Progressive Education Society's
Modern College of Arts, Science and Commerce,
Shivajinagar, Pune – 5

First Year of B.Sc. Data Science
Course under NEP 2025-26

Course Code: 24CsDscU1201
Course Name: Descriptive Statistics

Teaching Scheme: TH: 2 Hours/Week **Credit: 02**
Examination Scheme: CIA: 20 Marks **End-Sem: 30 Marks**

Prerequisite Courses:

- Basic Mathematics.

Course Objectives:

- To study different types of data.
- To Study Statistical Population and sample.
- To compute different types of measures of central tendency and dispersion.
- To Study the concept of skweness, kurtosis.
- To study kind and extent of relationship between two variable.

Course Outcomes:

CO No	Course Outcomes (COs)	Bloom's Cognitive Level
CO 1	Classify data sets according to types of variables and data.	2
CO 2	Compute appropriate measures of central tendency for given data sets.	3
CO 3	Compute measures of dispersion and analyze data variability.	3
CO 4	Calculate different types of moments for data sets.	3
CO 5	Interpret skewness and kurtosis for real-life data analysis.	4
CO 6	Determine the kind and extent of relationship between two variables.	4

Course Contents

Unit 1	**Population and Sample**	**03L**
	1.1 Types of characteristics: Qualitative: Measurement scale: Nominal scale, ordinal scale (Use of Likert scale). Quantitative: Discrete and continuous variables : Measurement scale (Using interval scale, ratio scale) 1.2 Types of data: (a) Primary data, Secondary data. (b) Cross-sectional data, Time series data, Directional data. 1.3 Notion of a statistical population and sample: Finite population, infinite population, homogeneous population and heterogeneous population.	
Unit 2	**Measures of Central Tendency**	**10L**
	2.1 Concept of central tendency of statistical data, Statistical averages, characteristics of a good statistical average. 2.2 Arithmetic Mean (A.M.): Definition, effect of change of origin and scale, combined mean of a number of groups, merits and demerits, trimmed arithmetic mean. 2.3 Mode and Median: Definition, merits and demerits. Empirical relation between mean, median and mode. 2.4 Partition values: Quartiles, Deciles and Percentiles. Geometric Mean (G.M.): Definition, merits and demerits. Harmonic Mean (H.M.): Definition, merits and demerits. Order relation between arithmetic mean, geometric mean, harmonic mean. 2.5 Weighted Mean: weighted A.M., G.M. and H.M. 2.6 Situations where one kind of average is preferable to other.	
Unit 3	**Measures of Dispersion**	**07L**
	3.1 Concept of dispersion, characteristics of good measure of dispersion. 3.2 Range, Semi-interquartile range (Quartile deviation): Definition, merits and demerits. 3.3 Mean absolute deviation: Definition, merits and demerits, minimality property (without proof) 3.4 Variance and standard deviation: Definition, merits and demerits, effect of change of origin and scale, combined variance for n groups (derivation for two groups).Mean squared deviation: Definition, minimality property of mean squared deviation. 3.5 Measures of dispersion for comparison: Coefficient of range, coefficient of quartile deviation and coefficient of mean deviation, coefficient of variation (C.V.).	
Unit 4	**Moments, Skewness and Kurtosis**	**06L**

	4.1 Moments: Raw moments (m'_r) for ungrouped and grouped data. Central moments (m_r) for ungrouped and grouped data, Effect of change of origin and scale. Relations between central moments and raw moments upto 4^{th} order. 4.2 Concept of skewness of frequency distribution, positive skewness, negative skewness, symmetric frequency distribution and its relation with central tendency. Bowley's coefficient of skewness: Bowley's coefficient of skewness, Karl Pearson's coefficient of skewness, Measures of skewness based on moments (β_1,γ_1).	
	4.3 Concepts of kurtosis, Types of kurtosis: Leptokurtic, Mesokurtic and Platykurtic, Measures of kurtosis based on moments (β_2,γ_2). 4.4 Properties of β_1 and β_2 (i) $\beta_2 \geq 1$ (ii) $\beta_2 \geq \beta_1+1$ (Statement of results related measures of skewness and kurtosis).	
Unit 5	**Correlation**	**04L**
	5.1 Measure of linear dependence between two variables: Covariance between two variables (m11) : Definition, computation, effect of change of origin and scale. Karl Pearson's coefficient of correlation (r) : Definition, Properties: (i) $-1 \leq r \leq 1$ (ii) Effect of change of origin and scale Computation for ungrouped data and grouped frequency distributed data with interpretation. 5.2 Spearman's rank correlation coefficient: Definition, computation and interpretation (without ties). In case of ties, compute Karl Pearson's correlation coefficient between ranks. (Derivations are not expected)	

References:

1. Agarwal B. L. (2003). Programmed Statistics, 2nd edition, New Age International Publishers, NewDelhi.
2. Das (2009). Statistical Methods, Tata Mcgraw Hill *Publishing.*
3. *Dixit P.G, Prayag V.R.,.Descriptive Statistics,Nirali prakashan., Pune*
4. Goon, A. M., Gupta, M. K. and Dasgupta, B. (2016). Fundamentals of Statistics, Vol. 1, 6th Revised Edition, The World Press Pvt. Ltd., Calcutta.
5. Gore.Anil, Pranjapesharayu ,Kulkarni Madhav. Statistics for everyone. SIPFAcadamy publisher,Nashik
6. Gupta, S. C. and Kapoor, V. K. (2000). Fundamentals of Mathematical Statistics, 10th Edition, Sultan Chand and Sons Publishers, New Delhi.
7. Krzanowski (2007). Statistical Principals and Techniques in Scientific and Social Research, Oxford University Press Inc., New York.
8. Mohanty (2016). Basic Statistics, Scientific Publisher
9. Mukhopadhyay P. (2015). Applied Statistics ,*Publisher*: Books & Allied (*P*) Ltd.
10. Purohit, S. G., Gore S. D.,Deshmukh S. R (2008). Statistics Using R, Narosa Publishing House, NewDelhi
11. Rastogi (2015). Biostatistics, 3rd Edition, Publisher Medtec
12. Robert S. Witte, John S. Witte (2017). Statistics, *Publisher*: Wiley
13. *Roussas*, George G. (2016). First course in mathematical *statistics*.2nd edition *Publisher:* Academic Press.

Progressive Education Society's
Modern College of Arts, Science and Commerce,
Shivajinagar, Pune – 5

First Year of B.Sc. Data Science
Course under NEP 2025-26

Course Code: 24CsDscU1202
Course Name: Lab Course on 24CsDscU1201

Teaching Scheme: TH: 4 Hours/Week **Credit: 02**
Examination Scheme: CIA: 20 Marks **End-Sem: 30 Marks**

Pre requisites :

- Basic Mathematics.

Course Objectives:

- To learn graphical representation, summary statistics of the data.
- To Study the concept of Skweness, Kurtosis and correlation.
- To learn attributes and its measure of association.

Course Outcomes:

CO No	Course Outcomes (COs)	Bloom's Cognitive Level
CO 1	Represent data using tables, charts, and graphs.	2
CO 2	Compute and interpret measures of central tendency and dispersion.	3
CO 3	Apply appropriate statistical measures to different data sets.	3
CO 4	Compute measures of skewness and kurtosis.	3
CO 5	Analyze correlation between variables using appropriate methods.	4
CO 6	Evaluate attributes and measures of association for categorical data.	5

Course Contents

Sr. No.	Title of Experiment/ Practical
1	Diagrammatic and Graphical representation of statistical data-I: Bar diagram, Histogram, Stem and leaf.
2	Diagrammatic and Graphical representation of statistical data-II: Frequency Polygon, Frequency curve, Ogive curve (less and more than type)

3	Diagrammatic and Graphical representation of statistical data-III: Pie chart, Sub-divided bar diagram, Multiple bar diagram, Dual axis plot
4	Computation of measures of central tendency and dispersion (ungrouped data). Use of an appropriate measure and interpretation of results and computation of partition values.
5	Computation of measures of central tendency and dispersion (grouped data)-I Use of an appropriate measure and interpretation of results and computation of partition values.
6	Computation of measures of central tendency and dispersion (grouped data)-II Use of an appropriate measure and interpretation of results and computation of partition values.
7	Measures of Skewness, Kurtosis based quartiles, Box and Whisker plot and detection of outliers.
8	Moments, Measures of Skewness, Kurtosis based moments.
9	Scatter diagram, computation of correlation coefficient
10	Computation of Spearman's Rank Correlation Coefficient.
11	Represent Attributes data in contingency table, Computation of frequencies of attributes.
12	Check Consistency of data, independence of attributes, Computation of Yule coefficient of association and it's interpretation.
13	Project work (equivalent to 3 practicals)

Progresive Education Society's
Modern College of Arts, Science and Commerce,
Shivajinagar, Pune – 5

First Year of B.Sc. Data Science
(Course under NEP 2025-26)

Course Code: 24CsDscU2201
Course Name: Discrete Probability and Probability Distributions

Teaching Scheme: TH: 2 Hours/Week **Credit: 2**
Examination Scheme: CIA: 20 Marks **End-Sem: 30 Marks**

Course Objectives: To Introduce to students:

- the basic concepts of probability, axiomatic approach of probability, concept of independence of two events, conditional probability and Bayes law
- the concept of random variable and univariate probability distribution.
- the concept of independence of variables involved.
- Standard discrete probability distributions based on finite and infinite sample space.
- Interrelations between Standard discrete probability distributions.

Course Outcomes:

CO No	Course Outcomes (COs)	Bloom's Cognitive Level
CO 1	Differentiate between random and non-random experiments.	2
CO 2	Apply concepts of probability, conditional probability, and independence of events.	3
CO 3	Use Bayes' theorem to compute probabilities in real situations.	3
CO 4	Define random variables and construct their probability distributions.	1
CO 5	Analyze and apply standard discrete probability distributions in real-life contexts.	4
CO 6	Evaluate limiting distributions and their applications.	5

Course Contents:

Unit 1	Probability	08L
	1.1 Definitions of - (i) Sample space, (ii) Discrete sample space (finite and countably infinite) (iii) Continuous sample space, (iv) Events, (v) Elementary event, (vi) Complement of an event, (vii) Certain event, (viii) Impossible event. etc. 1.2 Concept of occurrence of an event, Algebra of events. 1.3 Classical definition of probability and its limitations. 1.4 Probability model, equiprobable and non-equiprobable sample space. 1.5 Axiomatic definition of probability, theorems on probability, computation of probability of an event. 1.6 Conditional probability and Independence of two events. 1.7 Pair-wise independence and mutual independence for three events. Multiplication theorem for two events. 1.8 Partition of the sample space. Bayes' theorem (Statement only), and its applications.	
Unit 2	**Univariate Probability Distributions**	05 L
	Concept and definition of a discrete random variable. 2.2 Probability mass function (p.m.f.) and distribution function (d.f.), FX(·), of discrete random variable, Characteristic properties of distribution function. 2.3 Mode and median of a univariate discrete probability distribution. Definition of expectation (mean) of a random variable its properties, moment generating function (m.g.f.) and cumulant generating function(c.g.f.). 2.4 Definitions of variance and its properties. standard deviation (s.d.) and coefficient of variation (c.v.) of univariate probability distribution Definition of raw, central moments and factorial raw moments of univariate probability distributions and their interrelations. Examples and problem	
Unit 3	**Discrete Probability Distributions based on finite sample space**	10L
	3.1 Degenerate distribution (one point distribution) P.m.f.: P(X=c) =1, mean and variance. 3.2 Uniform discrete distribution on integers 1 to n: p.m.f. $P(X=x) = \begin{cases} \frac{1}{n} & , x = 1, 2, \cdots, n. \\ 0 & \text{otherwise}. \end{cases}$ mean, variance, M.G.F, distribution of sum of two discrete uniform random variable. variance, real life situations, comments on mode andmedian. 3.3 Bernoulli distribution: p.m.f.	

	$P(X=x)=\begin{cases} p^x(1-p)^{1-x} & ,x=0,1 \\ & 0<p<1 \\ 0 & \text{otherwise} \end{cases}$ mean, variance, M.G.F. and moments. Distribution of sum of two independent Bernoulli random variable. 3.4 Binomial distribution: p.m.f. $P(X=x)=\begin{cases} \binom{n}{x} p^x q^{(n-x)} & ,x=0,1,\ldots,n. \\ & ,0<p<1 \quad ,q=1-p \\ 0 & \text{otherwise}. \end{cases}$ Notation: X ~B(n,p). , situations where this distribution is applicable. mean, variance, recurrence relation for successive probabilities, computation of probabilities of different events, mode of the distribution, m.g.f., moments, skewness (comments when p = 0.5, p > 0.5, p < 0.5), additive property (statements only) 3.5 Hypergeometric Distribution: p.m.f $P(X=x)=\begin{cases} \dfrac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} & ,x=0,1,\cdots\cdots \min\{n,M\}, N>M \\ 0 & \text{otherwise} \end{cases}$ Notation: X ~H (N, M, n), situations where this distribution is applicable 3.6 Mean and variance, binomial approximation to hypergeometric distribution (Statement only). Computation of probability, 3.7 Examples and problems.	
Unit 4	**Univariate discrete probability distributions based on countably infinite sample space.**	07 L
	4.1 **Poisson distribution**: Notation: X ~P(λ). $p(x)=\begin{cases} \dfrac{e^{-\lambda}\lambda^x}{x!} & , \quad x=0,1,2\ldots\ldots, \quad \lambda>0 \\ 0 & \text{otherwise} \end{cases}$ Situations where this distribution is applicable. Mean, variance. m.g.f, moments, skewness, kurtosis, recurrence relation, conditional distribution X | X + Y , additive property, Poisson distribution as a limiting form of binomial distribution (Statement only). 4.2 **Geometric distribution**: Notation: X ~G(p), Geometric distribution on support (0, 1, 2…) , with p.m.f.	

$$p(x) = \begin{cases} pq^x, & x = 0, 1, 2\ldots, 0 < p < 1,\ q = 1 - p. \\ 0 & \text{otherwise} \end{cases}$$

Geometric distribution on support (1,2,...)with p.m.f.

$$p(x) = \begin{cases} pq^{x-1}, & x = 1, 2\ldots \quad 0 < p < 1,\ q = 1 - p. \\ 0 & \text{otherwise} \end{cases}$$

distribution function, recurrence relation, situations where this distribution is applicable. Mean, variance. m.g.f , moments , lack of memory property. Examples and problems

4.3 **Negative Binomial Distribution**:

Probability mass function (p.m.f.)

$$P(X = x) = \binom{x+k-1}{x} p^k q^x \quad ; x = 0,1,2, \ldots \ ; 0 < p < 1 \ ; q = 1 - p \ ; k > 0$$
$$= 0 \quad ; \quad otherwise.$$

Notation: $X \sim N(k, p)$.

Nature of p. m. f. negative binomial distribution as waiting time distribution. M.G.F. (State the formula only)., mean, variance, Relation between geometric and negative binomial distribution. Poisson approximation to negative binomial distribution (Statement only). Real life situations.

References:

1. Gupta, S. C. and Kapoor, V. K. (2000). Fundamentals of Mathematical Statistics, 10th Edition, Sultan Chand and Sons Publishers, New Delhi.
2. Agarwal B. L. (2003). Programmed Statistics, 2nd edition, New Age International Publishers, New Delhi.
3. B.L.S. Prakasarao, (2008). A First Course in Probability and Statistics, World Scientific Publishing Company.
4. Ross S. (2002). A First Course in Probability, 6th edition, Pearson Education, Inc. & Dorling Kindersley Publishing, Inc.
5. *Roussas*, George G. (2016). First course in mathematical *statistics*.2nd edition *Publisher:* Academic Press.
6. Goon, A. M., Gupta, M. K. and Dasgupta, B. (2016). Fundamentals of Statistics, Vol. 1, 6th Revised Edition, The World Press Pvt. Ltd., Calcutta.
7. Mukhopadhyay P. (2015). Applied Statistics, *Publisher*: Books & Allied (*P*) Ltd.

Progressive Education Society's
Modern College of Arts, Science and Commerce,
Shivajinagar, Pune – 5

First Year of B.Sc. Data Science
(Course under NEP 2025-26)

Course Code: 24CsDscU2202
Course Name: Lab Course on 24CsDscU2201

Teaching Scheme: TH: 4 Hours/Week **Credit: 2**
Examination Scheme: CIA: 20 Marks **End-Sem: 30 Marks**

Pre-requisites: Knowledge of the topics in Major theory (section I and II)

- **Course outcome:**

CO No	Course Outcomes (COs)	Bloom's Cognitive Level
CO 1	Solve problems on permutations, combinations, and probability.	3
CO 2	Apply conditional probability and independence in problem solving.	3
CO 3	Fit and evaluate binomial, Poisson, and negative binomial distributions.	3
CO 4	Apply probability distributions to real-life scenarios.	4
CO 5	Simulate and model sampling from discrete distributions.	4
CO 6	Develop mini-projects using probability distribution concepts.	6

Sr.No.	Title of Experiment/ Practical
1	Problems on permutations and combinations- I
2	Problems on permutations and combinations- II
3	Problems on probability and conditional probability I
4	Problems on probability and conditional probability II
5	Problems of independence and Bayes' theorem.
6	Fitting of binomial distribution and computation of expected frequencies.
7	Fitting of Poisson distribution and computation of expected frequencies.
8	Fitting of negative binomial distribution and computation of expected frequencies
9	Applications of Bernoulli and binomial distribution.

10	Applications of uniform and Poisson distribution.
11	Applications of hypergeometric distribution.
12	Applications of negative binomial and geometric distribution.
13	Model sampling from binomial, Poisson, Discrete Uniform distribution.
14	Model sampling from Geometric and Negative Binomial distribution.
15	Project