statistik'** or the French word **'statistique'**, each of which means a **political state.** In the ancient times the scope of Statistics was primarily limited to the collection of the data of population or wealth of the country by the governments for framing military and economic policies.

According to history, evidence of collection of population statistics in ancient India is found during the reign of **Chandragupta Maurya.** Kautilya's Arthashastra elaborates the management of statistical facts about State administration in the country. The tax system on lands was reformed by Todarmal during the reign of Akbar. Abul Fazal (in 1656-97) in his book Ain-e-Akbari has written a detailed account of the administrative and statistical surveys conducted during reign of Akbar.

At present Statistics has grown at a phenomenal rate. The application of Statistics is done in almost all the disciplines of knowledge. Statistics is widely used as a valuable tool in the analysis of the problems in physical and social sciences, commerce and economics and also in science and technology. The subjects like Medicine use statistics to draw valid and meaningful conclusions.

The 'theory of probability' originated as the 'theory of game and chances', during the 17^{th} century by James Bernoulli (1654-1705), De-Morve (1667-1754), Laplace (1749-1827), Gauss (1777-1855). It is De-Morve who discovered the 'Normal curve'. Sir Francis Galton (1820-1911) developed 'Regression theory' in the field of Biometry. Karl Pearson (1857-1936) came out with latest discovery ' Chi-square test of goodness fit ' which is of very high value in research activities in applied science. His contribution to 'Correlation theory' is also prominent. The discovery of Student's t-distribution by W.S.Gosset(1876-1937) is a mile stone in the growth of statistics.

Among the modern stalwarts of Statistics, **Sir Ronald A. Fisher (1890-1962)** is considered as **Father of Statistics**. He developed the theories

which could be used in various fields like Biometry, Psychology, Education, Agriculture, Commerce, etc..

The Indian scholar Prashantha Chandra Mahalanobis (1893–1972) called as father of Indian statistics and was the founder of Indian Statistical Institute .

Definitions of statistics :

The word Statistics represents different meanings with reference to different contexts are as given below :

(i) Numerical statement of facts or Statistical data .

Statistics stands for the numerical information from which certain conclusions can be drawn. In this sense the word Statistics is used as "figures or numerical information".

For example,:

1. "Sales statistics of different products of a company for past few years are recorded for a compartive study."
2. The heights of a group of 5 students in cms are : 156, 168, 180, 153, 172.

In these sentences the term statistics refers to numerical statements.

(ii) Statistical methods.

Just like "Accounts" stands for the principles and the methods of accounting, ledger, balance sheet etc., "Statistics" stands for the methods of collection, presentation, analysis and interpretation of numerical information. That is, to collect, represent, analyze and draw valid conclusions on the basis of numerical data.

Statistics as numerical statements of facts: (Plural sense)

Prof. Horace Secrist defined statistics as ' Statistics may be defined as the aggregate of facts, affected to a marked extent by multiplicity of causes, numerically expressed, enumerated or estimated according to a

reasonable standard of accuracy, collected in a systematic manner, for a predetermined purpose and placed in relation to each other '.

The above definition is of a comprehensive nature and explains which type of numerical data can be called as Statistics. The features or characteristics of Statistics are as follows :

1. **Statistics are aggregate of facts.**
   Statistics are a group of facts. Single or unrelated figures are not statistics, for the simple reason that such figures cannot be compared. For example, the age of a person in years is not statistics, but a series relating to the ages of a group of persons will be called statistics

2. **Statistics are affected to a marked extent by multiplicity of causes:**
   Generally speaking, statistical data are affected to a considerable extent by a number of causes. For instance, statistics of prices are affected by conditions of supply, demand, imports, exports and many other factors. Similarly, the statistics of wheat production is based on the factors such as rainfall, method of cultivation, soil fertility, quality of seeds, usage of manure etc.

3. **Statistics are numerically expressed.**
   All statistics are numerical statements, that is, statements are expressed in numbers. Qualitative expressions like good, bad, young, old etc., do not constitute statistics. Similarly, statements such as the population of India is rapidly increasing or the production of rice is sufficient etc., do not constitute statistics. Whereas the statement like Indian population is growing at the rate of 1.94 per 1000 per year, is statistics.

4. **Statistics are enumerated or estimated according to a reasonable standard of accuracy.**
   Statistical data must be reasonably accurate because, statistical data are used as a basis for statistical enquiry. If the scope of our enquiry is limited then we will collect the data by the method of actual counting. On the other hand, if the scope of enquiry is wide then we will collect

the data by the method of estimation. An estimate cannot be as accurate or precise as actual measurement. Hence, reasonable standard of accuracy must be observed.

5. Statistics are collected in a systematic manner.

It is essential that the data must be collected according to planned and scientific methods. Facts and figures collected in a haphazard manner would likely to lead a misleading result.

6. Statistics are collected for a pre-determined purpose.

The purpose of collecting statistical data must be decided in advance. It should be well-defined and specific. Data collected in an unsystematic manner and without complete awareness of the purpose will be confusing and cannot be made for valid conclusions.

7. Statistics should be placed in relation to each other.

Statistical data are collected mostly for the purpose of comparison. For the purpose of comparison, it is necessary that data must be homogeneous. It would be meaningless to compare the heights of men with heights of trees.

The statistical data are quantitative or numerical, that is, they can be counted, measured, compared, analyzed and are used to draw conclusions and make wise decisions.

Statistics as statistical methods : [singular sense]

The various activities of Statistics are collection, classification and tabulation of numerical facts, description and comparison, analysis and interpretation. A scientific enquiry is conducted for the purpose of decision-making in various fields. Statistics provides tools in performing these activities. The choice of methods will depend upon the nature of the data as well as the purpose for which it is collected. Hence, Statistics can be characterized as an applied science which helps in drawing conclusions.

A.L. Bowley defined Statistics as 'The Science of counting'. Afterwards,

he redefined it as ' The science of averages ' .

Boddington defined Statistics as 'The science of estimates and probabilities'.

A more precise and the exact definition given by Croxton and Cowden is "Statistics is the science of collection, presentation, analysis and interpretation of numerical data."

According to the above definitions we find that,

1. Statistics is an operational technique.
2. It helps in processing the raw data.
3. It is basically a tool of analysis.
4. The processing is done by the scientific methods of analysis.
5. Interpretation follows analysis.

Main Divisions of Statistics: (Branches)

Statistics have two branches, namely **Statistical Methods** and **Applied Statistics.**

Statistical methods are also called mathematical statistics or theory of statistics. It deals with the procedure of statistical analysis of numerical data. It is a tool for decision making.

Applied Statistics is the application of statistical methods to concrete situations like Agriculture, Industry, Population, Medicines etc. Some branches of Applied Statistics are,

1. Biometry (Bio Statistics which deals problems in Biology)
2. Demography (Study of human population)
3. Econometrics is a Statistical tool to prove or disprove the theories of Economics
4. Statistical Quality Control (control of quality of manufactured goods before marketing)
5. Actuarial Science (Statistics in the field of Insurance)
6. Stylometry (Statistics in Literature)
7. Psychometry (which deals with psychological problems).

Functions of Statistics :

The followings are the important functions of Statistics :

1. **It presents the facts in a definite form :**
   It is with the help of figures we can represent the things in their true form. Without a statistical study our ideas are vague and indefinite. So the facts should be presented in definite form. If conclusions are given in numbers, then they are more convincing than the conclusions are expressed on the basis of quality. The statements like, there is lot of unemployment in India or population is increasing at a faster rate are not in definite form. The statement, price level in 2010 is 15% less as compared to the price level in 2008, is in a definite form.

2. **It simplifies and condenses the size of the data :**
   Statistics presents large volumes of complex data in meaningful and understandable form.

3. **It facilitates comparison :**
   Statistics makes the comparison of particular phenomenon possible through various techniques like averages, time series, index numbers, measures of dispersion, measures of skewness etc. Comparison enables us to understand the behavior of data over a given time period

4. **It enriches our knowledge and widens our experience :**
   Statistics provides opportunities to enrich our knowledge and enhance our experience with the help of statistical techniques and methodology.

5. **It helps in formulating policies :**
   Statistics facilitates the formulation of various economic and other policies at State, National or Global level. Business organizations also make use of statistics in the formulation of polices in the area of finance, marketing and personnel.

6. **It helps in Business forecasting :**
   Forecasting refers to the process of predicting future events. Statistics facilitates business forecasting through various techniques as time series, extrapolation etc..

7. It helps in framing and testing of hypothesis :

Statistics facilitates the formulations of hypothesis, testing the validity of hypothesis and developing a theory on the basis of results.

Limitations of Statistics :

Some limitations of Statistics are as follows :

1. It does not deal with individuals :

Statistics deals with aggregate of facts and not with individuals. Individual do not constitute statistical data. For example, Lakshmi is a tall girl is not a statistical statement.

2 It does not deal with qualitative data :

As we have seen, statistics is a collection of numerical facts and the tools of statistics can be applied only for quantitative measurements. It is not possible to use these tools for qualitative characteristics like beauty, complexion, melody, character, taste, intelligence etc., which can be compared but cannot be measured in figures. Therefore in a study of such attributes, statistical methods are not of much use

3. Statistical laws are true only on an average :

Natural sciences are exact in the sense that their results are universally true. Statistical laws are not exact. For instance, the average number of accidents in a certain road is four per week does not mean that ' if all the four accidents have occurred in the first day of the week then there will be no more accidents in that week '.

4. Only experts can make the best possible use of statistics :

Statistics can be used by experts only. Those who are not aware of statistical methods cannot make the best possible use of available data.

5. For statistical analysis, uniformity and homogeneity of data is essential :

It is essential that data must be uniform and homogenous. Heterogeneous data are not comparable. It would be of no use to compare the heights of trees with heights of men because these data are of heterogeneous in nature.

Scope and applications of Statistics :

Statistics/statistical tools are applied in Governments, Business, Industry, Management, Commerce, Economics, Sociology, Biology, Agriculture, Medicine, Psychology and Insurance etc. . Some of them are discussed below :

Statistics in State :

Originally, Statistics was used in the state affairs. Later on it was used in all government departments, such as Finance, Education, Population study, Transport, Post and Telegraph, Commerce, Defense, Agriculture etc.. There are central government departments like C.S.O. (Central Statistical Organization), N.S.O. (National Sample survey Organization),

R.G.I.(Registrar General of India), I.I.P.S. (Indian Institute of Population Studies), which collects the data concerned to different fields. The data supplied by these organizations are useful in framing suitable policies in solving the problems of food shortage, poverty, unemployment etc..

Statistics in Business :

Statistical methods are used by businessmen in making decisions, estimation and comparison etc.. Also, in the business activities such as production, finance, sales, accounting, purchase, quality control, marketing etc., statistical methods are extensively used. Statistical methods are used in forecasting the future trends and tendencies.

Statistics in Economics :

Statistical methods are useful in understanding economic problems, formulating economic policies. Five-year plans, Savings, Taxation, Exports and Imports etc and their progress are evaluated by statistical methods. Econometrics is a branch of statistics, which deals with the application of statistical methods in the field of Economics.

Statistics in Science :

The physical and natural sciences like Astronomy, Engineering, Biometry (Biostatistics) and Pharmaceuticals, statistical methods are used. Also, in Agricultural science, statistics are widely used. Thus, the famous saying is that ' Sciences without statistics bear no fruit; Statistics without sciences has no root'.

Distrust of statistics :

The statement - 'There are three types of lies - lies, damned lies and statistics - wicked in the order of their naming '-mentions the lack of confidence in the subject among the few. This distrust may be the reaction due to the manipulations of the data, which may seem to be accurate, done by corrupt persons for some selfish motives.

Causes of distrust :

Distrust of statistics arises due to the direct or indirect influence of certain causes. These causes can be :

1. Figures can be easily believed.
2. Ignoring the limitations of statistics.
3. Misuse of the figures.
4. Inadequate samples.
5. Lack of subject knowledge.

Remedies to Remove Distrust :

1. Need of caution.
2. Statistical limitations should be taken into consideration.
3. Self-restraint. That is, self control in statistical fallacies.
4. Statistics must be used by experts.
5. Analytical study of data before its use.

Some basic concepts :

Units or Individuals :

The objects whose characteristics are studied in any statistical survey are called units or Individuals.

Population or Universe :

The totality of units under consideration is called **population or universe.** A population which contains finite (countable) number of units is called a **finite population.** A population which contains infinite (uncountable) number of units is called an **infinite population.**

Ex : 1. The population of fully grown trees in Lalbagh is a finite population.

2. The population of number of trees and plants in Lalbagh is an infinite population.

Sample:

The representative units of a population is called sample.

Units in a population will possess certain characteristics. Some characteristics are numerically measurable and some are not.

Numerically measurable characteristics are called **Quantitative Characteristics.**

The characteristics which are not numerically measurable are called **Qualitative Characteristics.**

In a population some characteristics remain same for all the units and some others vary from unit to unit

Variable :

A quantitative characteristic which varies from unit to unit is a **variable.**

Ex : Height/Weight of students, Number of students etc.

Attribute :

A qualitative characteristic which varies from unit to unit is an **attribute.**

Ex : Skin colour/Intelligence/Ability of students etc.

Discrete Variable :

A variable which assumes only specified values within a given range is a **discrete variable.**

Ex : Number of benches in a class room, Number of road accidents in a day etc.

Continuous Variable :

A variable which assumes all the possible values within a given range is a **continuous variable.**

Ex: Income of persons, Weight of students etc.

Numerically expressed data are called **Quantitative Data.**

The data which are not numerically expressed are called **Qualitative data.**

Measurement of quantitative data is very easy but, measurement of qualitative data requires different type of scale of measurement. The two scales of measurement are:

1. Nominal scale.
2. Ordinal Scale.

A number assigned for every unit for identification of different categories is called a **nominal scale.**

Ex : 1. Roll numbers assigned to students of a college.

2. Among male population married assigned number 0 and unmarried assigned 1.

Numbers assigned to observations that can be arranged in ascending or descending order is called an **ordinal scale.**

Ex: 1. Ranks given to students according to their performance.

2. In a shop three varieties of rice are assigned 1, 2 and 3 according to their quality.

Questions

1. Write a note on the origin and the development of science of Statistics
2. What are the two senses in which the term 'Statistics' is generally used ?
3. Distinguish between ' Statistics as numerical data' and ' Statistics as statistical methods' with examples.
4. State Prof. Horace Secrist's definition of Statistics.
5. Mention the characteristics of Statistics.
6. Briefly explain the characteristics of statistics.
7. State 'A.L.Bowley's' and 'Croxton and Cowden's' definition of Statistics.
8. Write a brief note on the main divisions of Statistics.
9. Mention any two branches of Applied Statistics.
10. What is Biometry ?
11. Define Econometrics.
12. What are the functions of Statistics ?
13. What are the limitations of Statistics ?
14. What is the role of statistics in Business, Commerce and Economics?
15. What are the causes of distrust of Statistics ?
16. What are the remedies to remove the distrust of Statistics ?
17. Explain the following terms with examples:
    a) Population and Sample.
    b) Qualitative and Quantitative characteristics.
    c) Qualitative and Quantitative data.
    d) Variable and Attribute.
    e) Discrete and Continuous variables.

* * * * *

Unit-II

ORGANIZATION OF DATA

Statistical Enquiry :

An enquiry means search for knowledge or truth. A Statistical enquiry means a search for knowledge through statistical methods. For example, the study of scoring pattern of students in an examination, study of population of India, study of wage structure of a factory workers, enquiry of the number of children per couple of a certain locality etc. Broadly there are two stages of statistical enquiry.

They are :

1. Planning and preparation.
2. Execution of the survey.

1. Planning and preparation :

For the success of any statistical enquiry a perfect planning and preparation is essential. In planning the following factors are to be carefully considered.

a) Object of an enquiry

b) Scope of the enquiry

c) Units used for collection and measurement

d) Sources of data

e) Method of collection of data

f) Framing a format

g) Accuracy level

h) Type of enquiry

2. Execution of the survey :

After a perfect planning, the next stage is execution. The various steps to be considered are as follows :

a) Setting a team of administrators
b) Designing of questionnaire
c) Selection and training of enumerators
d) Field work by enumerators and supervision
e) Follow up work in the case of non response
f) Analysis of collected data
g) Preparation of final report

Collection of data :

The first step of statistical enquiry is to collect facts and figures relating to a particular survey, whether the enquiry is of business , economics or social science. The **investigator** is the person who conducts the statistical enquiry. The person who collects the information for the investigator is called **Enumerator;** he/she should be a trained and an efficient statistician. The statistician collects and analyze the characteristics under study for further statistical analysis. The **respondents** (**informants**) are the persons from whom the information will be collected. **Collection** of data is the process of enumeration together with the proper recording of results. The success of an enquiry is based upon the proper collection of data.

In statistical survey the data are of two types :

I. Primary data
II. Secondary data

I. Primary data :

Data which are collected for the first time , directly from the field by the investigator is called **primary data**. They are original in nature. The data is also called the first hand data. The data are the raw materials of the enquiry. For example , if an individual or an office collects the data to study a particular problem directly from the field, it will be a primary data.

Methods of collection of primary data

For collection of primary data the investigator may choose any of the following methods depending upon object of an enquiry.

1. Direct personal observation.
2. Indirect oral interview.
3. Information through agencies.
4. Mailed questionnaires.
5. Schedules sent through enumerator.

1. Direct personal observation

Here, the investigator extracts the required information by personal observation of the units. The investigator must be a keen observer. He cross–examines the informant and records the necessary information. In this method the enumerator may collect biased information. However the method is less costly when the field of enquiry is small.

The method is adopted in the following cases ,

a) Where greater accuracy is needed.

b) Where the field of enquiry is not large.

c) Where confidential data are to be collected.

d) Where sufficient time is available.

Merits :

i. True and reliable data can be collected.

ii. In this method degree of accuracy can be high.

iii. Uniformity and homogeneity can be maintained.

Demerits :

i. Where the area is large this method is not suitable.

ii. This method is expensive and time consuming.

iii. An untrained investigator will not bring good result.

2. Indirect oral interview

If the informant is not ready to give information, the method of indirect oral interview can be followed. Under this method the investigator approaches the witnesses or third parties, who are in touch with the informant. Persons who are supposed to have knowledge about the problem under investigation are interrogated and the desired information is collected.

This method is more suitable , where the area to be studied is small. It is used when direct information cannot be obtained.

Merits :

i. This method is simple and convenient.

ii. This method is free from the bias and prejudice of the respondent.

Demerits :

i. Interview with an improper man will spoil the result.

ii. Witness may supply biased information according to their interests

3. Information through agencies

Under this method, local agents or correspondents will be appointed. They collect the information and transmit it to the investigator. They may apply any method according to the intensity of the situation. Agents who collect information from the informants are generally called **correspondents.**

This method is generally adopted in those cases where the information is to be obtained at regular intervals from a wide area. News papers are the examples of collection of information through agencies.

Merits :

i. This method is very cheap and economical.

ii. It is useful where information is needed regularly.

Demerits

i. The information may be biased.

ii. It is difficult to maintain the degree of accuracy and uniformity.

4. Mailed questionnaire

Questionnaire is a list of questions where the answers are filled by the informants and these answers are the required information for the investigation

The questionnaire is sent to the respondents, who are expected to write the answers in the space provided in the questionnaire.

The questionnaire will be sent to the informants through mail. A covering letter is also sent along with the questionnaire, requesting the informants to extend their full cooperation by giving the correct information and return the duly filled questionnaire with in a fixed time. This method is appropriate in cases where informants are spread over a wide area and are literates.

Merits :

i. This method is most economical.

ii. It saves manpower.

iii. It can be widely used, where the area of investigation is large.

Demerits :

i. This method cannot be used if the informants are illiterates.

ii. In this method many informants will not respond.

iii. In the case of non response, follow up work is essential.

5. Schedules sent through enumerators

It is a widely used method of collection of primary data. A number of enumerators are selected and trained for this purpose.

Here, the trained investigator collects data through enumerators. The enumerators contact the informants and collect the information which

is called schedule method. **Schedule** is a list of questions where the facts will be supplied by informants and recorded by enumerator.

Merits

i. This method is very useful where the informants are illiterates.

ii. In this method the rate of non response is less.

Demerits

i. In this method the training of enumerators is essential.

ii. This method is time consuming.

iii. Personal bias of the enumerators may lead to failure of enquiry.

Framing a Questionnaire

While drafting a questionnaire, the following general principles are to be considered

1. The number of questions should be as less as possible.
2. The questions should be simple to understand.
3. Questions should be arranged logically.
4. Answers to the questions should be short (Yes or No type).
5. As far as possible, questions regarding personal matters should be avoided
6. Any clarifications, if necessary regarding any of the questions should be supplied in the form of foot note.
7. Necessary instructions should be given to the informants.
8. Questions which require mathematical rigour should be avoided.
9. A questionnaire should be attractive.
10. The questions are so framed that the validity of the information supplied by the informants can be cross checked.
11. Questionnaire should also assure that the information supplied will be kept confidential and shall not be used for disadvantage.

Format of a questionnaire

DATE:.....

QUESTIONNAIRE REGARDING INTRODUCTION OF ENGLISH AT SCHOOL LEVEL

Please Note: The investigator wants to know about the introduction of English at school level. You are requested to co–operate kindly in the successful conduct of the survey by supplying information to the best possible degree of accuracy.

1. Name :_____
2. Age / Date of birth :_____ YY/MM/Date
3. Marital Status : a) Married ( ) b) Unmarried ( )
4. Sex : a) Male ( ) b) Female ( )
5. Qualification :
   a) Primary ( )
   b) Secondary ( )
   c) Higher Secondary ( )
   d) Pre University ( )
   e) Degree ( )
   f) Post graduation ( )
6. Which state do you belong?
   a) Karnataka ( ) b) Others (specify) _____
7. Mother tongue:
   a) Kannada ( ) b) Others (specify) _____
8. Can you,
   a) Speak Kannada : Yes ( ) No ( )
   b) Read Kannada : Yes ( ) No ( )
   c) Write Kannada : Yes ( ) No ( )
9. Which medium did you studied in high school?
   a) Kannada ( ) b) English ( )
10. Which medium of instruction do you prefer children of 1st standard?
    a) Kannada ( ) b) English ( )
11. If you prefer Kannada medium from 1st standard, from which standard do you prefer English medium?
    a) 5th standard ( ) b) 8th standard ( ) c) After SSLC ( )
12. Why you prefer (two sentences)_____.

II. Secondary data :

Secondary data is the data which has been already collected and analyzed by enumerators. Secondary data is also called second hand data.

Sources of secondary data

The various sources of secondary data can be divided into two categories.

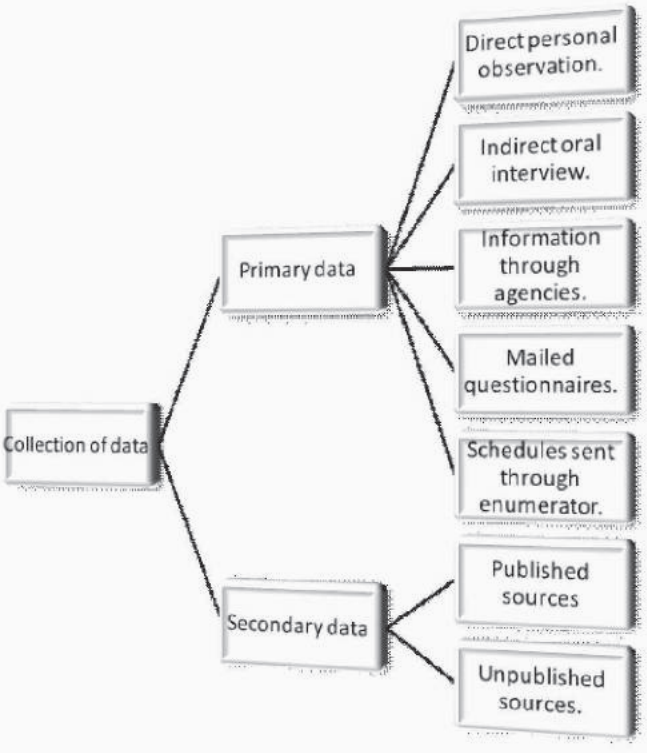
1. Published sources.
2. Unpublished sources.

1. Published Sources

a) Official publications of international bodies such as I.M.F. , I.B.R.D., U.N.O. etc.
b) Official publications of central and state governments.
c) Journals , Newspapers and periodicals.
d) Websites of various organizations on the internet.

2. Unpublished Sources

a) Records maintained at government offices , municipal offices etc.
b) Records maintained by research institutions , research scholars etc.



Census enumeration and sample survey

Information regarding an enquiry can be collected in two ways. They are :

1. Census enumeration.
2. Sample survey.

1. Census enumeration

A complete enumeration of each and every unit of the population is called **census enumeration.** For example, 'population census of India' which will be conducted once in every ten years , is a census enumeration.

Merits

i. The results are more accurate and reliable.

ii. The data are collected from each and every unit of the population.

iii. It provides the detailed study of all the units of the population.

iv. Census method is free from sampling errors.

Demerits

i. Non sampling errors are likely to be more in census enumeration.

ii. It requires more money , labour and time.

iii. It is not possible in some circumstances where the universe is vast .

iv. While procuring the data if the units are damaged, census enumeration is not suitable.

2. Sample survey

The representative units (a part or a portion) of a population is called **sample.** An enumeration based on a sample is called **sample survey.**

The theory of sampling has been developed recently, but it is not new. In our day to day life , we have been using sampling theory without

knowing about it For example , a housewife tests a small quantity of vegetable before purchase (i.e. she accepts or rejects a lot by sample survey).

Merits

i. This method requires less labour, less time and is economical.

ii. Sample survey is more scientific.

iii. This method is applied for those units which are destructive in nature.

iv. Sample survey is free from non sampling errors.

Demerits

i. Sample survey requires adoption of appropriate sampling method and appropriate methods of analysis.

ii. If the population is too heterogeneous in nature, the use of sampling procedure is impossible.

iii. Sampling errors are part and parcel of sample survey.

Pilot survey :

A survey conducted before any general survey is called a pilot survey.

Sampling :

The process of extracting sample from the population is called sampling.

Methods of sampling :

The following are some of the methods of sampling.

1. Simple random sampling.
2. Systematic sampling.
3. Stratified sampling.

1. Simple random sampling

It is a technique where the sample is drawn in such a way that each and every unit of the population will have an equal and independent chance of being included in the sample. Several methods have been

adopted for random selection of the sample. The most popular and simplest method is lottery method. In this method all the items of the population are numbered then select the required number of samples by drawing numbers.This lottery method cannot be used when the population is infinite. An alternative method is that of using the random number tables.

2. Systematic sampling

Systematic sampling is a procedure of drawing samples by selecting equally separated units in a numerically , alphabetically or geographically arranged population.If we want to select a sample of 'n' items from a population of N units, under thismethod , the population will be divided in to equally separated n groups of ' k ' items each.

Where, $k = \frac{N}{n}$

k = sample interval , N = Population size, n = Sample size.

One number will be selected between 1 and k by lottery method. Let it be 'a'.

Then the sample units are : a , a+k , a+2k,... , a+ (n – 1)k.

For example , if N = 1000, n = 100, then $k = \left(\frac{N}{n}\right) = \left(\frac{1000}{100}\right) = 10$

If a = 5 , then the sample units will be 5, 15, 25 ,..... 995.

3. Stratified sampling

When the population is heterogeneous or of different segments then this method is applied. First , the population is divided in to a number of sub groups called **strata**. Each stratum is homogeneous. From each stratum appropriate number of units are randomly selected. This method is called **Stratified sampling.**

Difference between census enumeration and sample survey

Census Enumeration	Sample survey
1. Enumeration of each and every unit of the population is called census enumeration .	1. Enumeration of a few representative units of the population is called sample survey .
2. Non sampling errors are likely to be more .	2. Sampling errors are more .
3. This method is not scientific .	3. This method is more scientific .
4. This method is impossible if the population is infinite .	4. This method is more suitable if the population is infinite .
5. This method requires more money , time and labour .	5. This method is economical .
6. In destructive cases this method cannot be used .	6. In destructive cases this method is only method which can be used .

Statistical Error (Sampling Errors) :

Statistical error is the difference between the estimated value and the actual value.

Causes of Errors :

1. Errors of origin :

   The error that occurs due to improper definition of statistical units , defective questionnaire and wrong method of enquiry is called the **Error of origin.**

2. **Error of inadequacy :**

   The error that occurs due to incomplete data or insufficient data is called **Error of inadequacy**.

3. **Error of manipulation :**

   The error that occurs at the time of analysis (clerical errors) is called **Error of manipulation.**

Biased and unbiased errors :

There are two classes of sampling errors. They are **biased errors** and **unbiased errors.**

Biased errors :

The errors that occur with the notice of the investigator are called **biased errors.** These errors are prejudiced errors.

Unbiased errors :

The errors that occur without the notice of the investigator are called **unbiased errors.** These errors occur due to chance causes (which cannot be controlled).

Measurement of errors

There are two types of measurements :

1. Absolute error
2. Relative error

1. Absolute Error : is the arithmetic difference of actual value and the estimated value.

Absolute Error = Actual value – Estimated value

A.E = a – e

2. Relative Error: is the ratio of absolute error and the estimated value.

Note : As the sample size increases, the sampling error decreases.

*****

Questions

1. What is meant by a statistical enquiry ?
2. Mention the two stages of statistical enquiry.
3. Mention the points that are to be considered in planning an enquiry.
4. Mention the points that are to be considered in execution of the survey.
5. Write a brief note on statistical enquiry and its stages.
6. What is primary data ?
7. Mention the methods of collection of primary data.
8. Explain the methods of collection of primary data with relative merits and demerits ?
9. Define secondary data.
10. Mention the sources of secondary data
11. What is a questionnaire ?
12. What are the guidelines for the construction of a questionnaire ?
13. Define sample.
14. Define sampling.
15. Define sample survey.
16. Define census enumeration.
17. Mention the merits and demerits of census enumeration.
18. Mention the merits and demerits of sample survey.
19. Distinguish between census enumeration and sample survey.
20. What is a strata ?
21. Briefly explain the three methods of sampling.
22. What is meant by statistical error ?
23. What are the causes for the occurrence of statistical error ?
24. Distinguish between biased and unbiased errors.
25. Mention the methods of measurement of errors.
26. What is meant by absolute error ? (Define absolute error)
27. What is meant by relative error ? (Define relative error)

Unit-III

CLASSIFICATION AND TABULATION

Introduction :

In the previous chapters we discussed about the different methods of data collection. Usually the data collected is voluminous and complex in nature. It cannot be understood easily and also not suitable for statistical analysis. Such data needs to be arranged, grouped and classified according to similar characteristics possessed by the items of the data. The classification and tabulation are the process of organizing the data.

Classification of data :

According to Prof. Horace Secrist ' classification is the process of arranging the data in to sequences and groups according to their common characteristics or separating them into different but related part '.

Definition :

"**Classification** is the process of arranging the data in to groups or classes according to common characteristics possessed by the items of the data ". That is, sorting of different items (observations) on the basis of similar characteristics possessed by the items constituting in the data.

Objectives :

The main objectives of the classification are :

1. **To reduce the size of the data :** Classification reduces the vast and voluminous data in to handy size.
2. **To bring the similarities together :** Classification simplifies the data by arranging similarities and dissimilarities contained in the data.
3. **To facilitate comparison** : Classification helps in comparison of data characteristics.
4. **To give importance to significant features of the data :** Unnecessary and irrelevant details in the data are eliminated at the time of classification and make the data more precise and compact.

5. **To enable further statistical analysis** : Classification of the data enables further statistical analysis and hence to draw meaningful conclusions.

Types of classification :

Generally classification of the data is made on the following basis :

1. Chronological (temporal) classification.
2. Geographical (spatial) classification.
3. Qualitative classification.
4. Quantitative classification.

1. Chronological classification :

Chronology means orderly arrangement of statistical data according to time, date of happening of an event. If the data are arranged over a different period of time, then the type of classification is called **chronological classification.**

Example : Population of India from 1931 to 2011.

Years	1931	1941	1951	1961	1971	1981	1991	2001	2011
Population (in millions)	279	319	361	439	548	683	846	1002	1210

2. Geographical classification :

If the data is classified on the basis of geographical or location or area-wise (such as cities, districts, states etc.), it is called **geographical (spatial) classification.**

Example 1. Production of sugar in some states of India :

Name of the State	Uttar Pradesh	Bihar	Tamil Nadu	Maha rashtra	Other states	Total
Sugar production (In million tons)	55	20	25	15	12	127

Example 2. According to 2011 census, population of some states of India is:

Sl. No.	Name of the state	Population (in crores)	Rank in 2011
1	Uttar Pradesh	19.95	1
2	Maharashtra	11.23	2
3	Bihar	10.38	3
4	West Bengal	9.13	4
5	Andhra Pradesh	8.47	5
6	Madhya Pradesh	7.26	6
7	Tamil Nadu	6.86	7
8	Rajasthan	6.11	8
9	Karnataka	6.04	9
10	Gujarat	4.19	10

3. Qualitative classification :

If the data is classified on the basis of the qualitative characteristics or attributes (such as sex, literacy, employment and religion etc.), is called **qualitative classification.** Classification of units on the basis of a single characteristic is **simple or one way classification.**

Example 1.

The residents of a certain locality can be classified according to their occupations as:

Occupations	Number of residents
Office assistants	
Business men	
Teachers	
Bank Employees	
Total	

The classification of units on the basis of single characteristic (attribute) into two classes is called **Dichotomous** classification. That is the data are classified in to two classes as one group possessing (presence) and other not possessing (absence) the quality.

Example 2.

The members of a club can be classified on the basis of sex as:

Members		
Male	Female	Total
80	20	50
45	5	50

Similarly, the classification of units on the basis of attributes in to two or more characteristics is called **Manifold classification.** These are also called as **contingency tables.**

Example 3.

Population of a village can be classified according to Sex, Literacy, Employment and Marital status as follows :

Population	Men	Literates	Employed	Married
				Single
			Unemployed	Married
				Single
		Illiterates	Employed	Married
				Single
			Unemployed	Married
				Single
	women	Literates	Employed	Married
				Single
			Unemployed	Married
				Single
		Illiterates	Employed	Married
				Single
			Unemployed	Married
				Single

4. Quantitative classification :

The classification of the units on the basis of quantitative characteristics or variable (such as height, weight, wages, age in years, number

of children, number of phone calls, number of births, number of deaths etc.), is called **quantitative classification**.

Example : 1.

The group of students may be classified according to their heights as:

Height (cm)	140-150	150-160	160-170	170-180	Total
No. students	6	24	18	2	50

Example : 2.

Number of Apples per box are given below:

No. of Apples per box	5	6	7	8	9	10	Total
No. of boxes	5	8	13	10	6	3	45

The above two types of classifications are frequency distributions.

Thus, a **frequency distribution** is a systematic presentation of the values taken by a variable along with their frequencies. **Frequency** refers to the number of times an observation is repeated. The number of observations corresponding to a particular class is known as the **class Frequency.** Class frequency is a positive integer including zero.

From the above examples, we can explain the following terms :

Number of apples per box, number of children, height (cm), weight (kg) are variables and number of students, number of boxes are frequencies.

While framing a frequency distribution, if class intervals are not considered, is called **discrete frequency distribution.**

Example :

The number of families according to number of children.

No. of children	0	1	2	3	4	5	6	Total
No. of families	10	40	80	100	250	150	50	680

While framing a frequency distribution, if class intervals are considered, is called **continuous frequency distribution.**

Example :

The weight (kg) of persons

Weight (kg)	30-40	40-50	50-60	60-70	70-80	Total
No. of persons	10	15	40	45	20	130

Formation of Discrete frequency distribution :

Here, three columns are formed-Variable, Tally bars, Frequency. In the first column, values of given variable are written without repetition in an order. For each value a tally/stroke is marked against that value in the second column. In this way tally scores are marked for all values. For easy counting the tallies are put as a group of 5 (𝍸). Finally count the number of tally bars corresponding to each value of the variable in third column. It is known as frequency. The total frequency (N) is equal to the total number of observations.

Example : 1.

In a survey of 40 families in a certain locality, the number of children per family was recorded and the following data were obtained.

1, 0, 3, 2, 1, 5, 6, 2, 2, 1, 0, 3, 4, 2, 1, 6, 3, 2, 1, 5, 3, 3, 2, 4, 2, 2, 3, 0, 2, 1, 4, 5, 3, 3, 4, 4, 1, 2, 4, 5.

Represent the data in the form of a discrete frequency distribution.

Solution :

Frequency distribution of the number of children.

Number of Children(x)	Tally Marks	Frequency(f)
0	|||	3
1	𝍸 ||	7
2	𝍸 𝍸	10
3	𝍸 |||	8
4	𝍸 ||	6
5	||||	4
6	||	2
	Total	40

Example 2 :

Number of teaching staff working in 20 different colleges was recorded as below:

15 12 18 10 15 12 20 25 18 10 15 12 15 20 25 15 18 20 15 18

Prepare a frequency table.

Solution :

The difference between the highest and lowest value in the data is very less, and number of repeated values are also less, so use discrete frequency table. third column. It is known as frequency. The total frequency (N) is equal to the total number of observations.

Discrete frequency table representing the number of teaching staff in different colleges.

No. of teaching staff(X)	Tally Marks	No. of colleges(f)
10	||	2
12	|||	3
15	~~||||~~ |	6
18	||||	4
20	|||	3
25	||	2
Total		20

Formation of Continuous frequency distribution :

Suitable class intervals are formed on the basis of the magnitude of the data. For each value a tally mark is marked against the class in which it falls. This process is continued until all the values are exhausted. The tallies of each class are counted and written as frequency of that class.

To construct a continuous frequency distribution table, it essential to know the following factors :

Range :

It is the difference between highest and lowest value in the data.

i.e., Range = H.V. - L.V.

Class :

The sub range is called class.

Class limits :

The lowest and the highest values which are taken to define the boundaries of a class are **class limits.** The lowest value is called **lower limit (L.L)** and the highest value is **upper limit (U.L).**

Example: (30-40), (40-50)...are class limits.

The lowest value of the class is called **lower limit (L.L.)** as 30, 40... are lower limits and the highest value of the class is **upper limit (U.L.) 40, 50**...are upper limits.

Inclusive class :

In a class, if lower as well as upper limits are included in the same class, such a class is called **Inclusive class**. Here, upper limit of a class is not equal to the lower limit of the next class. Example : 0 - 9, 10-19,... are inclusive classes.

Exclusive class :

In a class, if the lower limit is included in the same class and upper limit is excluded from that class but included in the next class, such a class is called **Exclusive class**. Here, upper limit of a class is equal to lower limit of the next class.

Example : 30-40, 40-50, ... are exclusive classes.

Correction factor :

It is half of the difference between lower limit of a class and upper limit of the preceding class. Thus,

$$\text{Correction factor (C.F.)} = \frac{\text{Lower limit of a class} - \text{Upper limit of the precending class}}{2}$$

To get exclusive class intervals from inclusive class intervals, add C. F. to all upper limits and subtract C. F. from all lower limits.

Example: Convert the flowing inclusive class intervals to exclusive class intervals.

C - I	10 – 19	20 – 29	30 – 39	40 - 49

$$\text{C.F.} = \frac{20-19}{2} = \frac{1}{2} = 0.5$$

By subtracting it from all lower limits we get lower limits as 9.5, 19.5, 29.5, 39.5 and adding it to all upper limits we get upper limits as 19.5, 29.5, 39.5, 49.5

Therefore the exclusive class intervals are: 9.5-19.5, 19.5-29.5, 29.5-39.5 and 39.5-49.5

Open-end classes :

In a class, if the lower or upper limit of the class is not specified such a class is called open-end class. For example : less than (below) or more than(above) a particular class limit.

The frequency distribution based on open-end classes is called **open-end frequency distribution.**

Example : Open-end frequency distribution.

Class Interval	Frequency
Less than 20	8
20-30	15
30-40	23
40-50	12
50-60	9
More than 60	3

Mid-point (class mark) :

The central value of a class is called **mid-point or class mark.** It is the average of class limits.

$$\text{i.e., } m \text{ or } x = \frac{\text{Lower Limit} + \text{Upper Limit}}{2}$$

Example : Mid-point of the class [10-20) is $m = \frac{LL+UL}{2} = \frac{10+20}{2} = 15$

Width (size) of the class :

The difference between the upper and lower limits of a class is called **width of the class**. It is denoted by **c** or **i**.

For example, the width of the class interval (30-40) is 40-30 = 10.

Note: In case of inclusive classes, width of the class intervals has to be calculated, after converting the inclusive classes in to exclusive form.

Number of classes :

The number of classes can be obtained by using the Prof. Sturge's Rule

Number of classes (K) = 1 + 3.322 log N; N: Number of observations.

The width of the class can also be obtained by :

$$\text{Width of the class} = c = \frac{\text{Range}}{\text{Number of classes(K)}}$$

Cumulative frequency :

The added up frequencies are called cumulative frequencies.

There are two types of cumulative frequencies : (i) less than type (ii) more than type.

The number of observations (frequencies) below a certain limit is **less than cumulative frequency** (L.C.F). The frequency distribution formed for less than cumulative frequencies against upper class limits, is, Less than Cumulative Frequency Distribution.

The number of frequencies above a certain limit is **more than cumulative frequency**. The frequency distribution formed for more than cumulative frequencies against lower class limits is more than Cumulative Frequency Distribution.

Example:

Frequency distribution		Less than cumulative frequency distribution		More than cumulative frequency distribution	
Weight (kg)	Number of persons	Weight (kg)	Number of persons	Weight (kg)	Number of persons
30 - 40	10	Less than 40	10	More than 30	60
40 - 50	15	Less than 50	25	More than 40	50
50 - 60	20	Less than 60	45	More than 50	35
60 - 70	15	Less than 70	60	More than 60	15

Frequency density:

The frequency per unit of class interval is the **frequency density** (f/c). Or, it is the ratio of the class frequency to the width of that class interval.

$$\text{i.e., Frequency density} = \frac{\text{Frequency of the class}}{\text{width of the class}}$$

It is used to compare the concentration of the frequencies of different classes for a given frequency distribution.

Frequency density can be calculated as :

Class Interval	Frequency (f)	Width of the class (c)	Frequency density (f/c)
0-10	10	10	$\frac{10}{10} = 1$
10-30	15	20	$\frac{15}{20} = 0.75$
30-50	40	20	2
50-60	45	10	4.5
60-65	20	5	4

Relative frequency:

It is the ratio of frequency of the value of the variable to the total frequency.

$$\text{i.e., Relative frequency (R f)} = \frac{\text{Frequency of the value of the variable}}{\text{Total frequency } (N)}$$

The relative frequencies are calculated as:

Observation (No. of heads)	Frequency (f)	Relative frequency(f/N)
2	1	$\frac{1}{4}$ = 0.25
1	2	$\frac{2}{4}$ = 0.50
0	1	$\frac{1}{4}$ = 0.25
Total	N = 4	1

Principles (rules) of classification :

There are no strict rules for classification. However, the following general guidelines may be considered while classifying the data -

1. The number of classes should generally between 4 and 15.
2. Exclusive classes should be formed for better continuity between the class intervals.
3. The width of the classes should be usually kept constant throughout the distribution or may be different for the proper meaning of subject matter of the data.
4. Avoid open-end classes, because the mid values of such class cannot be easily computed.
5. The classes should be arranged in ascending or descending order.
6. The lower limit of the first class should be multiple of 5.

Example :1.

Let us consider the weights (in kg) of 50 students of a college.

42	62	46	54	41	37	54	44	32	45
47	50	58	49	51	42	46	37	42	39
54	39	51	58	47	64	43	48	49	48
49	61	41	40	58	49	59	57	37	34
56	38	45	52	46	40	63	41	51	41

Solution :

Here, Range = H. V. – L. V. = 64 – 32 = 32

The number classes as per Sturge's rule are obtained as follows:

Number of classes (K) $= 1 + 3.322 \log N$

$= 1 + 3.322 \log (50)$

$= 6.64 = 7$ classes (Approx.)

Size of classes $c = \frac{\text{Range}}{\text{Number of class}} = \frac{32}{7} = 5$ (Approx.)

The size of each class is 5. Thus, the required continuous frequency distribution is prepared as :

Table showing the frequency distribution of weights of students.

Class Interval (weight in kg)	Tally marks	Frequency (f) (Number of students)
30-35	||	2
35-40	𝍸 |	6
40-45	𝍸 𝍸	12
45-50	𝍸 𝍸	14
50-55	𝍸 |	6
55-60	𝍸 |	6
60-65	|||	4
	Total	50

Example 2 :

Following are the marks obtained by the students in a certain test. Prepare a frequency distribution with an interval of 10 marks each using inclusive class intervals.

37 49 54 51 37 15 12 33 23 25 18 35 33 42 45

55 69 63 46 29 18 37 46 59 29 35 27 45 47 65

Solution :

In the above data lowest value = 12 and highest value = 69.

Here, the size of each class as 10 (given). Therefore the classes are [10-19],[20-29]...[60-69].

Thus, the required continuous frequency distribution is prepared as:

Frequency distribution of Marks of students

Marks (Cl)	Tally marks	Number of students (f)
10-19	||||	4
20-29	~~||||~~	5
30-39	~~||||~~ ||	7
40-49	~~||||~~ ||	7
50-59	||||	6
60-69	|||	3
	Total	30

Formation of Bi-variate frequency distribution :

In many situations simultaneous study of two variables becomes necessary. For example, we want to classify data relating to the height and weights of a group of individuals, income and expenditure of a group of individuals, ages of husbands and wives, number of children and ages of mothers etc..

A frequency distribution formed on the basis of two related variables is called **bi-variate frequency distribution.**

Method of construction:

Step : 1. By adopting rules of classification appropriate class intervals are formed for both variables (one may be discrete and other may be continuous). If the one variable (say X) is grouped into ' m ' classes and the other variable (say Y) is grouped into ' n ' classes, then the bi-variate table will consist of ' m x n ' cells.

Step : 2. Represent the class intervals corresponding to one variable in first column of the table and class intervals corresponding to another variable in first row of the table.

Step : 3. By going through the different pairs of the values of (X, Y) variables, put a tally mark in the appropriate cell (row corresponding to one variable and column corresponding to another variable). Using

tally marks frequencies of each cell are calculated. The row wise and column wise total of these frequencies is made to get marginal frequencies. The total of marginal frequencies is equal to the total number of paired observations.

In a bivariate frequency distribution, the frequency distribution of only one of the variables is considered, it is **marginal frequency distribution.**

Example 1. Given below are the Ages (years) of mothers and the number of children. Prepare a bi-variate frequency table

Sl. No.	Ages of Mothers (Years)	Number of children		Sl. No.	Ages of Mothers (Years)	Number of children
1	28	1		11	27	0
2	37	0		12	39	3
3	42	3		13	23	4
4	25	2		14	33	2
5	29	1		15	36	1
6	47	5		16	32	6
7	37	6		17	22	3
8	35	2		18	29	2
9	23	2		19	38	1
10	41	1		20	48	5

Solution :

Bi-variate frequency table showing Ages (years) of mothers and Numberof children.

No. of children → / Ages (years) ↓	0	1	2	3	4	5	6	Total
20-25			| (1)	| (1)				2
25-30	| (1)	||	||		| (1)			6
30-35			| (1)				| (1)	2
35-40	| (1)	||	| (1)	| (1)			| (1)	6
40-45		| (1)		| (1)				2
45-50		| (1)				| (1)		2
Total	2	6	5	3	1	1	2	N=20

From the above table marginal frequency table of Ages of mothers can be prepared as:

Age (in years)	20-25	25-30	30-35	35-40	40-45	45-50	Total
No. of mothers	2	6	2	6	2	2	N=20

Similarly, marginal frequency table of Number of children can be prepared as:

No. of children	0	1	2	3	4	5	6	Total
No. of mothers	2	6	5	3	1	1	2	20

Example 2. The data given below relate to the Height and Weight of 20 persons. Construct a Bi-variate frequency table with class interval for height as 115-125,125-135,...and weight as 62-64, 64-66... write down the marginal distribution of X and Y.

Sl.No.	1	2	3	4	5	6	7	8	9	10
Height (cm)	170	135	136	137	148	121	117	128	143	129
Weight (kg)	70	65	65	64	69	63	65	70	71	62
Sl.No.	11	12	13	14	15	16	17	18	19	20
Height (cm)	163	139	122	134	140	132	120	148	129	152
Weight (kg)	70	67	63	68	67	69	65	68	67	67

Solution :

Bivariate frequency table showing height and weight of persons

Weights(kg) → Height(cm) ↓	62-64	64-66	66-68	68-70	70-72	Total
115-125	|| (2)	|| (2)				4
125-135	| (1)		| (1)	|| (2)	| (1)	5
135-145		||| (3)	|| (2)		| (1)	6
145-155			| (1)	|| (2)		3
155-165					| (1)	1
165-175					| (1)	1
Total	3	5	4	4	4	N=20

From the above table marginal frequency table of height is:

Height (in cm)	115 - 125	125 - 135	135 - 145	145 - 155	155 - 165	165 - 175	Total
No. of person	4	5	6	3	1	1	N=20

Similarly, from the above table marginal frequency table of weight is:

Weight (in kg)	62-64	64-66	66-68	68-70	70-72	Total
No. of persons	3	5	4	4	4	20

Tabulation

The collected and the classified data can also be presented in the form of a table. A tabular presentation provides better understanding, comparison, and analysis to draw valid conclusions. Here, the data will be arranged in the form of table containing rows and columns. Thus, a statistical table makes it possible for the investigator to present a huge mass of data in a detailed and orderly form. Classification and Tabulation are not two distinct processes. Before tabulation, data are classified and then displayed under different columns and rows of a table.

Tabulation is a process of systematic arrangement of the classified data in rows and columns, in the form of a table.

Objectives :

The Statistical arrangement of the data in a tabular form serves the following objectives :

1. It simplifies the complex data.
2. To facilitate comparison.
3. To give an identity to the data.
4. To reveal the trend and tendencies of the data.
5. For diagrammatic and graphical representation a tabulated data is essential.

6. It facilitates computation of various statistical measures like averages, dispersion, correlation etc..

Parts of a Table :

The various parts of the table are depending on the nature and scope of the data.

1. Table Number :

A number should be given to each and every table, in order to distinguish and also for easy reference. Usually the table number is written at the right hand corner at the top of the table.

2. Title of the table :

A title should be given to the table. The title should indicate in short the contents of the table, which will be written at the top in bold letters.

3. Head note or sub-title :

It is a brief explanatory note or statement given just below the heading of the table put in a bracket. The statistical units of measurements, such as in '000s, Rs., Million tons, crores, Kg, etc..

4. Captions :

Column headings are called captions. They explain what the column represents. Captions are always written in one or two words on the top of each column.

5. Stubs :

Row headings are called Stubs. They explain what the row represents. Stubs are usually written in one or two words at the left extreme side of each row.

6. Body of the table :

The body of the table contains numerical data. The table should be well balanced. i.e., length and breadth should have appropriate ratio in accordance to the size of the paper.

7. Foot note :

It is generally written below the table explaining in brief and precisely, clarifying anything special referring to the table such as conclusions, remarks, inferences, omissions, additions, abbreviations etc..

8. Sources :

Below foot note or below the table, source of the data may be mentioned for verification to the reader.

General format of a table

Table No.

Title

(Sub-title / Head note:)

Stub headings (Row headings)	Captions (columns headings)				
Stub Entries					
			Body of the table		

Foot note:

Source:

Rules/Requisites of tabulation :

A good statistical table is not only grouping of data into columns and rows but also summarizes the total information in an easy way, with minimum possible space. Though, there is no hard and fast rule for preparation of a table yet a few general points should be kept in mind which are given below :

1. Size :

The size of the table should be according to the size of the paper with more rows than columns. Exchange of the data can be done by altering the column and rows. A sufficient space should be provided in a particular cell to enter any new or to alter any affected figures.

2. Logical order :

The stubs and captions should be arranged logically, according to the alphabetical, chronological, geographical order. Items are to be arranged according to the description of stubs and captions.

3. The table should not be **overloaded** with number of characteristics, rather can be prepare another table. i.e., In case it is not possible to accommodate all the information in a single table, it is better to have two or more related tables.
4. The table should be **complete** in all respects.
5. No cell should be left blank. Either, put (—) dash or write N.A. (Not Available).
6. **Miscellaneous column** can be provided for the presentation of ratios, percentages etc., of the data.
7. **Ditt**o (") marks should not be used, as they may confuse with the row or column.
8. **Footnote :** It is used to explain omissions, any remarks regarding the numerical entries of the data in the table.
9. **Source :** The data obtained from publications, organizations, Journals etc. should be mentioned.

Type of Tables :

Tables can be classified according to their purpose, nature of data or number of characteristics as follows :

1. Simple or one-way table
2. Two way table
3. Manifold table

1. Simple or one-way Table:

If only one characteristic is represented in a table, it is called a simple or one-way table. For example, the employees of an organization can be presented according to their ages or marital status or category of work or any other.

Example 1. The residents of a certain locality can be classified according to their occupations

Table-1: The blank table given below represents the number of residents with different occupations in a locality.

Occupations	Number of residents
Office assistants	
Business men	
Teachers	
Bank Employees	
Total	

2. Two-way table :

If two characteristics are represented in a table, it is called two-way table. Here, either stub or caption or both is divided into two co-ordinate parts.

Example 2. The residents of a certain locality classified according to their occupations and sex is :

Table 2: The blank table given below represents the number of residents with different occupation and according to their sex in a locality

Occupation	Number of residents		
	Men	Women	Total
Office assistants			
Business men			
Teachers			
Bank Employees			
Total			

Example 3. Prepare a blank table to represent the students in a college according to (i) Faculty : Arts, Commerce, Science and (ii) Sex

Solution: The blank table given below represents the number of students in different faculties and sex

Faculty	Number of students		
	Boys	Girls	Total
Arts			
Commerce			
Science			
Total			

3. Manifold Table :

If more than two characteristics are represented in a table, it is called manifold table. Here stubs or captions are divided into two or more parts.

Example 4. Prepare a blank table showing the particulars relating to the residents of a certain locality according to :

(i) Occupation : office assistants, business men, teachers, bank employees.

(ii) Sex : men and women.

(iii) Marital status : married, single.

Solution: The blank table represents the residents of a locality according to their occupation, sex and marital status.

Occupation	Number of residents								
	Men			Women			Total		
	Married	Single	Total	Marrie	Single	Total	Marrie	Single	Total
Office									
Business men									
Teachers									
Bank									
Total									

Foot note:
Source:

Example 5. Prepare a blank table showing the particulars relating to the students of a college classified according to :

(i) Faculty : Commerce, Science.

(ii) Sex : boys and girls.

(iii) Age group : Age-below 18 years, 18 - 20 years, above 20 years.

Solution: The blank table represents the number of students according to faculty, sex and age.

Age group (years)	Faculty								
	Commerce			Science			Total		
	Boys	Girls	Total	Boys	Girls	Total	Boys	girls	Total
Below 18									
18 - 20									
20 and above									
Total									

Foot note:

Source:

Example 6. In a sample study about the food habits of residents of a Agrahar village, the following data were observed. 55% of the residents were males; 85% were vegetarians ; only 12% were non-vegetarian females. Tabulate the above information.

Solution : Table showing the sex-wise distribution of vegetarians of Agrahar village (figures in percentages).

Food habit	Male	Female	Total
Vegetarians	52*	33*	85
Non-Vegetarian	3*	12	15*
Total	55	45*	100

Foot Note:*45-12=33; 85-33=52; 55-52=3; 12+3=15

Example 7 :

In a state there were 30 lakh people. Out of these, 10 lakh people live in urban areas and the rest in rural areas. In urban areas there were 7 lakh male people, out of which 2.5 lakh are illiterate. In urban areas 2 lakh ladies were illiterates. In rural areas there were 15 lakh male people out of which 5 lakh were literate, in rural areas illiterate ladies were 3 lakh. Tabulate the above information.

Solution: Table representing the residents of a state in urban and rural area according to sex and literacy (Figures in Lakhs).

Area of Living	Men			Women			Total		
	Literates	Illiterates	Total	Literates	Illiterates	Total	Literates	Illiterates	Total
Urban	2.5	4.5*	7	1*	2	3*	3.5*	6.5*	10
Rural	5	10*	15	2*	3	5*	7*	13*	20*
Total	7.5*	14.5*	22*	3*	5*	8*	10.5*	19.5*	30

Foot Note: *computed figures, remaining are given figures.

Source: Imaginary.

Example 8 :

In a sample study about food habits in two towns, the data was obtained:

Town A : 50 % persons were males
30 % were non-vegetarians
18 % male non-vegetarians

Town B : 45% persons were males
25 % were non-vegetarians
16 % were male non-vegetarians;

tabulate the above data.

Solution :Table-8: Table showing the sex-wise distribution of vegetarians in two villages

Food Habit	Village A			Village B		
	Males	Females	Total	Males	Females	Total
Vegetarian	32*	38*	70*	29*	46*	75*
Non-vegetarians	18	12*	30	16	9*	25
Total	50	50*	100	45	55*	100

Foot Note: * computed figure.

Questions

1. What is classification of the data ?
2. What are the objectives of classification ?
3. Explain geographical classification and chronological classification of data with examples.
4. Explain qualitative and quantitative classification with examples.
5. Give the formula of Sturge's to find the number of classes. Or Give the formula used to determine the number of classes.
6. Define the following terms :
   i. Frequency, class frequency
   ii. Class limits
   iii. Width of the class
   iv. Class mid point
   v. Class interval-inclusive, exclusive and open-end
   vi. Correction factor
   vii. Cumulative frequency- less than and more than
   viii. Frequency density
   ix. Relative frequency
   x. Frequency distribution - Discrete, Continuous, Univariate, Bi-variate, Marginal.

7. For what purpose is correction factor used, in frequency distribution ?
8. What are the guidelines of classification ?
9. What is the tabulation of the data ?
10. What are the parts of a table ?
11. What is the purpose of table number in tabulation ?
12. What are captions and stubs of a table ?
13. What is head note of table ?
14. What is indicated by source note of a table ?
15. Give a general format of a table.
16. What are the requisites of a good table ? or What are the General rules to the tabulation ?
17. Elucidate the difference between classification and tabulation.
18. The number of doctors serving in forty different hospitals were recorded as below :

    6 7 5 7 6 3 9 8 6 7 5 7 6 8 5 8 5 9 5 6
    5 9 6 6 4 4 7 5 5 8 5 3 3 8 4 4 3 4 4 3

    Prepare a discrete frequency table.
19. Following is the data regarding the number of meteors observed in 30 different days to an observatory : 7, 4, 2, 5, 9, 7, 4, 3, 8, 6, 9, 10, 3, 3, 5, 9, 5, 7, 6, 3, 8, 4, 3, 7, 9, 8, 10, 6, 8, 7. Tabulate above observation.
20. From the following paragraph prepare a discrete frequency table with the number of letters present in the words.

    Success in the examination confers no right to appointment unless government is satisfied, after such enquiry as may be deemed necessary that the candidate is suitable for appointment to the public service.

21. From the following paragraph prepare a frequency table of the number of the letters in a word, ignoring the punctuation marks.

"By statistics we mean aggregates of facts affected to a marked extent by multiplicity of causes numerically expressed, enumerated or estimated according to a reasonable standard of accuracy, collected in a systematic manner for a pre-determined purpose and placed in relation to each other.

22. A review of the first 30 pages of a statistics book reveals the following printing mistakes

0	1	3	3	2	5	6	0	1	0	4	1	1	0	2
3	2	5	0	4	2	2	3	3	4	6	1	4	3	2

Prepare a frequency distribuion of printing mistakes.

23. Following are the ages (in years) of school children of a lower primary class at Sandeep International school.

12	5	13	12	10	11	7	9	6	10	9	13	5	10	7
14	6	11	13	7	9	8	11	10	8	12	13	9	6	14

Prepare a frequency table using inclusive class interval of width 3years.

24. Following is the data regarding the I.Q of 30 children of a school.

106	118	112	120	103	105	99	139	65	113	108	116	93	100	120
100	108	100	112	110	109	117	98	103	99	110	95	101	105	98

Prepare a frequency distribution with class intervals: less than or equal to 90, 90-110, 110-120, more than or equal to 120.

25. From the following data of the hourly wages (in rupees) of workers employed in a certain factory, construct a frequency table with classes 40-49, 50-59 and so on.

76	63	83	75	61	41	115	82	60	40	74	42
78	95	56	77	78	65	67	50	84	76	100	53
100	81	59	73	54	79	79	80	104	69	68	79
72	80	70	69	64	42	76	84	90	77	49	66
73	71	94	78	86	51	96	103	52	79	50	72

26. Given below are the daily wages in rupees of 36 workers in a factory manufacturing plastic products.

100 115 120 125 92 140 150 162 189
165 200 220 250 240 300 320 270 280
400 382 288 235 225 312 270 250 242
344 248 188 220 240 212 224 325 425

Form a frequency distribution taking lowest class-interval as 90-140 with the magnitude of 50.

27. Below are given the marks obtained by a batch of 22 students in Mathematics and Statistics:

Roll No.	1	2	3	4	5	6	7	8	9	10	11
Marks in Mathematics	53	54	52	32	30	60	47	46	35	28	25
Marks in Statistics	58	55	25	32	20	85	44	80	33	72	10
Roll No.(continued)	12	13	14	15	16	17	18	19	20	21	22
Marks in Mathematics	42	33	48	72	51	45	33	65	29	35	40
Marks in Statistics	42	15	46	50	64	39	33	30	36	40	52

Prepare a bivariate frequency table.

28. Below are the ages of husbands and wives prepare a bivariate frequency distribution with suitable width :

Sl. No.	Age of Husband (in years)	Age of Wife (in years)
1	28	24
2	42	37
3	29	25
4	47	37
5	35	23
6	41	39
7	33	23
8	36	32
9	29	24
10	38	38

Sl. No.	Age of Husband (in years)	Age of Wife (in years)
11	27	23
12	40	30
13	26	25
14	41	35
15	25	21
16	38	34
17	31	20
18	35	29
19	29	25
20	47	34

29. Prepare a blank table to show the distribution of population of a village according to
    i. Sex : men, women
    ii. Age group : below 18 years, 18-30 years and 30years & above
    iii. Marital status : married, single.
30. Draw a blank table to show the distribution of employees of a factory according to-
    i. Sex : men, women
    ii. Religion : Hindu, Muslim, Christian
    iii. Wages :below Rs.5000, Rs.5000-10000 and Rs.10000 & above.
31. Prepare a blank table to show the population of a town according to-
    i. Sex : men, women
    ii. Age group in years : (0-25), (25-50), 50 & above
    iii. Periods : 2008, 2009, 2010.
32. Draw a blank table to show the students of a college according to-
    i. Faculty : arts, commerce & science
    ii. Sex : boys, girls
    iii. Class : I puc, II puc
33. Prepare a blank table to show the employees of a commercial bank according to
    i. Sex : men, women
    ii. Cadre : officers, cashiers, clerks
    iii. Age group (in years) : below 30, 30-40 and 40 & above
34. In a sample study regarding smoking habit in a town, the following data were obtained :

    Men population = 58%

    Smokers = 22%

    Men smokers = 18%

    Tabulate the above data.

35. Tabulate the following data about the coffee drinking habit in two towns A and B :

Town A	Town B
55% were males	52% were males
28% were coffee drinkers	25% were coffee drinkers
18% were male coffee drinkers	16% were male coffee drinkers

36. Number of students ina college in 2005 was 510. Of these 480 were boys. In 2010 the number of boys increased by 100% and that of girls increased by 300% of that of 2005. In 2012 the total number of students in the college was 1200, the number of boys being double the number of girls. Tabulate the above information.

37. Out of a total number of 10,000 candidates who have applied for jobs in a government department, 6,854 were males, 3,146 were graduates and others non-graduates. The numbers of candidates with some experience were 2,623 of whom 1,860 were male. The number of male graduates was 2,012. The number of graduates with experience was 1,093 that include 323 females. Tabulate the above information.

38. Present the following information in the table.

    In the year 2005 there were 200 banks out of which 50 were private banks, in these fifty banks 10 had A.T.M facility and none of the nationalized banks had this facility. In the year 2010, there was an increase of 50 more banks in the private sector and 180 in nationalized banks. Total number of banks had A.T.M. facility in2010 is 100 and 75 private sector banks has this facility.

*****

Unit-IV

DIAGRAMMATIC AND GRAPHIC PRESENTATION OF DATA

Even though tabular presentation of data is more convincing, they are not interesting because of numerical figures. To make the data easy to understand even by a layman at a glance, diagrams and graphs are used. Just one diagram is enough to represent a given data more effectively than thousand words.

So, in this chapter, let us illustrate different ways of presentation of data through diagrams and graphs.

Importance/Need :

Diagrams and graphs are extremely useful because of the following reasons:

1. They are attractive and hence diagrams and graphs are commonly used in news papers and magazines for the purpose of advertisements and campaign.
2. They give bird's-eye view of the entire data at a glance.
3. They can be easily understood by common man.
4. They can be remembered for a longer period of time.
5. They facilitate comparison.

General Rules for constructing diagrams and graphs :

1. Every diagram should have a suitable **title** and is written above it.
2. **Proper scale** according to the size of the paper should be selected.
3. It should be **neat and attractive.**
4. It should not be **overloaded** with more information.
5. To indicate **different parts,** suitable shades, colours and crossings should be used.
6. An **index** indicating different shades, colours, crossing etc. used should be shown clearly.

7. It should be **complete** in all respects.
8. It should be **simple** and **self explanatory.**

Types of diagrams :

In practice, a very large variety of diagrams are in use and new ones are constantly being added. For the sake of convenience and simplicity, they may be divided under the following heads :

1. One Dimensional Diagrams
2. Two Dimensional Diagrams
3. Three Dimensional diagram
4. Pictograms and Cartograms.

According to present syllabus, our discussion is confined only to one and two dimensional diagrams.

1. One Dimensional Diagrams :

The diagrams drawn by considering only one dimension are called one dimensional diagrams. They are mostly bar diagrams. Here, only the height (length) of the bars is considered.

Some of the one dimensional diagrams are :

(a) Simple Bar diagram

(b) Multiple Bar diagram

(c) Component/Subdivided Bar diagram

(d) Percentage Bar diagram.

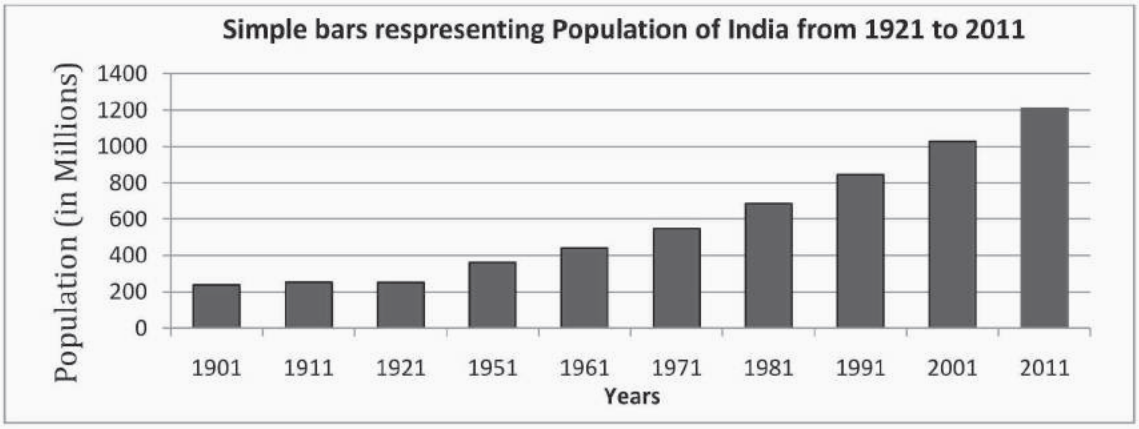
(a) Simple Bar Diagram :

Here, to represent the data rectangular bars of equal width and height proportional to the magnitude of the items are drawn with proper scale. Width of the bar is taken merely for attraction, but it is nothing to do with the data. It is used for the comparative study of two or more items of a single variable. For example, the figures of imports, exports, and population etc., for few years are to be represented by simple bars.

Example1 : Following figures represent the decadal change of population of India. Draw a simple bar diagram.

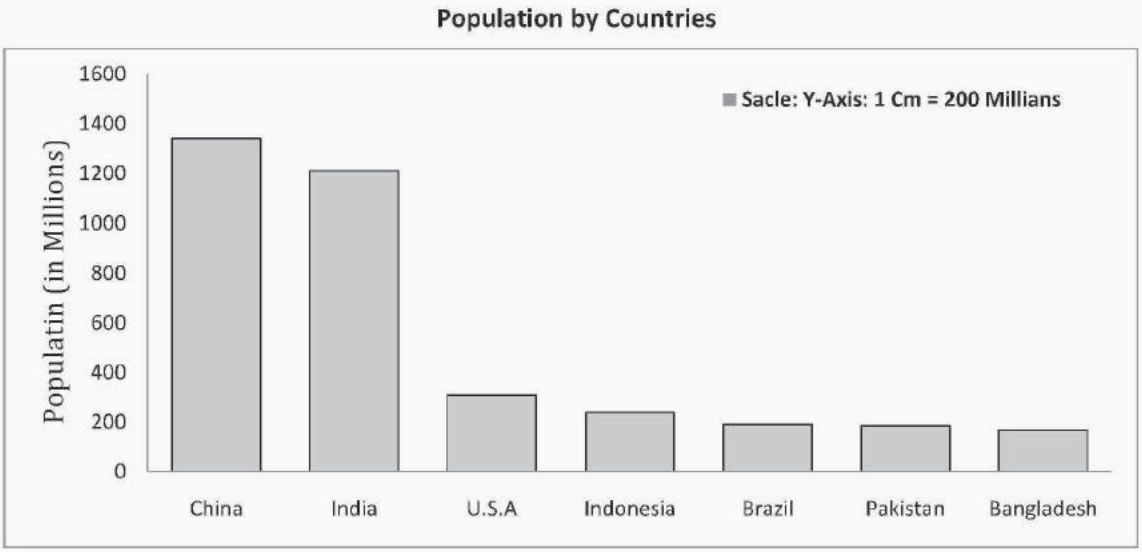
Year	1901	1911	1921	1931	1941	1951	1961	1971	1981	1991	2001	2011
Population (Million)	238	252	251	278	318	361	439	548	688	846	1028	1210

Solution :



Example2: Represent the data regarding the population of different countries by simple bar diagram.

Country	China	India	U.S.A	Indonesia	Brazil	Pakistan	Bangladesh
Population (Million)	1341	1210	309	238	191	185	165



Note: While constructing simple (multiple or compound) bar diagrams, the scale is determined on the basis of the highest value in the data.The space between consecutive bars (sets of bars) should be same.

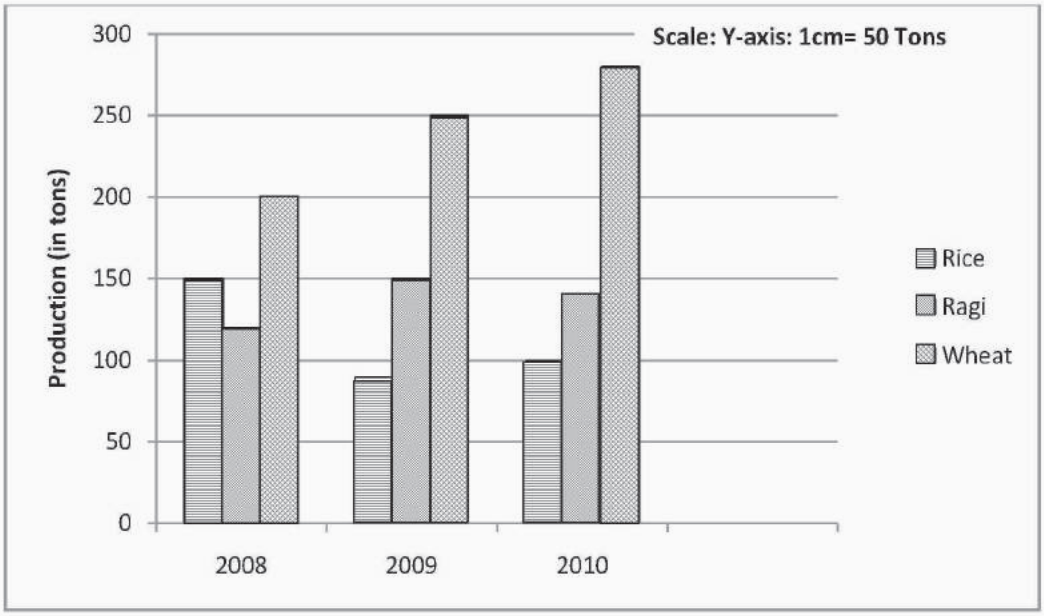
(b) Multiple (Compound) Bar Diagrams :

Multiple bar diagrams are used to represent various sets of interlinked data. Here, a set of adjacent bars (one for each component) is drawn for comparison. Bars are differentiated with shades or colours and an index is provided for easy reference of bars.

Example: 1 Following is the data showing foodgrain production in different years. Draw a multiple bar diagram to represent the data.

Foodgrain	Productin (in tons)		
	2008	2009	2010
Rice	150	90	100
Ragi	120	150	140
Wheat	200	250	280

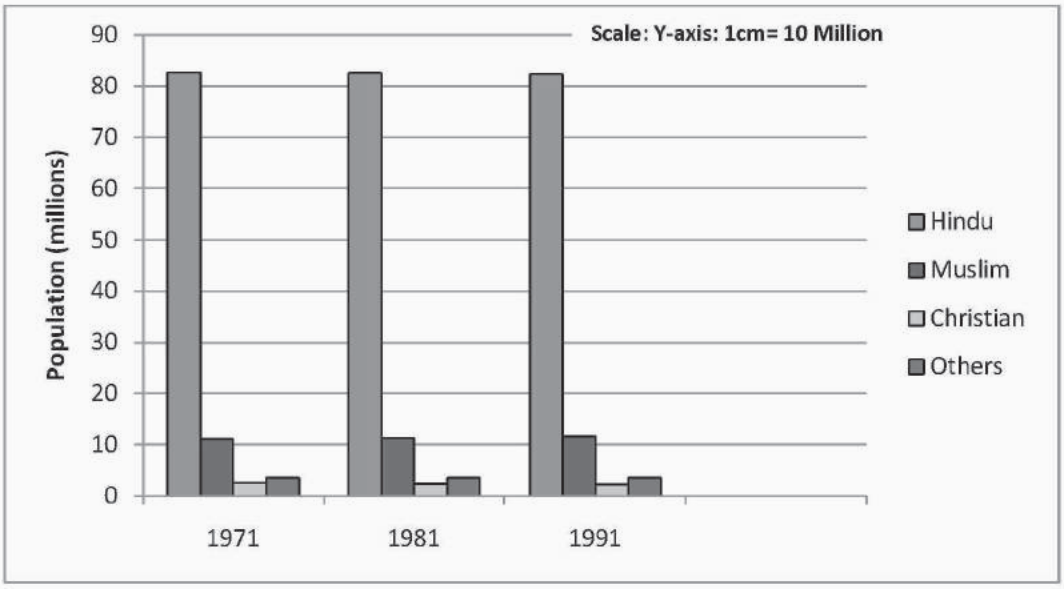
Solution : Multiple bar diagram represents foodgrain production for ree different years.



Example 2. Draw a multiple bar diagram from the following data :

Religion	Population (millions)		
	1971	1981	1991
Hindu	82.7	82.6	82.41
Muslim	11.2	11.4	11.67
Christian	2.6	2.4	2.32
Others	3.5	3.6	3.6

Solution : Multiple Bar diagram representing the Religion wisepopulation for the year s1971, 1981, 1991



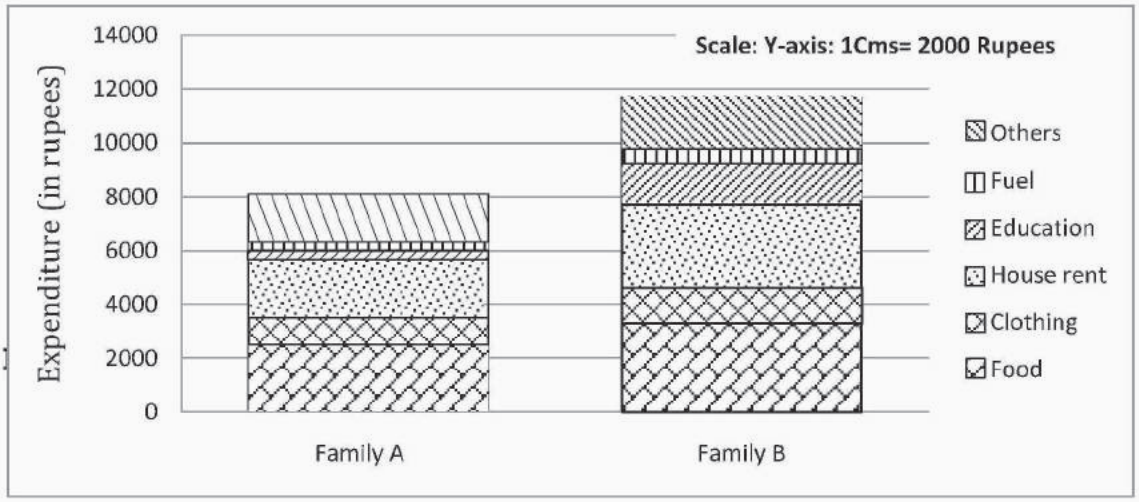
(c) Component/Sub-divided Bar Diagrams :

Here, simple bars are drawn for the total magnitude of the variables, further the bars are divided into various segments, each segment representing a given component of the total. To distinguish between the various components different shades or colours are used. For easy reference of the components an index is provided. These are useful in comparing the total magnitudes, along with the components.

Example 1. For the following data regarding the expenditure of families A and B. Represent the data by sub-divided bar diagram.

Items	Expenditure in Rupees	
	Family A	Family B
Food	2580	3350
Clothing	880	1250
House rent	2200	3100
Education	360	1550
Fuel	280	450
Others	1800	2000

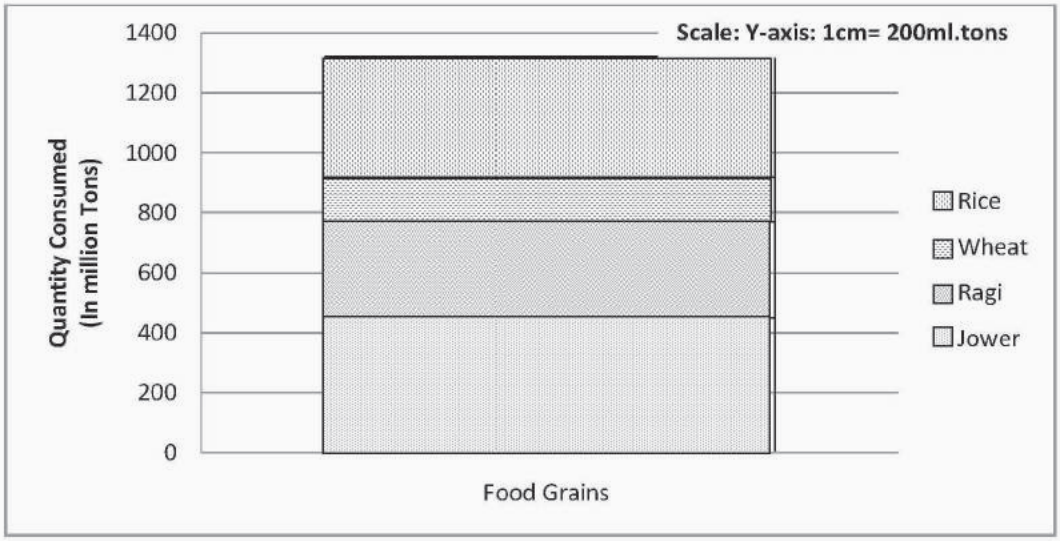
Solution : SUB-DIVIDED BAR DIAGRAM REPRESENTING THE EXPENDITURES OF FAMILY A AND B



Example 2. Following data represents the major consumption of food grains in Karnataka state. Represent the data by a subdivided bar diagram.

Food Grains	Quantity Consumed (In million Tons)
Jower	450
Ragi	320
Wheat	150
Rice	400
Total	1320

Component bar diagram representing consumption of food grains in Karnataka State.



(d) Percentage Bar Diagrams :

This is another form of component bar diagram. Here, the components are not the actual values but percentages of the whole. Bars are drawn with height equal to 100 each and bars are divided into segments according to their values expressed as percentage of the respective totals. The different subdivisions are shaded properly and an index is provided for easy reference of the components. These are useful in comparing percentage components. Here, the percentage (%) of an individual component $= \frac{\text{Individual value}}{\text{Total value}} \times 100$

Example 1. Following are the marks obtained by two students A and B in an annual examination. Represent the data by percentage bar diagram.

Subjects	Marks of students	
	Student A	Student B
Kannada	72	82
English	85	92
Statistics	97	95
Economics	88	90
Business	90	87
Accountancy	94	98
Total	526	544

Solution:

Convert marks in to percentages: % of a subject = $\frac{\text{Individual value}}{\text{Total value}} \times 100$

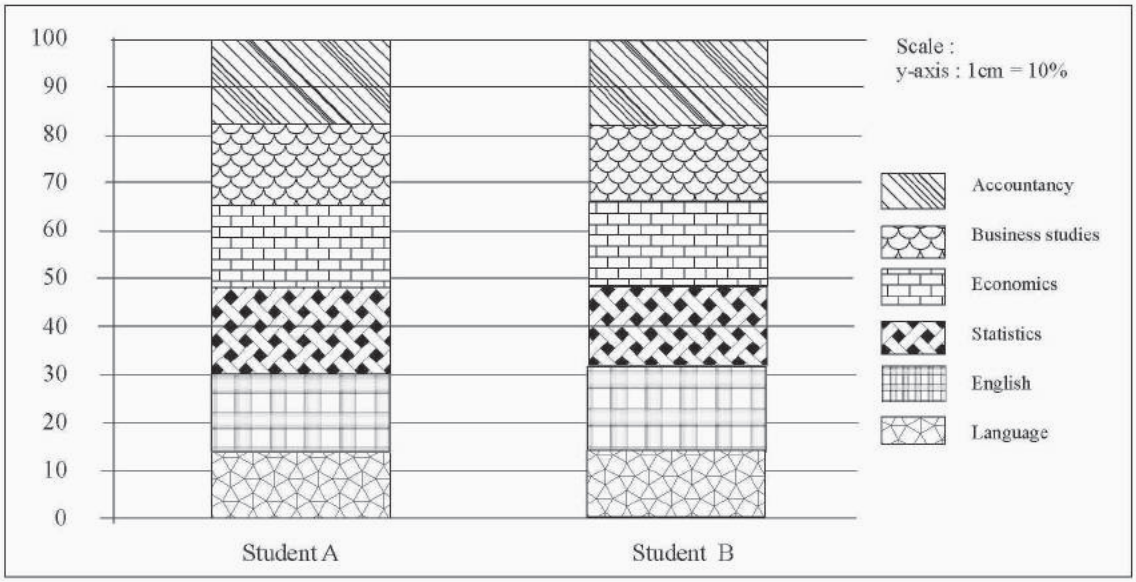
Student A: Kannada = $\frac{72}{526} \times 100 = 13.69 \cong 14\%$

Student B: Statistics = $\frac{95}{544} \times 100 = 17.46 \cong 17\%$

Similarly , other values can be calculated.

Subjects	Student A			Student B		
	Marks	In %	Cumulative %	Marks	In %	Cumulative %
Kannada	72	14	14	82	15	15
English	85	16	30	92	17	32
Statistics	97	18	48	95	17	49
Economics	88	17	65	90	17	66
Business studies	90	17	82	87	16	82
Accountancy	94	18	100	98	18	100
Total	526	100	-	544	100	-

Percentage Bar Diagram showing the marks of the students A and B



2. Two Dimensional Diagrams :

The diagrams drawn by considering two dimensions are called two dimensional diagrams. These are also called as area diagrams. In two dimensional diagrams, both height and width of the bars are taken into account and area generated by them is proportional to the magnitude of the data. Some of the two dimensional diagrams are- (a) Rectangles (b) Squares (c) Circles (d) Pie diagram chart. According to the present syllabus, we study only pie-diagram.

Pie-diagram (pie chart, circular diagram) :

In case of bar diagram, a bar is sub-divided into segments to represent its components. Where, in case of pie-diagram a circle is sub-divided into sectors to represent its components. Where, the area of the sector will be proportional to the magnitude of the components. This will be done by computing the angles of each sector by using the formula :

$$\text{Angle of an individual component} = \frac{\text{Individual value}}{\text{Total value}} \times 360^0$$

Example 1.

The following table shows the cost structure of Indian Hotel Industry in percentages.

Cost components	Total expenses (%)
Administrative	30
Employees payments	20
Repairs and	14
Food and Beverages	12
Power	16
Selling expenses	8

Draw a Pie diagram to represent the data.

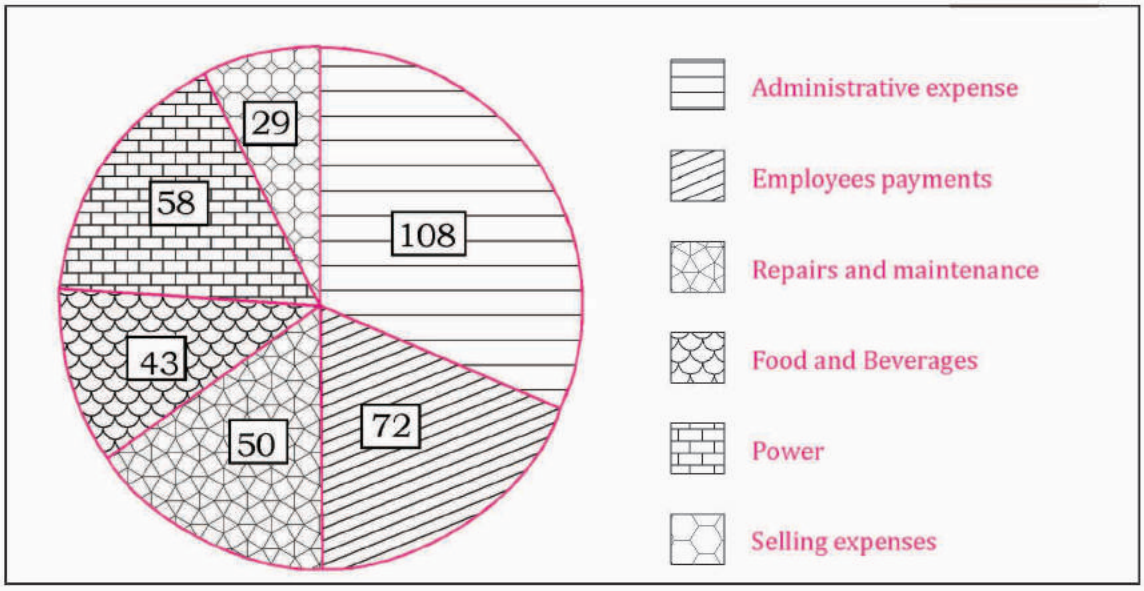
Solution :

$$\text{Angle of an individual component} = \frac{\text{Individual value}}{\text{Total value}} \times 360^0$$

Administrative expense= $\frac{30}{100} \times 360^0 = 108^0$

Cost components	Total expenses (%)	Angles
Administrative	30	108
Employees payments	20	72
Repairs and	14	50
Food and Beverages	12	43
Power	16	58
Selling expenses	8	29
Total	100	360

Pie - diagram show sthe expenses of Indian hotel industry

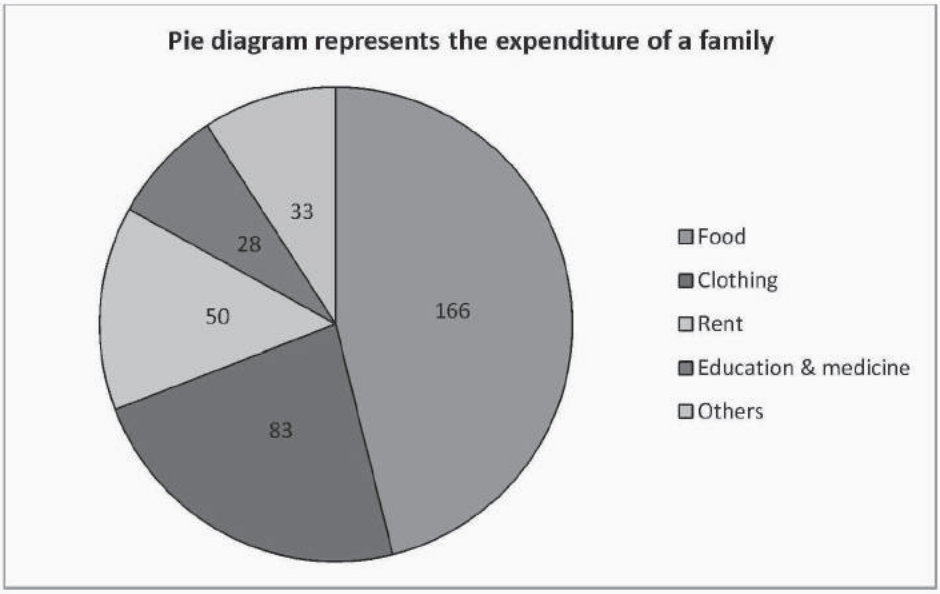


Example 2. The following data relates to the expenditure of a family per month. Draw a pie chart.

Items	Expense (in Rs.)
Food	3000
Clothing	1500
Rent	900
Education & medicine	500
Others	600

Solution :

Items	Expense (in Rs.)	Angles
Food	3000	166
Clothing	1500	83
Rent	900	50
Education & medicine	500	28
Others	600	33
Total	6500	360



Graphical presentation of frequency distributions

There are many ways of presentation of frequency distributions in the form of graphs. Some commonly used graphs are :

1. Histogram
2. Frequency Polygon
3. Frequency Curve
4. Ogives.

Histogram :

Histogram is constructed for a continuous frequency distribution. "A histogram is a set of adjacent rectangles, whose height and width are