

**Anekant Education Society's
Tuljaram Chaturchand College of Arts, Science and
Commerce, Baramati**

Autonomous

Course Structure for F. Y. B. Sc. (Computer Science) STATISTICS

Semester	Paper Code	Title of Paper	No. of Credits
I	CSST-1101	Statistical Methods -I	2
	CSST-1102	Probability and some discrete probability distributions	2
	CSST-1103	Practical-I	2

Programme Outcomes:

PO1. Computer Knowledge : Apply the knowledge of mathematics, statistics and computer fundamentals to IT application

PO2. Design / Development of solution: Design solution for IT applications using latest technologies and develop and implement the solutions using various latest language.

PO3. Modern tool usage : Create, select and apply appropriate techniques, resources and modern engineering and IT tools including prediction and modelling to complex IT applications with an understanding of the limitations.

PO4. Environment and sustainability : Understand the impact of the IT analyst solutions in societal and environmental contexts, and demonstrate the knowledge and need for sustainable development.

PO5. Ethics: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.

PO6. Individual and Team work : Function effectively as an individual, and as a member or leader in diverse team, and in multidisciplinary settings.

PO7. Innovation, employability and Entrepreneurial skills : Identify opportunity, pursue that opportunity to create value and wealth for the betterment of the individual and society at large. Develop the capacity to study and research independently that will help to develop skills for transition to employment in hardware/software companies.

SYLLABUS (CBCS) FOR F. Y. B. Sc. (Computer Science) STATISTICS
(w.e. from June, 2019)
(2019 Pattern)
Academic Year 2019-2020

Class : F.Y. B. Sc. (Computer Science) (Semester- I)

Paper Code: CSST-1101

Paper : I

Title of Paper: Statistical Methods I

Credit : 2 credits

No. of lectures: 36

A) Learning Objectives:

1. To compute various measures of central tendency, dispersion, skewness and kurtosis.
2. Compute the correlation coefficient for bivariate data and interpret it.

B) Course Outcome

By the end of the course, students will be able to:

CO1. Categorize different types of data accurately.

CO2. Data representation and visualization

CO3. Recognize real-world examples of structured, unstructured, and semi-structured data.

CO4. Do computation of various measures of central tendency.

CO5. Do computation of various measures of dispersion.

CO6. Understand Concept of inventory which are used in software development

CO7. Grasp the significance of dispersion measures in data analysis.

TOPICS/CONTENTS:

UNIT1: Data Representation

(6L)

- 1.1 Definition, importance, scope and limitations of statistics w.r.to computer science.
- 1.2 Data Condensation: Types of data (Primary and secondary), Attributes and variables, discrete and Continuous variables, classification and construction of frequency distribution.
- 1.3 Graphical Representation: Histogram, Frequency polygon, Frequency curve, Ogive Curves, Steam and leaf chart.
- 1.4 Numerical problems related to real life situations.

UNIT2: Measures of central tendency

(8L)

- 2.1 Concept of central tendency, requisites of good measures of central tendency.

- 2.2 Arithmetic mean: Definition, computation for ungrouped and grouped data, combined mean, weighted mean, merits and demerits.
- 2.3 Median and Mode: Definition, formula for computation for ungrouped and grouped data, graphical method, merits and demerits. Empirical relation between mean, median and mode (without proof)
- 2.4 Partition Values: Quartiles, Percentiles, Deciles, Box Plot.
- 2.5 Numerical problems related to real life situations.

UNIT3: Measures of Dispersion (6L)

- 3.1 Concept of dispersion and measures of dispersion, requisites of good measures of dispersion, absolute and relative measures of dispersion.,
- 3.2 Range and Quartile Deviation: definition for ungrouped and grouped data and their coefficients, merits and demerits.,
- 3.3 Variance and Standard deviation: definition for ungrouped and grouped data, coefficient of variation, combined variance & standard deviation, merits and demerits.
- 3.4 Numerical problems related to real life situations.

UNIT4: Moments, Skewness and Kurtosis (8L)

- 4.1 Raw and central moments: definition for ungrouped and grouped data (only first four moments), relation between central and raw moments upto fourth order. (without proof)
- 4.2 Measures of Skewness: Types of skewness, Pearson's and Bowley's coefficient of skewness, Measures of skewness based on moments.,
- 4.3 Measures of Kurtosis: Types of kurtosis, Measures of kurtosis based on moments
- 4.4 Numerical problems related to real life situations

UNIT 5: Correlation (For bivariate raw data) (8L)

- 5.1 Concept of bivariate data, scatter diagram, concept of correlation, positive correlation, negative correlation, zero correlation.
- 5.2 Karl Pearson's coefficient of correlation, properties of correlation coefficient, interpretation of correlation coefficient.
- 5.3 Spearman's rank correlation coefficient (formula with and without ties).
- 5.4 Numerical problems

References:

- 1 Fundamentals of Applied Statistics (3rd Edition), Gupta and Kapoor, S.Chand and Sons, New Delhi, 1987.

- 2 An Introductory Statistics, Kennedy and Gentle.
- 3 Statistical Methods, G.W. Snedecor, W.G. Cochran, John Wiley & sons, 1989.
- 4 Introduction to Linear Regression Analysis, Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining, Wiley

Programme Outcomes and Course Outcomes Mapping :

Course Outcomes	Programme Outcomes(POs)						
	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	3	-	-	-	-	-	-
CO2	-	3	2	-	-	-	-
CO3	-	-	-	-	-	2	-
CO4	3	-	-	-	-	-	-
CO5	-	-	-	-	-	-	3
CO6	-	-	-	3	-	-	-
CO7	-	-	-	-	2	-	-

PO1. Computer Knowledge:

CO1. Categorize different types of data accurately. (Weightage: 3)

Justification: This CO involves the application of mathematical and statistical knowledge to categorize data accurately, aligning with the need for computer knowledge.

CO4. Do computation of various measures of central tendency. (Weightage: 3)

Justification: Central tendency measures involve mathematical computations, showcasing the application of mathematical knowledge in IT applications.

PO2. Design/Development of solution:

CO2. Data representation and visualization. (Weightage: 3)

Justification: Data representation and visualization are crucial aspects of designing solutions for IT applications, and this aligns with the design and development outcome.

PO3. Modern tool usage: Create, select and apply appropriate techniques, resources and modern engineering and IT tools including prediction and modeling to complex IT applications with an understanding of the limitations.

CO2. Data representation and visualization. (Weightage: 2)

Justification: The use of modern tools in data representation and visualization is essential for creating and implementing IT solutions.

PO4. Environment and sustainability:

CO6. Understand the Concept of inventory which is used in software development. (Weightage: 3)

Justification: Understanding inventory in software development is critical for optimizing resources, aligning with the need for sustainability.

PO5. Ethics:

CO7. Grasp the significance of dispersion measures in data analysis. (Weightage: 2)

Justification: Dispersion measures in data analysis involve ethical considerations in interpreting and presenting data accurately.

PO6. Individual and Teamwork:

CO3. Recognize real-world examples of structured, unstructured, and semi-structured data. (Weightage: 2)

Justification: Recognizing different types of data is crucial for effective collaboration and teamwork in diverse settings.

PO7. Innovation, employability and Entrepreneurial skills: CO5. Do computation of various measures of dispersion. (Weightage: 3)

Justification: Computation of measures of dispersion is essential for identifying opportunities, creating value, and making informed decisions, aligning with the focus on innovation and employability.

SYLLABUS (CBCS) FOR F. Y. B. Sc. (Computer Science) STATISTICS
(w.e. from June, 2019)
(2019 Pattern)

Academic Year 2019-2020

Class : F.Y. B. Sc. (Computer Science) (Semester- I)

Paper Code: CSST-1102

Paper : II

Title of Paper: Probability and some discrete probability distributions

Credit : 2 credits

No. of lectures: 36

A) Learning Objectives:

The main objective of this course is to acquaint students with some basic concepts of probability, axiomatic theory of probability, concept of random variable, probability distribution (univariate) and some standard continuous distributions.

B) Course Outcome:

By the end of the course, students will be able to:

CO1. Distinguish between random and non-random experiments.

CO2. Find the probabilities of various events.

CO3. Obtain probability distribution of univariate continuous random variables.

CO4. Use distributions in real life situations.

CO5. Probability Distributions give up the possible outcome of any random event.

CO6. Apply Binomial distribution and Poisson distribution in real life situations.

CO7. Studying probability will help children to develop critical thinking skills and to interpret the probability that surround us daily.

TOPICS/CONTENTS:

Unit-1. Sample space and Events:

(6L)

1.1 Concepts of experiments, deterministic and nondeterministic experiments.

1.2 Definitions: Sample space, Types of sample space, Event, Types of Events:
Elementary event, Complementary event, Sure event, Impossible event.

1.3 Concept of occurrence of an event, Equally-likely events

1.4 Algebra of events (Union, Intersection, Complementation).

1.5 Definitions of Mutually exclusive events, Exhaustive events.

1.6 Algebra of events and its representation of events in set theory notation:

Occurrence of the following events:

- a) at least one of the given events
- b) none of the given events
- c) all of the given events
- d) mutually exclusive events
- e) mutually exhaustive events
- f) exactly one event out of the given events.

1.7 Numerical problems related to real life situations.

Unit-2. Probability:

(13L)

2.1 Concept of Multiplication principle and Permutation and Combination

2.2 Classical Probability: Classical definition of probability, examples, Probability model, probability of an event, examples. Axiomatic definition of probability.

Proof of the results:

- i) $0 \leq P(A) \leq 1$,
- ii) $P(A) + P(A^c) = 1$
- iii) $P(\Phi) = 0$
- iv) If $A \subset B$, $P(A) \leq P(B)$
- v) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (Addition theorem of probability) and its generalization (Statement only).

Numerical problems related to real life situations.

2.3 Conditional Probability

- Concepts and definitions of Conditional Probability
- Definition of conditional probability of an event.
- Multiplication theorem for two events. Examples.
- Partition of sample space.
- Idea of Posterior probability, Statement and proof of Bayes' theorem, examples on Baye's theorem.

2.4 Independence of Events

- Concept of Independence of two events.
- Proof of the result that if A and B are independent then, i) A and B^c , ii) A^c and B iii) A^c and B^c are independent.
- Pair wise and Mutual Independence for three events.
- Numerical problems related to real life situations.

UNIT 3: Discrete random Variable

(4L)

3.1 Definition of random variable and discrete random variable

3.2 Definition of probability mass function, distribution function and its properties

3.3 Definition of expectation and variance, theorems on expectation

3.4 Numerical problems related to real life situations

**UNIT 4: Standard Discrete distribution
(13L)**

- 4.1 Uniform Distribution : definition, mean, variance
- 4.2 Bernoulli Distribution : definition, mean, variance, additive property
- 4.3 Binomial Distribution : definition, mean, variance, additive property
- 4.4 Poisson Distribution : definition, mean, variance, mode, additive property, limiting case of $B(n, p)$
- 4.5 Numerical problems related to real life situations

References:

1. Bhat B. R., Srivenkatramana T and Madhava Rao K. S. (1997): Statistics: a Beginner’s Text, Vol. II, New Age International (P) Ltd.
2. Goon A. M., Gupta M. K., Das Gupta B. (1999): Fundamentals of Statistics, Vol.II, World Press, Calcutta.
3. Hogg R. V. and Crag R. G.: Introduction to Mathematical Statistics Ed.4.
4. Gupta and Kapoor : Fundamentals of Mathematical Statistics, Sultan Chand and Sons, New Delhi.
5. Meyer P.L.(1970): Introductory Probability and Statistical Applications, Edition Wesley.
6. Rohatgi V. K. and Saleh A. K. Md. E. (2002): An Introduction to probability and statistics. John wiley & Sons (Asia)
7. Gupta and Kapoor : Fundamentals of Applied Statistics, Sultan Chand and Sons, New Delhi.

Programme Outcomes and Course Outcomes Mapping :

Course Outcomes	Programme Outcomes (POs)						
	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	-	-	-	-	-	-	-
CO2	-	3	-	-	-	-	-
CO3	1	-	2	-	-	-	-
CO4	-	-	-	-	-	-	3
CO5	-	-	-	-	-	-	-
CO6	-	-	-	-	-	-	-
CO7	-	-	-	-	2	2	-

PO1. Computer Knowledge

CO3. Obtain probability distribution of univariate continuous random variables (Weightage: 1)

Justification: While computer knowledge may not be explicitly required for understanding probability distributions, it can be beneficial for handling data and simulations related to probability.

PO2. Design / Development of Solution

CO2. Find the probabilities of various events (Weightage: 3)

Justification: Designing and developing solutions often involves assessing the probabilities of different events for informed decision-making.

PO3. Modern Tool Usage

CO3. Obtain probability distribution of univariate continuous random variables (Weightage: 2)

Justification: Modern tools are often used to handle and analyze probability distributions, especially in the case of continuous random variables.

PO4. Environment and Sustainability

No direct mapping observed

Justification: The provided COs do not explicitly address environmental or sustainability aspects.

PO5. Ethics

CO7. Studying probability will help children develop critical thinking skills and interpret the probability that surrounds us daily (Weightage: 2)

Justification: Understanding probability involves critical thinking skills and ethical considerations when interpreting and communicating probabilities.

PO6. Individual and Team Work

CO7. Studying probability will help children develop critical thinking skills and interpret the probability that surrounds us daily (Weightage: 2)

Justification: The study of probability can involve both individual and collaborative efforts in interpreting and analyzing real-world scenarios.

PO7. Innovation, Employability, and Entrepreneurial Skills

CO4. Use distributions in real life situations (Weightage: 3)

Justification: Applying probability distributions in real-life situations requires innovative thinking and contributes to employability skills.

SYLLABUS (CBCS) FOR F. Y. B. Sc. (Computer Science) STATISTICS
(w.e. from June, 2019)
(2019 Pattern)

Academic Year 2019-2020

Class : F.Y. B. Sc. (Computer Science) (Semester- I)
Paper Code: CSST-1103
Paper : III Title of Paper: Practical-I
Credit : 2 credits No. of lectures: 40

Pre requisites: Knowledge of the topics in the theory papers.

A) Learning Objectives:

The main objective of this course is to acquaint students with computation ability with interpreting summary Statistics

B) Course Outcome:

By the end of the course, students will be able to:

- CO1.** Represent statistical data diagrammatically and graphically .
- CO2.** Compute various measures of central tendency, dispersion, moments, skewness and kurtosis.
- CO3.** Compute correlation coefficient, regression coefficients and to interpret the results.
- CO4.** Interpret summary Statistics of computer output
- CO5.** Use statistical tools like graphical representation, summary statistics of data with the help of MS-Excel and R-software
- CO6.** Understand conditional probability will help your business not only maximize its revenue and net profit ratio but also help you improve your cash flow.
- CO7.** Gather information in a measured and systematic manner to ensure accuracy and facilitate data analysis

INDEX

Sr.No.	Title of Experiments	Page No.	Remark
1.	Measures of Central Tendency and Dispersion – I		
2.	Measures of Central Tendency and Dispersion – II		
3.	Problems on Simple Probability, Conditional Probability, Bayes' Theorem and Independence of Event		
4.	Measures of Skewness and Kurtosis		
5.	Correlation		
6.	Fitting of Binomial distribution		
7.	Fitting of Poisson distribution		
8.	Diagrammatic Representation and Descriptive Statistics using R		
9.	Computation of probabilities of discrete probability distributions using R		
10	Project: Data collection, its condensation and representation		

Programme Outcomes and Course Outcomes Mapping :

Course Outcomes	Programme Outcomes (POs)						
	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	-	-		-	-	-	-
CO2	-	2	-	-	-	-	-
CO3	-	-	-	-	-	-	-
CO4	-	-	-	-	-	-	-
CO5	3	-	3	-	-	-	-
CO6	-	-	-	-	-	-	3
CO7	-	-	-	-	2	2	-

PO1. Computer Knowledge

CO5. Use statistical tools like graphical representation, summary statistics of data with the help of MS-Excel and R-software (Weightage: 3)

Justification: This course outcome directly involves utilizing statistical tools, including MS-Excel and R-software, demonstrating proficiency in computer knowledge.

PO2. Design / Development of Solution

CO2. Compute various measures of central tendency, dispersion, moments, skewness and kurtosis (Weightage: 2)

Justification: Designing solutions often requires a thorough understanding of statistical measures to make informed decisions.

PO3. Modern Tool Usage

CO5. Use statistical tools like graphical representation, summary statistics of data with the help of MS-Excel and R-software (Weightage: 3)

Justification: Utilizing modern statistical tools like MS-Excel and R-software aligns with the modern tool usage outcome.

PO4. Environment and Sustainability

No direct mapping observed

Justification: The provided COs do not explicitly address environmental or sustainability aspects.

PO5. Ethics

CO7. Gather information in a measured and systematic manner to ensure accuracy and facilitate data analysis (Weightage: 2)

Justification: Gathering information in a measured and systematic manner reflects ethical considerations in data analysis.

PO6. Individual and Team Work

CO7. Gather information in a measured and systematic manner to ensure accuracy and facilitate data analysis (Weightage: 2)

Justification: Individual and team work are involved in systematically gathering and analyzing data.

PO7. Innovation, Employability, and Entrepreneurial Skills

CO6. Understand conditional probability will help your business not only maximize its revenue and net profit ratio but also help you improve your cash flow (Weightage: 3)

Justification: Understanding conditional probability demonstrates innovative thinking for business improvement and aligns with employability skills.