Anekant Education Society's

Tuljaram Chaturchand College of Arts, Science and Commerce, Baramati

(Empowered Autonomous)

Four Year B.Sc. Degree Program in Statistics

(Faculty of Science & Technology)

CBCS Syllabus

S.Y.B.Sc. (Statistics) Semester -IV

Department of Statistics

Tuljaram Chaturchand College of Arts, Science and Commerce, Baramati

Choice Based Credit System Syllabus (2023 Pattern)

(As Per NEP 2020)

To be implemented from Academic Year 2024-2025

Anekant Education Society's

Tuljaram Chaturchand College, Baramati

(Empowered Autonomous)

Board of Studies (BoS) in Statistics

From 2022-23 to 2024-25

Sr. No.	Name	Designation
1.	Prof. Dr. Vikas C. Kakade	Chairman
2.	Prin. Dr. Avinash S. Jagtap	Member
3.	Dr. Neeta K. Dhane	Member
4.	Dr. Vaishali V. Patil	Member
5.	Mrs. Sarita D. Wadkar	Member (Ad hoc)
6.	Dr. Chandrashekhar P. Swami	Member
7.	Dr. Priti S. Malusare	Member (Ad hoc)
8.	Dr. Nilambari A. Jagtap	Member (Ad hoc)
9.	Miss. Kalyani C. Kale	Member (Ad hoc)
10.	Dr. Pooja S. Gaikwad	Member (Ad hoc)
11.	Dr. Akanksha S. Kashikar	Vice-Chancellor Nominee
12.	Prin. Dr. Rajendra G. Gurao	Expert from other University
13.	Dr. Rohan Koshti	Expert from other University
14.	Mr. Saurabh Kadam	Industry Expert
15.	Dr. Jaya L. Limbore	Meritorious Alumni
16.	Dr. Trupti Arekar	Invitee Member
17.	Miss. Priya N. Rakate	Invitee Member
18.	Ms. Shital B. Choudhar	Invitee Member
19.	Ms. Tejashri D. Kawade	Invitee Member
20.	Miss. Rupali S. Kale	Invitee Member
21.	Miss. Ankita S. Yadav	Invitee Member

22.	Miss. Birnale Sneha Sanjay (M. Sc. Part- II)	Student Representative
23.	Miss. Prabhune Utkarsha Shrinivas (M. Sc. Part- II)	Student Representative
24.	Miss. Chavan Vaishnavi Ajit (TYBSc)	Student Representative
25.	Miss. Gawade Vaishnavi Tukaram (TYBSc)	Student Representative

Anekant Education Society's
Tuljaram Chaturchand College of Arts, Science & Commerce, Baramati.
(Empowered Autonomous)

Tuljaram Chaturchand College of Arts, Science & Commerce, Baramati is an empowered autonomous & dynamic institute and has successfully implemented the National Education Policy-2020 since the academic year 2023-24. We are updating our academic policies as per local needs keeping in view the global perspectives. Accordingly, we have updated our program outcomes as per the graduate attributes defined in New Education Policy. In general, program outcomes are categorized into two categories as disciplinary & interdisciplinary outcomes and generic outcomes.

Program Outcomes for B.Sc.

PO.1. **Comprehensive Knowledge and Understanding**: Graduates will possess a profound understanding of their field of study, including foundational theories, principles, methodologies, and key concepts, within a broader multidisciplinary context.

PO.2. **Practical, Professional, and Procedural Knowledge**: Graduates will acquire practical skills and expertise essential for professional tasks within their field. This includes knowledge of industry standards, best practices, regulations, and ethical considerations, with the ability to apply this knowledge effectively in real-world scenarios.

PO.3. **Entrepreneurial Mindset and Knowledge**: Graduates will cultivate an entrepreneurial mindset, identifying opportunities, fostering innovation, and understanding business principles, market dynamics, and risk management strategies.

PO.4. **Specialized Skills and Competencies**: Graduates will demonstrate proficiency in technical skills, analytical abilities, problem-solving, effective communication, and leadership, relevant to their field of study. They will also adapt and innovate in response to changing circumstances.

PO.5. **Capacity for Application, Problem-Solving, and Analytical Reasoning**: Graduates will possess the capacity to apply learned concepts in practical settings, solve complex problems, and analyze

data effectively. This requires critical thinking, creativity, adaptability, and a readiness to learn and take calculated risks.

PO.6. **Communication Skills and Collaboration**: Graduates will effectively communicate complex information, both orally and in writing, using appropriate media and language. They will also collaborate effectively in diverse teams, demonstrating leadership qualities and facilitating cooperative efforts toward common goals.

PO.7. **Research-related Skills**: Graduates will demonstrate observational and inquiry skills, formulate research questions, and utilize appropriate methodologies for data collection and analysis. They will also adhere to research ethics and effectively report research findings.

PO.8. **Learning How to Learn Skills**: Graduates will acquire new knowledge and skills through self-directed learning, adapt to changing demands, and set and achieve goals independently.

PO.9. **Digital and Technological Skills**: Graduates will demonstrate proficiency in using ICT, accessing information sources, and analyzing data using appropriate software.

PO.10. **Multicultural Competence, Inclusive Spirit, and Empathy**: Graduates will engage effectively in multicultural settings, respecting diverse perspectives, leading diverse teams, and demonstrating empathy and understanding of others' perspectives and emotions.

PO.11. **Value Inculcation and Environmental Awareness**: Graduates will embrace ethical and moral values, practice responsible citizenship, recognize and address ethical issues, and take appropriate actions to promote sustainability and environmental conservation.

PO.12. **Autonomy, Responsibility, and Accountability**: Graduates will apply knowledge and skills independently, manage projects effectively, and demonstrate responsibility and accountability in work and learning contexts.

PO.13. **Community Engagement and Service**: Graduates will actively participate in community-engaged services and activities, promoting societal well-being.

Credit Distribution Structure for F.Y.B.Sc.-2023-2024 (Statistics)

Level	Semester	Major		Minor	GE/OE	VSC, SEC (VSEC)	AEC, VEC, IKS	OJT, FP, CEP, CC, RP	Cum. Cr./ Sem.	Degree/ Cum. Cr.
		Mandatory	Electives							
4.5	I	STA-101-MJM: Descriptive Statistics – I STA-102-MJM: Discrete Probability and Probability Distributions – I STA-103-MJM: Statistics Practical – I Credits-2+2+2	--	--	STA-116-OE: Commercial Statistics STA-117-OE: Introduction to MS-Excel and Statistical Computing Credit- 2+2	STA-121-VSC: Introduction to R Programming Language STA-126-SEC: Statistical Computing Using MS-Excel Credit- 2+2	ENG-131-AEC: Functional English-I STA-137-IKS: Evaluation of Science and Statistics in India EVS-135-VEC: Environmental Science Credit- 2+2+2	CC1: To be selected from the Basket Credit- 2	22	UG Certificate 44
	II	STA-151-MJM: Descriptive Statistics – II STA-152-MJM: Discrete Probability and Probability Distributions – II STA -153-MJM: Statistics Practical – II Credits-2+2+2	--	STA-161-MN: Basic Statistics Credits-2	STA-166-OE: Business Statistics STA-167-OE: Statistics Learning with Software Credit- 2+2	STA-171-VSC: - Data Analysis with R Software STA-176-SEC: Application of Statistics Using Advanced Excel Credit- 2+2	ENG-181-AEC: Functional English-II COS-185-VEC: Digital and Technological Solutions Credit- 2+2	**CC2:** To be selected from the Basket Credit- 2	22	
	Cum Cr.	12	--	2	8	8	10	4	44	

Credit Distribution Structure for S.Y.B.Sc.-2024-2025 (Statistics)

Level	Semester	Major		Minor	GE/OE	VSC, SEC (VSEC)	AEC, VEC, IKS	OJT, FP, CEP, CC, RP	Cum. Cr./ Sem.	Degree/ Cum. Cr.
		Mandatory	Electives							
5.0	III	STA-201-MJM: Continuous Probability Distributions – I (T) STA-202-MJM: Statistical Techniques- I(T) STA-203-MJM: Applied Statistics-I(T) STA-204-MJM Statistics Practical-III (P) **Credits-2+2+2+2**	--	STA-211-MN: Foundations of Probability: Theory and Applications (T) STA-212-MN: Minor Statistics Practical – I (P) **Credit-2+2**	STA-216-OE: Applied Statistical Techniques (T) **Credit- 2**	STA-221-VSC: Quantitative Techniques (T) **Credit- 2**	MAR-231-AEC: भाषिक उपयोजन व लेखन कौशल्ये (T) **Or** HIN-231-AEC: हिंदी भाषा : सृजन कौशल (T) **Or** SAN-231-AEC: प्राथमिक संभाषणकौशल्यम् (T) GEN-245-IKS: Indian Knowledge System (Generic) (T) **Credit- 2+2**	STA-235-FP: Project YOG/PES/CUL/NSS/NCC-239-CC: (T) To be selected from the Basket **Credit- 2+2**	--	**UG Diploma 46**
	IV	STA-251-MJM Continuous Probability Distributions – II (T) STA-252-MJM Statistical Techniques- II (T) STA-253-MJM Applied Statistics-II (T) STA -254-MJM: Statistics Practical – IV(P) **Credits-2+2+2+2**	--	STA-261-MN: Probability Distributions and Applications (T) STA-262-MN: Minor Statistics Practical – II (P) **Credit-2+2**	STA-266-OE: Practical Based on Applied Statistical Techniques (P) **Credit- 2**	STA-276-SEC: Programming in R and Introduction to Tableau, Power BI (P) **Credit- 2**	MAR-231-AEC: लेखन निर्मिती व परीक्षण कौशल्ये (T) **Or** HIN-231-AEC: हिंदी भाषा: संप्रेषण कौशल (T) **Or** SAN-231-AEC: प्रगत संभाषणकौशल्यम् (T) **Credit- 2**	YOG/PES/CUL/NSS/NCC-289-CC: (T) To be selected from the Basket **Credit- 2**	STA-285-CEP: Community Engagement Project (P) **Credit-2**	
	Cum Cr.	16	--	8	4	4	6	6	2	

Course Structure for F.Y.B.SC. Statistics (2023 Pattern)

Sem	Course Type	Course Code	Course Name	Theory / Practical	Credits
I	Major Mandatory	STA-101-MJM	Descriptive Statistics – I	Theory	02
	Major Mandatory	STA-102-MJM	Discrete Probability and Probability Distributions – I	Theory	02
	Major Mandatory	STA-103-MJM	Statistics Practical – I	Practical	02
	Open Elective (OE)	STA-116-OE	Commercial Statistics	Theory	02
	Open Elective (OE)	STA-117-OE	Introduction to MS-Excel and Statistical Computing	Practical	02
	Vocational Skill Course (VSC)	STA-121-VSC	Introduction to R Programming Language	Theory	02
	Skill Enhancement Course (SEC)	STA-126-SEC	Statistical Computing Using MS-Excel	Practical	02
	Ability Enhancement Course (AEC)	ENG-131-AEC	Functional English-I	Theory	02
	Value Education Course (VEC)	ENV-135-VEC	Environmental Science	Theory	02
	Indian Knowledge System (IKS)	STA-137-IKS	Evolution of Science and Statistics in India	Theory	02
	Co-curricular Course (CC)	--	To be selected from the Basket	Theory	02
				Total Credits Semester-I	**22**
II	Major Mandatory	STA-151-MJM	Descriptive Statistics – II	Theory	02
	Major Mandatory	STA-152-MJM	Discrete Probability and Probability Distributions – II	Theory	02
	Major Mandatory	STA-153-MJM	Statistics Practical – II	Practical	02
	Minor	STA-161-MN	Fundamental of Statistics	Theory	02
	Open Elective (OE)	STA-166-OE	Business Statistics	Theory	02
	Open Elective (OE)	STA-167-OE	Statistics Learning with Software	Practical	02
	Vocational Skill Course (VSC)	STA-171-VSC	Data Analysis with R Software	Practical	02
	Skill Enhancement Course (SEC)	STA-176-SEC	Application of Statistics Using Advanced Excel	Practical	02
	Ability Enhancement Course (AEC)	ENG-181-AEC	Functional English-II	Theory	02
	Value Education Course (VEC)	COS-185-VEC	Digital and Technological Solutions	Theory	02
	Co-curricular Course (CC)	--	To be selected from the Basket	Theory	02
				Total Credits Semester-II	**22**
				Cumulative Credits Semester I + Semester II	**44**

Course Structure for S.Y.B.SC. Statistics (2023 Pattern)

Sem	Course Type	Course Code	Course Name	Theory / Practical	Credits
III	Major Mandatory	STA-201-MJM	Continuous Probability Distributions – I	Theory	02
	Major Mandatory	STA-202-MJM	Statistical Techniques – I	Theory	02
	Major Mandatory	STA-203-MJM	Applied Statistics – I	Theory	02
	Major Mandatory	STA-204-MJM	Statistics Practical – III	Practical	02
	Minor	STA-211-MN	Foundations of Probability: Theory and Applications	Theory	02
	Minor	STA-212-MN	Minor Statistics Practical – I	Practical	02
	Open Elective (OE)	STA-216-OE	Applied Statistical Techniques	Theory	02
	Vocational Skill Course (VSC)	STA-221-VSC	Quantitative Techniques	Theory	02
	Ability Enhancement Course (AEC)	MAR-231-AEC HIN-231-AEC SAN-231-AEC	भाषिक उपयोजन व लेखन कौशल्ये हिंदी भाषा कौशल प्राथमिक संभाषणकौशल्यम्	Theory	02
	Field Project (FP)	STA-235-FP	Project	Practical	02
	Co-curricular Course (CC)	YOG/PES/CUL/NSS/NCC-239-CC	To be selected from the Basket	Theory	02
	Generic IKS Course (IKS)	GEN-245-IKS	Indian Knowledge System (Generic)	Theory	02
				Total Credits Semester-II	**24**
IV	Major Mandatory	STA-251-MJM	Continuous Probability Distributions – II	Theory	02
	Major Mandatory	STA-252-MJM	Statistical Techniques – II	Theory	02
	Major Mandatory	STA-253-MJM	Applied Statistics – II	Theory	02
	Major Mandatory	STA-254-MJM	Statistics Practical – IV	Practical	02
	Minor	STA-261-MN	Probability Distributions and Applications	Theory	02
	Minor	STA-262-MN	Minor Statistics Practical – II	Practical	02
	Open Elective (OE)	STA-266-OE	Practical Based on Applied Statistical Techniques	Practical	02
	Skill Enhancement Course (SEC)	STA-276-SEC	Programming in R and Introduction to Tableau, Power BI	Practical	02
	Ability Enhancement Course (AEC)	MAR-281-AEC HIN-281-AEC SAN-281-AEC	लेखन निर्मिती व परीक्षण कौशल्ये हिंदी भाषा: संप्रेषण कौशल प्रगत संभाषणकौशल्यम्	Theory	02
	Community Engagement Project (CEP)	STA-285-CEP	Community Engagement Project	Practical	02
	Co-curricular Course (CC)	YOG/PES/CUL/NSS/NCC-289-CC	To be selected from the Basket	Theory	02
				Total Credits Semester-IV	**22**
				Cumulative Credits Semester III + Semester IV	**46**

CBCS Syllabus as per NEP 2020 for S.Y.B.Sc. Statistics (2023 Pattern)

Name of the Programme : B.Sc. Statistics

Programme Code : USST

Class : S.Y.B.Sc.

Semester : IV

Course Type : Major Mandatory (Theory)

Course Code : STA-251-MJM

Course Title : Continuous Probability Distributions – II

No. of Credits : 02

No. of Teaching Hours : 30

Course Objectives:

1. To acquaint students with the Exact Sampling Distributions and their applications.
2. To understand concept of some continuous distributions with real life situations.
3. To find various measures of random variable and probabilities using its probability distributions.
4. To know the relations among the different distributions.
5. To understand the concept of transformation of univariate continuous random variables.
6. To study derived distributions and their applications.
7. To apply testing of hypothesis in real life situations.

Course Outcomes:

Students should be able to:

CO1. understand Chi-Square distribution, Student's t- distribution, Snedecor's F distribution.

CO2. compute means, mode, variance, moments, cumulants for above Distributions.

CO3. apply Exact Sampling Distributions.

CO4. know the relations among the different distributions.

CO5. learn exponential and Gamma distributions and its applications.

CO6. learn the concept of transformation of continuous random variables which help to study derived distributions.

CO7. understand the interrelations among, χ^2, t and F variates.

Topics and Learning Points

UNIT 1: Exponential Distribution **(06 L)**

1.1 Probability density function (p. d. f.) $f(x)=\begin{cases}\alpha e^{-\alpha x}; x\geq 0; \alpha>0\\ 0 \quad ; \quad otherwise\end{cases}$

Notation : X ~ Exp(α)

1.2 Nature of p. d. f., density curve, interpretation of α as rate and 1/α as mean, variance,

1.3 M.G.F., C.G.F., c.d.f., graph of c.d.f., lack of memory property, median, quartiles.

1.4 Distribution of min(X, Y) with X, Y i. i. d. exponential r. v. s.

UNIT 2: Gamma Distribution: **(06 L)**

2.1 Probability density function (p. d. f.) $f(x)=\begin{cases}\frac{\alpha^{\lambda}}{\Gamma\lambda}x^{\lambda-1}e^{-\alpha x} & ; x\geq 0;\ \alpha>0,\lambda>0\\ 0 & ; Otherwise\end{cases}$

Notation : $X \sim G(\alpha, \lambda)$.

2.2 Nature of probability curve, special cases: i) α=1 , ii) λ= 1,

2.3 M.G.F., C.G.F., moments, cumulants, β_1, β_2, γ_1, γ_2, mode, additive property.

2.4 Distribution of sum of n i. i. d. Gamma variates.

UNIT 3: Chi-square (χ_n^2) Distribution: **(08 L)**

3.1 Definition of χ^2 r. v. as sum of squares of i.i.d. standard normal variables, derivation of p.d.f. of χ^2 with n degrees of freedom (d.f.) using M.G.F., nature of p.d.f. curve, computations of probabilities using tables of χ^2 distribution. mean, variance, M.G.F., C.G.F., central moments, β_1, β_2, γ_1, γ_2, mode, additive property.

3.2 Normal approximation: $\frac{\chi_n^2-n}{\sqrt{2n}}$ with proof.

3.3 Distribution of $\frac{X}{X+Y}$ and $\frac{X}{Y}$, where X and Y are two independent chi-square random variables.

UNIT 4: Student's t-distribution: **(05 L)**

4.1 Definition of T r. v. with n d.f. in the form $\frac{U}{\sqrt{\chi_n^2/n}}$ where U $\rightarrow$ N(0, 1) and χ_n^2 is a χ^2 r. v. with n d.f. and U and χ_n^2 are independent r.v.s.

4.2 Derivation of p. d. f., nature of probability curve, mean, variance, moments, mode, use of tables of t-distribution for calculation of probabilities, statement of normal approximation.

UNIT 5: Snedecore's F-distribution: (05 L)

5.1 Definition of F r.v. with n_1 and n_2 d.f. as $F_{n_1,n_2} = \frac{\chi^2_{n_1}/n_1}{\chi^2_{n_2}/n_2}$ where $\chi^2_{n_1}$ and $\chi^2_{n_2}$ are independent chi-square r.v.s. with n_1 and n_2 d.f. respectively.

5.2 Derivation of p.d.f., nature of probability curve, mean, variance, moments, mode.

5.3 Distribution of $1/F_{n_1,n_2}$, use of tables of F-distribution for calculation of probabilities.

5.4 Interrelations among, χ^2, t and F variates.

References:

1. Barlow R. E. and Proschan Frank: Statistical Theory of Reliability and Life Testing. Holt Rinebart and Winston Inc., New Yark.
2. Sinha S. K.: Reliability and Life Testing, Second Edition, Wiley Eastern Publishers, New Delhi.
3. Parimal Mukhopadhyaya: An Introduction to the Theory of Probability. World Scientific Publishing.
4. Hogg R.V. and Criag A.T.: Introduction to Mathematical Statistics (Third edition), Macmillan Publishing, New York.
5. Gupta S. C. &Kapoor V.K: Fundamentals of Mathematical Statistics. Sultan Chand & sons, New Delhi.
6. Brase C. H. and Brase C. P. (2018), Understandable Statistics, Twelfth Edition, Cengage Learning, Biston
7. Moor D. S., Notz W. I., Flinger M. A., (2013), The Basic Practice of Statistics Sixth Edition, Freeman and Company New York

COs and POs Mapping

Course Outcomes	Programme Outcomes (POs)												
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PO13
CO1	3	2	1	3	3	2	3	2	2	1	1		
CO2	3	2	1	3	3	2	3	2	3	1	1		
CO3	3	3	1	3	3	2	3	2	3	1	1		
CO4	3	2	1	3	3	2	3	2	2	1	1		
CO5	3	2	1	3	2	1	3	2	2	1	1		
CO6	3	2	1	3	3	2	3	3	3	1	1		
CO7	3	2	1	3	3	2	3	2	2	1	1		
CO8	3	2	1	3	3	2	3	2	2	1	1		

CO9	3	2	1	3	3	2	3	2	3	1	1		

PO-CO Mapping and Justification

PO1: Comprehensive Knowledge and Understanding

CO1 (3): Strongly related because understanding Chi-Square, t, and F distributions involves mastering foundational theories in statistics.

CO2 (3): Strongly related because computing means, variance, moments, and cumulants requires deep knowledge of statistical distributions.

CO3 (3): Strongly related because applying exact sampling distributions requires a solid understanding of their foundational principles.

CO4 (3): Strongly related because knowing the relations among different distributions enhances comprehensive statistical knowledge.

CO5 (3): Strongly related as learning exponential and Gamma distributions, along with their applications, is essential for deep understanding of probability theory.

CO6 (3): Strongly related because understanding the transformation of continuous random variables contributes to advanced statistical knowledge.

CO7 (3): Strongly related as understanding the interrelations among $\chi 2$, t, and F variates is fundamental in theoretical statistics.

PO2: Practical, Professional, and Procedural Knowledge

CO1 (2): Moderately related because Chi-Square, t, and F distributions are used in practical applications like hypothesis testing.

CO2 (2): Moderately related because computing distribution parameters has direct relevance to professional tasks in data analysis.

CO3 (3): Strongly related because applying exact sampling distributions is essential in professional and real-world data analysis.

CO4 (2): Moderately related because understanding the relationships among distributions helps in choosing appropriate methods for practical applications.

CO5 (2): Moderately related because exponential and Gamma distributions are applied in reliability analysis, which is relevant to professional work.

CO6 (2): Moderately related as transformation of variables is important for practical problem-solving in data analysis.

CO7 (2): Moderately related because understanding the interrelations among distributions aids in applying appropriate tests in professional settings.

PO3: Entrepreneurial Mindset and Knowledge

CO1 (1): Partially related as Chi-Square, t, and F distributions don't directly promote entrepreneurial thinking but help with data-driven decision making.

CO2 (1): Partially related as computing statistical measures is more technical than entrepreneurial.

CO3 (1): Partially related as applying exact sampling distributions supports data analysis but doesn't foster entrepreneurial innovation directly.

CO4 (1): Partially related because understanding distribution relationships is more technical than entrepreneurial.

CO5 (1): Partially related as knowledge of exponential and Gamma distributions helps in business decision-making but not directly entrepreneurial.

CO6 (1): Partially related because transformations enhance technical problem-solving rather than entrepreneurial mindset.

CO7 (1): Partially related as interrelations among distributions are useful for technical decision-making rather than entrepreneurial innovation.

PO4: Specialized Skills and Competencies

CO1 (3): Strongly related because understanding these distributions is a specialized skill in statistics.

CO2 (3): Strongly related because computing moments, variance, and cumulants involves advanced technical skills.

CO3 (3): Strongly related because applying exact sampling distributions is an advanced skill required for specialized data analysis.

CO4 (3): Strongly related because understanding distribution interrelations is a specialized analytical competency.

CO5 (3): Strongly related as learning exponential and Gamma distributions is key for specialized analysis in fields like actuarial science or reliability.

CO6 (3): Strongly related because mastering transformations and derived distributions is a key technical competency.

CO7 (3): Strongly related because understanding the interrelations among $\chi 2$, t, and F variates is crucial for advanced statistical analysis.

PO5: Capacity for Application, Problem-Solving, and Analytical Reasoning

CO1 (3): Strongly related because applying Chi-Square, t, and F distributions in hypothesis testing requires strong analytical reasoning.

CO2 (3): Strongly related because computing means, variance, and moments requires problem-solving in statistical analysis.

CO3 (3): Strongly related as applying exact sampling distributions is key in solving real-world problems using statistical methods.

CO4 (3): Strongly related because understanding distribution relationships helps solve complex statistical problems.

CO5 (2): Moderately related as exponential and Gamma distributions have practical applications in real-world problems.

CO6 (3): Strongly related because transformations of continuous random variables are critical for solving problems involving derived distributions.

CO7 (3): Strongly related as understanding interrelations among $\chi 2$, t, and F distributions helps in solving complex statistical problems.

PO6: Communication Skills and Collaboration

CO1 (2): Moderately related because effective communication of statistical results requires understanding distributions like Chi-Square and t.

CO2 (2): Moderately related because interpreting and communicating statistical measures requires clarity.

CO3 (2): Moderately related as applying exact sampling distributions requires collaboration in research settings.

CO4 (2): Moderately related because understanding relationships among distributions supports clear communication of statistical concepts.

CO5 (1): Partially related because exponential and Gamma distributions are more technical and less related to direct communication skills.

CO6 (2): Moderately related because understanding transformations aids in clear communication of complex results.

CO7 (2): Moderately related as explaining interrelations among distributions requires strong communication skills.

PO7: Research-related Skills

CO1 (3): Strongly related because these distributions are fundamental in statistical research and hypothesis testing.

CO2 (3): Strongly related because computing statistical parameters is crucial for research analysis.

CO3 (3): Strongly related as applying exact sampling distributions is essential in statistical research methodologies.

CO4 (3): Strongly related because understanding relationships among distributions is critical for advanced research in statistics.

CO5 (3): Strongly related as exponential and Gamma distributions are widely used in various research fields.

CO6 (3): Strongly related because transformations are key in research for deriving new distributions and models.

CO7 (3): Strongly related because understanding interrelations among $\chi 2$, t, and F variates is essential for advanced research.

PO8: Learning How to Learn Skills

CO1 (2): Moderately related as learning these distributions helps students adapt to new statistical challenges.

CO2 (2): Moderately related because learning to compute statistical parameters fosters adaptive learning.

CO3 (2): Moderately related because applying sampling distributions encourages continuous learning in statistics.

CO4 (2): Moderately related because learning distribution relationships fosters adaptive thinking.

CO5 (2): Moderately related as understanding exponential and Gamma distributions involves continuous learning of new applications.

CO6 (3): Strongly related because mastering transformations requires adaptive learning to new statistical methods.

CO7 (2): Moderately related because understanding interrelations encourages further exploration in statistics.

PO9: Digital and Technological Skills

CO1 (2): Moderately related as technology is often used to apply Chi-Square, t, and F distributions.

CO2 (3): Strongly related because computing statistical measures often involves using statistical software.

CO3 (3): Strongly related because applying exact sampling distributions requires technological tools for data analysis.

CO4 (2): Moderately related because understanding relationships among distributions supports using statistical software.

CO5 (2): Moderately related as exponential and Gamma distributions are often applied using technology in data analysis.

CO6 (3): Strongly related because transformations of variables often require the use of software for implementation.

CO7 (2): Moderately related because understanding distribution interrelations can be enhanced using technological tools.

PO10: Multicultural Competence, Inclusive Spirit, and Empathy

CO1 (1): Partially related because statistical distributions are applied universally but not directly linked to multicultural competence.

CO2 (1): Partially related because computing statistical measures does not directly enhance multicultural competence.

CO3 (1): Partially related as exact sampling distributions are technical and not closely related to multicultural contexts.

CO4 (1): Partially related because understanding distribution relationships is more technical than multicultural.

CO5 (1): Partially related as exponential and Gamma distributions are more relevant to technical fields than multicultural settings.

CO6 (1): Partially related because transformations of variables are not directly linked to multicultural competence.

CO7 (1): Partially related as understanding interrelations among distributions is more about technical skills than empathy.

PO11: Value Inculcation and Environmental Awareness

CO1 (1): Partially related because understanding statistical distributions may involve ethical considerations but not environmental awareness.

CO2 (1): Partially related because computing statistical measures does not directly relate to value inculcation.

CO3 (1): Partially related as applying sampling distributions is technical and not strongly linked to environmental awareness.

CO4 (1): Partially related because understanding distribution relationships is technical rather than ethical.

CO5 (1): Partially related as learning exponential and Gamma distributions is not closely linked to ethical or environmental concerns.

CO6 (1): Partially related because transformations of variables are not directly related to value inculcation.

CO7 (1): Partially related as understanding interrelations among distributions is more about technical

CBCS Syllabus as per NEP 2020 for S.Y.B.Sc. Statistics (2023 Pattern)

Name of the Programme : B.Sc. Statistics

Programme Code : USST

Class : S.Y.B.Sc.

Semester : IV

Course Type : Major Mandatory (Theory)

Course Code : STA-252-MJM

Course Title : Statistical Techniques – II

No. of Credits : 02

No. of Teaching Hours : 30

Course Objectives:

1. To develop the ability to model relationships between multiple variables using multiple linear regression techniques, particularly in trivariate data scenarios.
2. To enable students to understand and apply the core principles of hypothesis testing, including the formulation and interpretation of null and alternative hypotheses.
3. To equip students with the skills to utilize sampling distributions and the Central Limit Theorem for estimating population parameters and conducting hypothesis tests.
4. To provide students with the knowledge and practice needed to apply the Z test for evaluating population means in various statistical contexts.
5. To train students in the application of the Z test for assessing population proportions, fostering accurate decision-making based on statistical evidence.
6. To familiarize students with the use of the t-test, both for independent and paired samples, in testing hypotheses related to population means.
7. To ensure students can apply the F test and chi-square test effectively, for analyzing population variances and testing the independence of categorical data, respectively.

4. To fit the appropriate time series model that can be used.**Course Outcomes:**

Student will be able to:

CO1. fit a multiple linear regression model in case of trivariate data

CO2. explain the fundamental principles of hypothesis testing, including formulating null and alternative hypotheses

CO3. Students will understand and apply sampling distributions to estimate population parameters and perform hypothesis tests, using concepts like the Central Limit Theorem.

CO4. apply Z test for testing population means

CO5. apply Z test for testing population proportions

CO6. apply t test for testing population means

CO7. apply F test for testing population variances

CO8. apply chi square test for testing the independence of two attributes

CO9. apply paired t test for comparing population means

Topics and Learning Points

Unit 1: Multiple Linear Regression Model (trivariate case) (12L)

1.1 Definition of multiple correlation coefficient $R_{i.jk}$ $i,j,k = 1,2,3.$

1.2 Properties of multiple correlation coefficient

1.2.1 $0 \leq R_{i.jk} \leq 1$ $i,j,k = 1,2,3$

1.2.2 $R_{i.jk} \geq Max\{|r_{ij}|, |r_{ik}|, |r_{ij.k}|, |r_{ik.j}|\}$ for $i \neq j \neq k. i,j,k = 1,2,3$

1.3 Interpretation of

1.3.1 coefficient of multiple determination $R^2_{i.jk}$

1.3.2 $R^2_{1.23}$ =1

1.3.3 $R^2_{1.23}$ =0

1.4 Definition of the partial correlation coefficient

1.5 Notion of multiple linear regression Yule's notation $R_{1.23}$

1.6 Fitting of regression plane of Y on X_1 and X_2, $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$, by the method of least squares; obtaining normal equations, solutions of normal equations.

1.7 Residuals: Definition, order, derivation of variance, properties.

1.8 Properties of the partial regression coefficient

1.8.1 $-1 \leq r_{ij.k} \leq 1$ for i,j,k=1,2,3: $i \neq j \neq k$

1.8.2 $b_{12.3} \times b_{21.3} = r^2_{12.3}$

Unit 2: Basic concept of Testing of Hypothesis and Sampling Distributions: (4 L)

2.1 Parameter, random sample from a distribution as i.i.d. r.v.s. $X_1, X_2, \ldots, Xn$, statistic, estimator, estimate, critical region. Statistical hypothesis, null and alternative hypothesis, one sided and two sided alternative hypothesis, p-value. Confidence interval.

2.2 Sampling distribution of a statistic. Distribution of sample mean $\bar{X}$ from normal, exponential and gamma distribution, Notion of standard error of a statistic.

2.3 Distribution of $\frac{nS^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \bar{X})^2$ for a sample from a normal distribution using orthogonal transformation. Independence $\bar{X}$ and S^2.

Unit 3: Large Sample Tests (Tests based on Normal distribution) : (4L)

3.1 Z-tests for population means:

3.1.1 One sample and two sample tests for one-sided and two-sided alternatives

3.1.2 Confidence Interval for Population Mean:100(1- α)% two sided confidence interval for single population mean (μ) and difference of population means $(\mu_1 - \mu_2)$ of two independent normal populations.

3.2 Z-tests for population proportions:

3.2.1 One sample and two sample tests for one-sided and two-sided alternatives

3.2.2 Confidence Interval for Population Proportion: 100(1- α)% two sided confidence interval for single population proportion (P) and difference of population proportions $(P_1 - P_2)$ of two independent normal populations.

Unit 4: Small Sample Tests (Tests based on Normal distribution) : (10L)

4.1 Tests based on Chi-square distribution:

4.1.1 *Test for independence* of two attributes

4.1.2 *Test for Goodness of Fit* (Without rounding off the expected frequencies) (Problems are not expected)

4.1.3 *Test for* $H_0 : \sigma^2 = \sigma_0^2$ against one-sided and two-sided alternatives when mean is known, mean is unknown.

4.2 Tests based on t-distribution: *t-tests for population means*:

4.2.1 One sample and two sample tests for one-sided and two-sided alternatives

4.2.2 Confidence Interval for Population Mean: 100(1- α) % two sided confidence interval for single population mean (μ) and difference of population means $(\mu_1 - \mu_2)$ of two independent normal populations.

4.3 *Paired t-test* for one-sided and two-sided alternatives.

4.4 Test based on F-distribution: Test for H_0: $\sigma_1^2 = \sigma_2^2$ against one-sided and two-sided alternatives when means are known and means are unknown.

References:

1. Goon A. M., Gupta, M. K. and Dasgupta, B. (1986), Fundamentals of Statistics, Vol. 2, World Press, Kolkata.
2. Gupta, S. C. and Kapoor, V. K. (2002), Fundamentals of Mathematical Statistics, (Eleventh Edition), Sultan Chand and Sons, 23, Daryaganj, New Delhi , 110002 .
3. Gupta, S. C. and Kapoor V. K. (2007), Fundamentals of Applied Statistics (Fourth Edition), Sultan Chand and Sons, New Delhi.

4. Gupta S. P. (2002), Statistical Methods (Thirty First Edition), Sultan Chand and Sons, 23, Daryaganj, New Delhi 110002.
5. Parimal Mukhopadhyaya: An Introduction to the Theory of Probability. World Scientific Publishing.
6. Hogg R.V. and Criag A.T.: Introduction to Mathematical Statistics (Third edition), Macmillan Publishing, New York.

COs and POs Mapping:

Course Outcomes	Programme Outcomes (POs)												
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PO13
CO1		2	1	3	2				2			2	
CO2	3					1	2	1					
CO3	3						3						
CO4		2		3	3				2			2	
CO5		2		3	3				2			2	
CO6		2		3	3				2			2	
CO7		2		3	3				2			2	
CO8		2		3	3				2			2	
CO9		2		3	3				2			2	

PO1: Comprehensive Knowledge and Understanding

- **CO2-3**: Explaining the fundamental principles of hypothesis testing is strongly related to understanding foundational theories and concepts in statistics.
- **CO3-3**: Understanding and applying sampling distributions also directly supports a profound understanding of statistical principles.

Justification: CO2 and CO3 strongly contribute to PO1, as they encompass the essential theoretical knowledge required for understanding hypothesis testing and sampling distributions.

PO2: Practical, Professional, and Procedural Knowledge

- **CO1-2:** Fitting a multiple linear regression model is moderately related as it involves applying statistical methods in a practical scenario.
- **CO4, CO5, CO6, CO7, CO8, CO9-2**: Applying various tests (Z, t, F, chi-square, paired t) is moderately related, as these are practical tools used in professional statistical analysis.

Justification: These course outcomes provide students with practical skills essential for conducting statistical analyses in professional contexts, aligning with PO2.

PO3: Entrepreneurial Mindset and Knowledge

- **CO1-1:** Fitting a multiple linear regression model could be partially related, as it aids in decision-making and predicting outcomes, which are relevant in entrepreneurial contexts.

Justification: CO1 can be linked to PO3 as regression analysis is useful in understanding market trends and making informed business decisions, though the connection is partial.

PO4: Specialized Skills and Competencies

- **CO1-3**: Strongly related as it demonstrates proficiency in regression analysis.
- **CO4, CO5, CO6, CO7, CO8, CO9-3**: All of these outcomes are strongly related, as they demonstrate the application of statistical tests, requiring specialized technical skills.

Justification: These COs require specialized analytical abilities and problem-solving skills, making them strongly related to PO4.

PO5: Capacity for Application, Problem-Solving, and Analytical Reasoning

- **CO1-2**: Moderately related as it requires applying knowledge to build regression models.
- **CO4, CO5, CO6, CO7, CO8, CO9-3**: Strongly related as they involve applying tests to solve real-world problems and analyze data.

Justification: These COs are central to developing the ability to apply learned concepts, solve problems, and use analytical reasoning in statistical contexts.

PO6: Communication Skills and Collaboration

- **CO2-1**: Partially related as explaining hypothesis testing requires clear communication of complex concepts.

Justification: CO2 is partially related as it involves explaining and discussing statistical concepts, although it does not directly focus on collaboration or broad communication skills.

PO7: Research-related Skills

- **CO3-3**: Strongly related as it involves understanding and applying sampling distributions, which is fundamental to research in statistics.
- **CO2-2**: Moderately related as hypothesis testing is a core component of research.

Justification: CO2 and CO3 are relevant to PO7, as they involve essential skills for conducting research, such as hypothesis formulation and the use of sampling distributions.

PO8: Learning How to Learn Skills

- **CO2-1**: Partially related as it involves learning and applying new concepts of hypothesis testing.

Justification: This CO involves continuous learning and application of new statistical methods, though it is only partially connected to the broader concept of learning how to learn.

PO9: Digital and Technological Skills

- **CO1-2**: Moderately related as regression analysis often involves using statistical software.
- **CO4, CO5, CO6, CO7, CO8, CO9-2**: Moderately related as these tests are commonly conducted using software.

Justification: The application of these COs often involves the use of digital tools and software, linking them to PO9.

PO10: Multicultural Competence, Inclusive Spirit, and Empathy

- **Not directly related** to any CO.

Justification: The course outcomes do not directly address multicultural competence or related skills.

PO11: Value Inculcation and Environmental Awareness

- **Not directly related** to any CO.

Justification: The course outcomes do not directly address values or environmental awareness.

PO12: Autonomy, Responsibility, and Accountability

- **CO1, CO4, CO5, CO6, CO7, CO8, CO9-2**: Moderately related as these involve taking responsibility for statistical analysis and ensuring accuracy.

Justification: These COs require students to independently conduct analyses and be accountable for their results, aligning with PO12.

PO13: Community Engagement and Service

- **Not directly related** to any CO.

Justification: The course outcomes do not directly engage with community service or related activities.

CBCS Syllabus as per NEP 2020 for S.Y.B.Sc. Statistics (2023 Pattern)

Name of the Programme	**:** B.Sc. Statistics
Programme Code	**:** USST
Class	**:** S.Y.B.Sc.
Semester	**:** IV
Course Type	**:** Major Mandatory (Theory)
Course Code	**:** STA-253-MJM
Course Title	**:** Applied Statistics – II
No. of Credits	**:** 02
No. of Teaching Hours	**:** 30

Course Objectives:

1. Develop a comprehensive understanding of quality management principles and the role of Statistical Process Control (SPC) in maintaining and improving process quality.
2. Apply the concept of 3σ control limits to design and analyze control charts for both variables and attributes, ensuring effective monitoring of process performance.
3. Differentiate between chance causes and assignable causes of variation in manufacturing or service processes, and understand their implications for process stability.
4. Master the techniques for constructing and interpreting control charts for variables and utilize these charts to detect process variations.
5. Gain proficiency in constructing and interpreting control charts for attributes to monitor product quality.
6. Calculate and interpret key process capability indices (e.g., Cp, Cpk) to evaluate the ability of a process to meet specification limits and identify areas for improvement.
7. Evaluate the stability and capability of processes through control charts and capability indices, enabling informed decisions on process improvement and quality assurance.

Course Outcome:

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

CO1. understand the concept of quality, SPC and process control tools ,

CO2. build the 3 σ control limits for variables and attributes control charts

CO3. compare chance causes and assignable causes of variation

CO4. construct and interpret control charts for variables

CO5. construct and interpret control charts for attributes

CO6. calculate and interpret process capability indices

CO7. assess the stability of a process using control charts and evaluate its capability using capability indices

Topics and Learning Points

UNIT – 1: Introduction to SPC **(4L)**

1.1 Meaning and Purpose of Statistical Process Control (SPC), Online process control methods (control charts) and offline process control methods (Sampling plans).

1.2 Seven Process Control (PC) Tools of SPC

(i) Check Sheet, (ii) Cause and effect diagram (CED),(iii) Pareto Diagram, (iv) Histogram, (v) Control chart, (vi)Scatter Diagram,(vii) Design of Experiments (DOE). (Only the introduction of 7 PC tools is expected)

UNIT – 2: Introduction to Control charts **(4L)**

2.1 Chance causes and assignable causes of variation, statistical basis of control charts, exact probability limits, k -sigma limits, justification for the use of 3- sigma limits

2.2 Criteria for detecting lack of control situations:

2.2.1 At least one point outside the control limits

2.2.2 A run of seven or more points above or below the central line.

2.2.3 Presence of a nonrandom pattern eg. cycle or linear trends etc.

2.2.4 Construction of control charts for (i) standards given, (ii) standards not given.

UNIT – 3 Control charts for variables **(8L)**

3.1 R chart and $\overline{X}$ chart: Purpose of R and $\overline{X}$ chart

3.2 Construction of R chart when the process standard deviation is specified: control limits, drawing of control chart, plotting of sample ranges. Drawing conclusion - determination of state of control process, corrective action if the process is out of statistical control.

3.3 Construction of $\overline{X}$ chart when the process average is specified: control limits, drawing of control chart, plotting of sample means. Drawing conclusion - determination of state of control of process, corrective action if the process is out of statistical control.

3.4 Construction of R chart when the process standard deviation (σ) is not given: control limits, drawing of control chart, plotting sample range values, revision of control limits if necessary, estimate of σ for future use.

3.5 Construction of $\bar{X}$ chart when the process average μ is not given : Drawing of control chart, plotting sample means, revision of control limits of $\bar{X}$ chart, if necessary.

UNIT 4: Control charts for Attributes (10L)

4.1 p – chart:

4.1.1 Construction of p-chart *when subgroup sizes are same and the standard is known*: control limits, drawing of control chart, plotting of sample fraction defectives. Determination of state of control of the process.

4.1.2 Construction of p-chart when *subgroup sizes are different* and the value of the process fraction defective P is not specified with separate control limits, drawing of control chart, plotting sample fraction defectives, determination of the state of control of the process. Interpretation of high and low spots. Identification of real-life situations.

4.2 C chart:

4.2.1 Construction of c-chart *when the standard is given*; control limits justification of 3 sigma limits, drawing of control chart, plotting number of defects per unit.

4.2.2 Construction of c chart *when the standard is not given*; control limits, explanation for the use of 3 sigma limits, drawing of control chart. Plotting number of defects per unit. Determination of state of control, interpretation of high and low spots.

UNIT 5: Capability Studies (4L)

Specification limits, natural tolerance limits, and their comparisons, decisions based on these comparisons, and estimate of percent defectives. Capability ratio and capability indices (Cp), capability performance indices Cpk with respect to machine and process, interpretation.

References:

1. Montgomery, D. C. (1983). Statistical Quality Control, John Wiley and Sons, Inc., New York.
2. Besterfield ,D.H. and Michna , C.B. et al. (2009). Total Quality Management, 3rd edition, Pearson Education, Delhi.34

3. Dodge, H.F. and Roming, H.G. Sampling Inspection tables, John Wiley and Sons, Inc. New York
4. Duncan A.J. (1974). Quality Control and Industrial Statistics,fourth edition D.B. Taraporewala Sons and Co. Pvt. Ltd., Mumbai.
5. Grant, E. L. and Leavenworth (1980). Statistical Quality Control, fifth edition, Mc-Graw Hill, New Delhi.
6. Johnson, N.L. and Kotz, S. (1993). Capability Studies, Chapman and Hall Publishers.
7. Brase C. H. and Brase C. P. (2018), Understandable Statistics, Twelfth Edition, Cengage Learning, Biston
8. Moor D. S., Notz W. I., Flinger M. A., (2013), The Basic Practice of Statistics Sixth Edition, Freeman and Company New York

Cos and POs Mapping:

Course Outcomes	Programme Outcomes (POs)												
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PO13
CO1	3			2				2					
CO2	2	3		3	3				2			2	
CO3	2				3		2		2	1			
CO4		3		3	3	2			2				
CO5		3		3	3	2			2				
CO6	2		2	3	3		2						
CO7	3	3	2	3	3		2	2		1	1	2	1

PO1: Comprehensive Knowledge and Understanding

CO1: 3 - Understanding quality, SPC, and process control tools provides a profound understanding of essential concepts in the field of quality management.

CO2: 2 - Building control limits require a solid understanding of statistical methodologies.

CO3: 2 - Differentiating between chance and assignable causes of variation is a key concept in SPC.

CO6: 2 - Calculating and interpreting process capability indices builds on foundational statistical knowledge.

CO7: 3 - Assessing process stability and capability requires a comprehensive understanding of SPC concepts.

PO2: Practical, Professional, and Procedural Knowledge

CO2: 3 - Building 3σ control limits is a practical skill crucial for real-world quality control tasks.

CO4: 3 - Constructing and interpreting control charts for variables directly applies professional knowledge in process control.

CO5: 3 - Constructing and interpreting control charts for attributes is essential for practical quality management.

CO7: 3 - Assessing process stability using control charts is a core practical competency in quality management.

PO3: Entrepreneurial Mindset and Knowledge

CO6: 2 - Calculating and interpreting process capability indices can inform decisions in process improvement and innovation, aligning with an entrepreneurial mindset.

CO7: 2 - Evaluating process capability can help identify opportunities for innovation and improvement in manufacturing or business processes.

PO4: Specialized Skills and Competencies

CO1: 2 - Understanding quality and SPC tools contributes to specialized knowledge in quality management.

CO2: 3 - Building control limits is a specialized skill critical for quality professionals.

CO4: 3 - Constructing and interpreting control charts for variables is a highly specialized skill in SPC.

CO5: 3 - Constructing and interpreting control charts for attributes also represents specialized technical knowledge.

CO6: 3 - Proficiency in calculating and interpreting process capability indices is a specialized technical competency.

CO7: 3 - Assessing process stability and capability requires a high level of specialized skill in SPC.

PO5: Capacity for Application, Problem-Solving, and Analytical Reasoning

CO2: 3 - Building control charts involves strong analytical reasoning and problem-solving skills.

CO3: 3 - Comparing causes of variation requires critical thinking and the ability to solve complex quality problems.

CO4: 3 - Interpreting control charts involves significant problem-solving and application of statistical reasoning.

CO5: 3 - Similarly, interpreting control charts for attributes requires strong analytical skills.

CO6: 3 - Calculating and interpreting capability indices demands the application of analytical reasoning to assess process performance.

CO7: 3 - Evaluating process stability and capability involves complex problem-solving skills.

PO6: Communication Skills and Collaboration

CO4: 2 - Interpreting control charts and communicating findings is essential for collaboration in quality teams.

CO5: 2 - Communicating the results of control charts for attributes is also crucial for team collaboration.

PO7: Research-related Skills

CO3: 2 - Comparing causes of variation involves observational and inquiry skills, which are essential in research.

CO6: 2 - Interpreting process capability indices often requires research-related skills, such as data collection and analysis.

CO7: 2 - Assessing stability and capability using control charts can be part of research and process improvement studies.

PO8: Learning How to Learn Skills

CO1: 2 - Understanding SPC concepts encourages lifelong learning and adaptability to new quality management techniques.

CO7: 2 - Evaluating process stability and capability encourages continuous learning and improvement.

PO9: Digital and Technological Skills

CO2: 2 - Building control limits typically involve the use of statistical software, requiring digital skills.

CO4: 2 - Constructing and interpreting control charts often requires proficiency with digital tools.

CO5: 2 - Similarly, working with control charts for attributes necessitates digital competence.

CO6: 2 - Calculating capability indices frequently involves the use of technology.

PO10: Multicultural Competence, Inclusive Spirit, and Empathy

CO3: 1 - Understanding variation and its causes can involve considering diverse perspectives in quality management.

PO11: Value Inculcation and Environmental Awareness

CO7: 1 - Assessing process capability and stability can include considerations of sustainability and ethical responsibility in process improvement.

PO12: Autonomy, Responsibility, and Accountability

CO2: 2 - Building control charts independently reflects autonomy and responsibility in professional tasks.

CO7: 2 - Evaluating process capability and stability requires accountability in ensuring accurate and reliable analysis.

PO13: Community Engagement and Service

CO7: 1 - Assessing and improving process capability can contribute to broader community or societal well-being by enhancing product quality and safety.

CBCS Syllabus as per NEP 2020 for S.Y.B.Sc. Statistics (2023 Pattern)

Name of the Programme : B.Sc. Statistics
Programme Code : USST
Class : S.Y.B.Sc.
Semester : IV
Course Type : Major Mandatory (Practical)
Course Code : STA-254-MJM
Course Title : Statistics Practical – IV
No. of Credits : 02
No. of Teaching Hours : 60

Course Objectives:

1. Apply R Software to perform hypothesis tests for means based on the normal distribution.
2. Use R Software to test proportions using the normal distribution.
3. Conduct t-tests using R Software for various statistical analyses.
4. Execute F-tests using R Software for comparing variances.
5. Implement chi-square tests using R Software for independence of attributes and goodness of fit.
6. Fit multiple regression models and compute correlation coefficients using R Software.
7. Compute and interpret probabilities of various statistical distributions using R Software.

Course Outcome:

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

CO1. Conduct hypothesis tests for means, proportions, and distributions using R Software.

CO2. Fit multiple regression models and compute correlation coefficients using R Software.

CO3. Compute and interpret probabilities for various statistical distributions using R Software.

CO4. Model and apply Exponential distributions in statistical analyses.

CO5. Construct and analyze variable and attribute control charts for process control.

CO6. Perform capability studies and apply Statistical Process Control techniques.

CO7. Apply statistical concepts through case study-based hypothesis testing and SPC.

Topics and Learning Points

Sr. No.	Title of Experiments
1	Test for means based on normal distribution (Also using R Software)
2	Test for proportions based on normal distribution (Also using R Software)
3	Test based on t distributions (Also using R Software)
4	Test based on F distributions (Also using R Software)
5	Tests based on chi-square distribution (Independence of attributes and Goodness of fit test) (Also using R Software)
6	Fitting of multiple regression plane and computation of multiple and partial Correlation coefficients using R Software.
7	Computations of probabilities of continuous probability distributions using R Software.
8	Fitting of Exponential Distribution (Also using R Software)
9	Model sampling from continuous probability distributions using R Software
10	Application of Exponential and Gamma distribution
11	Construction of $\bar{X}$ Chart and R Chart
13	Construction of p-chart
14	Construction of c-chart
15	Capability Studies

COs and POs Mapping:

Course Outcomes	**Programme Outcomes (POs)**												
	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**	**PO8**	**PO9**	**PO10**	**PO11**	**PO12**	**PO13**
CO1	3	3	1	2	3	1	2	2	3	1	1	2	1
CO2	3	3	1	3	3	1	2	2	3	1	1	2	1
CO3	3	3	1	2	3	1	2	2	3	1	1	2	1
CO4	3	2	1	2	2	1	1	2	2	1	1	2	1
CO5	2	3	1	3	2	2	2	1	2	1	1	2	1
CO6	2	3	1	3	2	2	2	1	2	1	1	2	1
CO7	2	2	1	2	3	2	2	2	2	1	1	2	1

PO1: Comprehensive Knowledge and Understanding

CO1: Conduct hypothesis tests for means, proportions, and distributions using R Software.

Weightage: 3

Justification: This outcome requires a deep understanding of statistical theories and methods, which are central to comprehensive knowledge in the field.

CO2: Fit multiple regression models and compute correlation coefficients using R Software.

Weightage: 3

Justification: Fitting regression models and computing correlation coefficients involve foundational statistical theories and principles.

CO3: Compute and interpret probabilities for various statistical distributions using R Software.

Weightage: 3

Justification: Understanding and interpreting probabilities is crucial for grasping statistical distributions and their applications.

CO4: Model and apply Exponential distributions in statistical analyses.

Weightage: 3

Justification: Applying specific distributions such as the Exponential distribution requires knowledge of their theoretical underpinnings.

CO5: Construct and analyze variable and attribute control charts for process control.

Weightage: 2

Justification: While control charts are based on statistical principles, the focus is more on practical application rather than deep theoretical knowledge.

CO6: Perform capability studies and apply Statistical Process Control (SPC) techniques.

Weightage: 2

Justification: Capability studies and SPC techniques rely on applied statistical knowledge rather than foundational theory.

CO7: Apply statistical concepts through case study-based hypothesis testing and SPC.

Weightage: 2

Justification: Application through case studies involves understanding and applying theoretical concepts in practical scenarios.

PO2: Practical, Professional, and Procedural Knowledge

CO1: Conduct hypothesis tests for means, proportions, and distributions using R Software.

Weightage: 3

Justification: Implementing hypothesis tests using R Software requires practical knowledge and adherence to professional procedures.

CO2: Fit multiple regression models and compute correlation coefficients using R Software.

Weightage: 3

Justification: Fitting models and computing coefficients involve applying professional and procedural knowledge in statistical analysis.

CO3: Compute and interpret probabilities for various statistical distributions using R Software.

Weightage: 3

Justification: Computing and interpreting probabilities involve practical knowledge and application of statistical procedures.

CO4: Model and apply Exponential distributions in statistical analyses.

Weightage: 2

Justification: Requires practical knowledge of applying specific statistical models and techniques.

CO5: Construct and analyze variable and attribute control charts for process control.

Weightage: 3

Justification: Constructing and analyzing control charts are practical tasks that require professional knowledge in quality control.

CO6: Perform capability studies and apply Statistical Process Control (SPC) techniques.

Weightage: 3

Justification: Involves applying SPC techniques and capability studies, which are crucial for professional and procedural knowledge in quality management.

CO7: Apply statistical concepts through case study-based hypothesis testing and SPC.

Weightage: 2

Justification: Application of statistical concepts through case studies involves practical and procedural knowledge but may not be as directly related to professional practices as other outcomes.

PO3: Entrepreneurial Mind set and Knowledge

CO1: Conduct hypothesis tests for means, proportions, and distributions using R Software.

Weightage: 1

Justification: While statistical analysis may support business decisions, it is not directly related to entrepreneurial mind set or innovation.

CO2: Fit multiple regression models and compute correlation coefficients using R Software.

Weightage: 1

Justification: Although regression analysis can support business insights, it is not directly related to entrepreneurial mind set.

CO3: Compute and interpret probabilities for various statistical distributions using R Software.

Weightage: 1

Justification: Similar to other outcomes, probability computations support business decisions but are not directly linked to an entrepreneurial mind set.

CO4: Model and apply Exponential distributions in statistical analyses.

Weightage: 1

Justification: Applying specific statistical distributions is more focused on technical skills rather than entrepreneurial knowledge.

CO5: Construct and analyze variable and attribute control charts for process control.

Weightage: 1

Justification: Process control is more about technical proficiency than fostering an entrepreneurial mindset.

CO6: Perform capability studies and apply Statistical Process Control (SPC) techniques.

Weightage: 1

Justification: SPC and capability studies support operational efficiency but are not directly related to entrepreneurial activities.

CO7: Apply statistical concepts through case study-based hypothesis testing and SPC.

Weightage: 1

Justification: Case studies may provide insights but are not directly linked to developing an entrepreneurial mind set.

PO4: Specialized Skills and Competencies

CO1: Conduct hypothesis tests for means, proportions, and distributions using R Software.

Weightage: 2

Justification: Requires specialized skills in hypothesis testing and data analysis.

CO2: Fit multiple regression models and compute correlation coefficients using R Software.

Weightage: 3

Justification: Fitting regression models and computing correlations are specialized skills in statistical analysis.

CO3: Compute and interpret probabilities for various statistical distributions using R Software.

Weightage: 2

Justification: Involves specialized knowledge in probability and statistical distributions.

CO4: Model and apply Exponential distributions in statistical analyses.

Weightage: 2

Justification: Requires specific skills in modelling and applying Exponential distributions.

CO5: Construct and analyze variable and attribute control charts for process control.

Weightage: 3

Justification: Specialized skills in constructing and analysing control charts are essential for quality control.

CO6: Perform capability studies and apply Statistical Process Control (SPC) techniques.

Weightage: 3

Justification: Performing capability studies and applying SPC techniques require specialized technical skills.

CO7: Apply statistical concepts through case study-based hypothesis testing and SPC.

Weightage: 2

Justification: Applying statistical concepts through case studies involves specialized knowledge but may not be as focused as other outcomes.

PO5: Capacity for Application, Problem-Solving, and Analytical Reasoning

CO1: Conduct hypothesis tests for means, proportions, and distributions using R Software.

Weightage: 3

Justification: Conducting hypothesis tests requires applying knowledge to solve statistical problems and reasoning analytically.

CO2: Fit multiple regression models and compute correlation coefficients using R Software.

Weightage: 3

Justification: Fitting regression models and computing correlations involve complex problem-solving and analytical reasoning.

CO3: Compute and interpret probabilities for various statistical distributions using R Software.

Weightage: 3

Justification: Computing and interpreting probabilities requires analytical skills and problem-solving abilities.

CO4: Model and apply Exponential distributions in statistical analyses.

Weightage: 2

Justification: Applying Exponential distributions involves problem-solving and analytical reasoning, though less complex compared to other outcomes.

CO5: Construct and analyze variable and attribute control charts for process control.

Weightage: 2

Justification: Constructing and analyzing control charts involves problem-solving within a specific context.

CO6: Perform capability studies and apply Statistical Process Control (SPC) techniques.

Weightage: 2

Justification: Capability studies and SPC techniques require problem-solving and analytical reasoning in quality management.

CO7: Apply statistical concepts through case study-based hypothesis testing and SPC.

Weightage: 3

Justification: Applying statistical concepts to case studies involves extensive problem-solving and analytical reasoning.

PO6: Communication Skills and Collaboration

CO1: Conduct hypothesis tests for means, proportions, and distributions using R Software.

Weightage: 1

Justification: The primary focus is on technical skills rather than communication or collaboration.

CO2: Fit multiple regression models and compute correlation coefficients using R Software.

Weightage: 1

Justification: Emphasis is on technical execution rather than communication.

CO3: Compute and interpret probabilities for various statistical distributions using R Software.

Weightage: 1

Justification: Focus is on technical calculations rather than communication.

CO4: Model and apply Exponential distributions in statistical analyses.

Weightage: 1

Justification: Technical application is the main focus, with limited emphasis on communication skills.

CO5: Construct and analyze variable and attribute control charts for process control.

Weightage: 2

Justification: Communicating control chart results effectively is important for quality control.

CO6: Perform capability studies and apply Statistical Process Control (SPC) techniques.

Weightage: 2

Justification: Capability studies and SPC require effective communication of findings within a professional context.

CO7: Apply statistical concepts through case study-based hypothesis testing and SPC.

Weightage: 2

Justification: Case studies often require effective communication and collaboration to present findings and analyses.

PO7: Research-related Skills

CO1: Conduct hypothesis tests for means, proportions, and distributions using R Software.

Weightage: 2

Justification: Hypothesis testing is relevant to research but not the sole focus.

CO2: Fit multiple regression models and compute correlation coefficients using R Software.

Weightage: 2

Justification: Regression models are used in research but the focus is on technical skills.

CO3: Compute and interpret probabilities for various statistical distributions using R Software.

Weightage: 2

Justification: Probabilities are important for research but are part of a broader skill set.

CO4: Model and apply Exponential distributions in statistical analyses.

Weightage: 2

Justification: Modelling distributions is relevant to research, though not central to all research activities.

CO5: Construct and analyze variable and attribute control charts for process control.

Weightage: 1

Justification: Focus is more on process control than on research-related skills.

CO6: Perform capability studies and apply Statistical Process Control (SPC) techniques.

Weightage: 1

Justification: More relevant to quality management than to research.

CO7: Apply statistical concepts through case study-based hypothesis testing and SPC.

Weightage: 3

Justification: Applying statistical concepts in case studies is closely related to research skills.

PO8: Learning How to Learn Skills

CO1: Conduct hypothesis tests for means, proportions, and distributions using R Software.

Weightage: 2

Justification: Learning to conduct various tests supports the development of self-directed learning skills.

CO2: Fit multiple regression models and compute correlation coefficients using R Software.

Weightage: 2

Justification: Learning to fit models and compute correlations enhances independent learning and problem-solving.

CO3: Compute and interpret probabilities for various statistical distributions using R Software.

Weightage: 2

Justification: Computing and interpreting probabilities contributes to learning and adapting to statistical methods.

CO4: Model and apply Exponential distributions in statistical analyses.

Weightage: 2

Justification: Applying specific statistical models helps in acquiring learning skills.

CO5: Construct and analyze variable and attribute control charts for process control.

Weightage: 1

Justification: Primarily focused on technical skills rather than learning how to learn.

CO6: Perform capability studies and apply Statistical Process Control (SPC) techniques.

Weightage: 1

Justification: Capability studies and SPC techniques are less related to learning how to learn.

CO7: Apply statistical concepts through case study-based hypothesis testing and SPC.

Weightage: 2

Justification: Applying concepts through case studies supports independent learning and adaptability.

PO9: Digital and Technological Skills

CO1: Conduct hypothesis tests for means, proportions, and distributions using R Software.

Weightage: 3

Justification: Utilizing R Software for hypothesis testing demonstrates proficiency in digital and technological tools.

CO2: Fit multiple regression models and compute correlation coefficients using R Software.

Weightage: 3

Justification: Requires advanced digital and technological skills for model fitting and analysis.

CO3: Compute and interpret probabilities for various statistical distributions using R Software.

Weightage: 3

Justification: Involves using technology to perform and interpret probability calculations.

CO4: Model and apply Exponential distributions in statistical analyses.

Weightage: 2

Justification: Involves applying digital tools for modelling but with a narrower focus.

CO5: Construct and analyze variable and attribute control charts for process control.

Weightage: 2

Justification: Constructing control charts involves technological skills in data analysis.

CO6: Perform capability studies and apply Statistical Process Control (SPC) techniques.

Weightage: 2

Justification: Requires using technological tools for analysis and process control.

CO7: Apply statistical concepts through case study-based hypothesis testing and SPC.

Weightage: 2

Justification: Applying statistical concepts through case studies involves technological skills in data analysis.

PO10: Multicultural Competence, Inclusive Spirit, and Empathy

CO1: Conduct hypothesis tests for means, proportions, and distributions using R Software.

Weightage: 1

Justification: The focus is on technical skills rather than multicultural competence.

CO2: Fit multiple regression models and compute correlation coefficients using R Software.

Weightage: 1

Justification: Primarily technical with limited focus on multicultural competence.

CO3: Compute and interpret probabilities for various statistical distributions using R Software.

Weightage: 1

Justification: Involves technical work rather than multicultural or inclusive skills.

CO4: Model and apply Exponential distributions in statistical analyses.

Weightage: 1

Justification: Focus is on statistical modelling rather than multicultural competence.

CO5: Construct and analyze variable and attribute control charts for process control.

Weightage: 1

Justification: The main focus is on technical skills in process control.

CO6: Perform capability studies and apply Statistical Process Control (SPC) techniques.

Weightage: 1

Justification: Emphasis on process control rather than multicultural or inclusive skills.

CO7: Apply statistical concepts through case study-based hypothesis testing and SPC.

Weightage: 1

Justification: Case studies focus more on technical application than multicultural competence.

PO11: Value Inculcation and Environmental Awareness

CO1: Conduct hypothesis tests for means, proportions, and distributions using R Software.

Weightage: 1

Justification: Technical focus with limited direct relevance to values and environmental awareness.

CO2: Fit multiple regression models and compute correlation coefficients using R Software.

Weightage: 1

Justification: Primarily technical work with little direct link to value inculcation.

CO3: Compute and interpret probabilities for various statistical distributions using R Software.

Weightage: 1

Justification: Focus on technical skills rather than values or environmental issues.

CO4: Model and apply Exponential distributions in statistical analyses.

Weightage: 1

Justification: Technical application with limited focus on values or environmental awareness.

CO5: Construct and analyze variable and attribute control charts for process control.

Weightage: 1

Justification: Primarily technical with limited connection to value inculcation.

CO6: Perform capability studies and apply Statistical Process Control (SPC) techniques.

Weightage: 1

Justification: Focuses on technical aspects of process control rather than values or environmental issues.

CO7: Apply statistical concepts through case study-based hypothesis testing and SPC.

Weightage: 1

Justification: Emphasis on technical application through case studies, with limited relevance to values and environmental awareness.

PO12: Autonomy, Responsibility, and Accountability

CO1: Conduct hypothesis tests for means, proportions, and distributions using R Software.

Weightage: 2

Justification: Conducting hypothesis tests requires a degree of responsibility and accountability for accurate results.

CO2: Fit multiple regression models and compute correlation coefficients using R Software.

Weightage: 2

Justification: Fitting models and computing correlations involve responsibility for ensuring correctness and accuracy.

CO3: Compute and interpret probabilities for various statistical distributions using R Software.

Weightage: 2

Justification: Requires accountability for accurate probability calculations and interpretations.

CO4: Model and apply Exponential distributions in statistical analyses.

Weightage: 2

Justification: Applying specific statistical models involves responsibility and accountability for accurate application.

CO5: Construct and analyze variable and attribute control charts for process control.

Weightage: 2

Justification: Constructing and analyzing control charts require accountability in process control and data analysis.

CO6: Perform capability studies and apply Statistical Process Control (SPC) techniques.

Weightage: 2

Justification: Capability studies and SPC techniques involve responsibility for accurate analysis and reporting.

CO7: Apply statistical concepts through case study-based hypothesis testing and SPC.

Weightage: 2

Justification: Applying statistical concepts through case studies involves autonomy in problem-solving and responsibility for results.

PO13: Community Engagement and Service

CO1: Conduct hypothesis tests for means, proportions, and distributions using R Software.

Weightage: 1

Justification: Technical focus with limited direct relevance to community engagement or service.

CO2: Fit multiple regression models and compute correlation coefficients using R Software.

Weightage: 1

Justification: Focus on technical skills rather than community engagement.

CO3: Compute and interpret probabilities for various statistical distributions using R Software.

Weightage: 1

Justification: Emphasis on technical skills with limited community service relevance.

CO4: Model and apply Exponential distributions in statistical analyses.

Weightage: 1

Justification: Technical focus with minimal connection to community engagement or service.

CO5: Construct and analyze variable and attribute control charts for process control.

Weightage: 1

Justification: Primarily technical with limited direct relevance to community service.

CO6: Perform capability studies and apply Statistical Process Control (SPC) techniques.

Weightage: 1

Justification: Focus on technical and process control aspects rather than community service.

CO7: Apply statistical concepts through case study-based hypothesis testing and SPC.

Weightage: 1

Justification: Applying concepts through case studies is more technical and less related to community engagement.

CBCS Syllabus as per NEP 2020 for S.Y.B.Sc. Statistics (2023 Pattern)

Name of the Programme : B.Sc. Statistics

Programme Code : USST

Class : S.Y.B.Sc.(Minor)

Semester : IV

Course Type : Major Mandatory (Theory)

Course Code : STA-261-MN

Course Title : Probability Distributions and Applications

No. of Credits : 02

No. of Teaching Hours : 30

Course Objectives:

1. To compute various measures of skewness and kurtosis.
2. Gain knowledge of discrete bivariate random variable and its probability distributions.
3. To obtain a probability distribution of bivariate random variable in the given situation.
4. Analyze the data with respect to bivariate discrete distributions.
5. Computation of correlation coefficients, Regression coefficients and their interpretation.
6. Students will develop a strong foundation to various discrete probability distributions.
7. Identify real life situations where Negative Binomial, Poisson and Geometric distribution can be applied, and compute the probabilities of the given event.

Course Outcomes:

Student will be able to:

CO1. describe bivariate random variable, joint distribution function, joint probability mass function, marginal and conditional distributions, mathematical expectation, and independence of random variables, and apply their properties in problem-solving.

CO2. apply binomial distribution in different real-life situations.

CO3. apply Hyper-geometric distribution in different real-life situations.

CO4. apply Poisson distribution in different real-life situations.

CO5. apply Geometric distribution in different real-life situations.

CO6. study the properties of the distributions.

CO7. study the interrelation between the standard probability distributions learnt above.

Topics and Learning Points

UNIT – 1 Moments, Skewness and Kurtosis (5L)

1.1. Raw moments $\mu_r' ; r = 1, 2, 3, 4$ for ungrouped data.

1.2. Central moments $\mu_r ; r = 1, 2, 3, 4$ for ungrouped data, effect of change of origin and scale on moments.

1.3. Relations between central moments and raw moments, up to 4^{th} order.

1.4. Skewness: Definition, types of skewness, measures of skewness;

1.4.1. Karl Pearson coefficient of skewness

1.4.2. Pearsonian coefficient of skewness

1.4.3. Bowley's coefficient of skewness

1.5. Box plot.

1.6. Kurtosis: Definition, types of kurtosis, measure of kurtosis based on moments and partition values.

1.7. Examples and problems.

Unit – 2 Bivariate discrete probability distribution (6L)

2.1 Definition of a bivariate discrete random variable (X, Y) on finite sample space, Joint p.m.f., and c.d.f., Properties of c.d.f. (without proof).

2.2 Computation of probabilities of events in the bivariate probability distribution

2.3 Marginal and conditional probability distribution, independence of two discrete r.v.s. Examples.

2.4 Mathematical Expectation: Definition of expectation of function of r.v. in bivariate distribution, Theorems on expectations: (i) E(X+Y) = E(X) + E(Y) (ii) E(XY) = E(X)·E(Y) when X and Y are independent, expectation and variance of linear combination of two discrete r.v.s.

2.5 Conditional mean, conditional variance, covariance and correlation coefficient, distinction between uncorrelated and independent variables,

2.6 Define m.g.f., properties of m.g.f., examples.

Unit-3: Correlation: (5L)

3.1. Covariance: Definition, computation, the effect of change of origin, and scale.

3.2. Correlation: Types of correlation, scatter diagram, Interpretation of correlation using Scatter diagram.

3.3. Karl Pearson's coefficient of correlation (r): Definition, computation for ungrouped data,

and interpretation. Properties: (i) $-1 \leq r \leq 1$ (ii) Effect of change of origin and scale (with proof).

3.4. Spearman's rank correlation coefficient: Definition, derivation of formula, computation, and interpretation (without ties). In case of ties, compute Karl Pearson's correlation coefficient between ranks. (Spearman's rank correlation coefficient formula with correction for ties not expected.)

Unit 4: Regression Analysis (7L)

4.1 Concept of regression, lines of regression, fitting of lines of regression by the least square method, interpretation of slope and intercept.

4.2 Regression coefficient (b_{yx} , b_{xy}): Definition, computation, properties (without proof). (i) $b_{yx}, b_{xy} = r^2$, (ii) $b_{yx}, b_{xy} \leq 1$, (iii) $b_{yx} = r\frac{\sigma_y}{\sigma_x}$, $b_{xy} = r\frac{\sigma_x}{\sigma_y}$ (iv) Effect of change of origin and scale, (v) Angle between the two lines of regression.

4.3 Mean Residual sum of squares (S.S) = $\frac{1}{n-2}\sum(y_i - \hat{y}_i)^2$, Residual plot and its interpretation.

4.4 Explained and unexplained variation, coefficient of determination.

4.5 Non-linear regression: (i) Second degree curve, (ii) Exponential curve of type $y = ab^x$, fitting of such curves by the least square method after logarithmic transformation, (iii) Logistic curve $y = \frac{k}{1+e^{a+bx}}$. Interpretation of b<0, b>0.Illustration of logistic curve. (Fitting of logistic curve is not expected).

4.6 Examples and problems.

Unit 5: Discrete Probability distribution (countably infinite sample space) (7L)

5.1 Poisson Distribution:

Review of random variable based on countably infinite sample space. Definition of Poisson with parameter λ. Mean, variance**,** mode, m.g.f., c.g.f. skewness, kurtosis, Recurrence relation for successive Probabilities, Additive property of Poisson distribution. Poisson distribution as a limiting case of Binomial distribution, examples. Conditional distribution of X given (X+Y) for Poisson distributions. Real life situations.

5.2 Geometric Distribution:

Definition of Geometric with parameter p in both cases with support {0, 1, 2,….} and with support{1, 2, …..}.Mean, Variance, distribution function, Lack of memory property, examples. Real life situations.

5.3 Negative Binomial Distribution

Probability mass function (p.m.f.).Nature of probability curve, negative binomial distribution as a waiting time distribution, Moment Generating Function (MGF), Cumulant Generating Function (CGF), mean, variance, skewness, kurtosis (recurrence relation between moments is not expected), additive property of NB (k, p).Relation between Geometric Distribution and Negative Binomial Distribution. Poisson approximation to Negative Binomial Distribution. Real life situations.

References:

1. Goon A. M., Gupta M. K., Das Gupta B. (1999): Fundamentals of Statistics, Vol.II, World Press, Calcutta.
2. Gupta and Kapoor: Fundamentals of Mathematical Statistics, Sultan Chand and Sons, New Delhi.
3. Sharma K. V. S. (2001) Statistics made it simple: Do it yourself on PC. Prentce Hall of India, New Delhi.
4. Gupta and Kapoor : Fundamentals of Applied Statistics, Sultan Chand and Sons, New Delhi.
5. B. L. Agarwal : Programmed Statistics, New Age International Publishers, New Delhi.
6. David Freedman, Robert Pisani, Roger Purves: Statistics
7. Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye: Probability & Statistics for Engineers & Scientists.
8. Brase C. H. and Brase C. P. (2018), Understandable Statistics, Twelfth Edition, Cengage Learning, Biston
9. Moor D. S., Notz W. I., Flinger M. A., (2013), The Basic Practice of Statistics Sixth Edition, Freeman and Company New York

Programme Outcomes and Course Outcomes Mapping:

CO-PO Mapping Table

COs \ POs	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PO13
CO1	3	2	2	3	3	2	3	2	2	1	1	2	1
CO2	2	3	2	3	3	2	2	2	2	1	1	2	1
CO3	2	3	2	3	3	2	2	2	2	1	1	2	1
CO4	2	3	2	3	3	2	2	2	2	1	1	2	1
CO5	2	3	2	3	3	2	2	2	2	1	1	2	1
CO6	3	2	2	3	3	2	3	2	2	1	1	2	1

CO7	3	2	2	3	3	2	3	2	2	1	1	2	1

This mapping table and the justifications show how each Course Outcome (CO) aligns with the Program Outcomes (POs) and the extent of their relationship.

Justification for Mapping PO and CO

PO1: Comprehensive Knowledge and Understanding

CO1 (3): Strongly related, as describing bivariate random variables and joint distribution functions requires a profound understanding of foundational concepts in probability and statistics.

CO2 (2): Moderately related because applying the binomial distribution involves using fundamental probability principles.

CO3 (2): Moderately related as understanding the hyper-geometric distribution builds on foundational probability theory.

CO4 (2): Moderately related because applying the Poisson distribution also relies on key statistical concepts.

CO5 (2): Moderately related since using the geometric distribution requires knowledge of probability basics.

CO6 (3): Strongly related, as studying the properties of distributions demands a comprehensive understanding of statistical principles.

CO7 (3): Strongly related due to the need to understand the interrelation between standard probability distributions, which involves a deep knowledge of the field.

PO2: Practical, Professional, and Procedural Knowledge

CO1 (2): Moderately related, as applying joint distributions and independence concepts in real-world problems requires practical skills.

CO2 (3): Strongly related, as applying the binomial distribution in real-life situations requires practical statistical knowledge.

CO3 (3): Strongly related due to the direct application of the hyper-geometric distribution in professional contexts.

CO4 (3): Strongly related, as the Poisson distribution is widely used in various practical and professional scenarios.

CO5 (3): Strongly related because using geometric distribution involves procedural knowledge applied in practical situations.

CO6 (2): Moderately related, as studying distribution properties aids in making informed decisions in professional settings.

CO7 (2): Moderately related because understanding the relationships between distributions

supports practical application in real-world problems.

PO3: Entrepreneurial Mindset and Knowledge

CO1 (2): Moderately related, as the ability to analyze data can help identify business opportunities.

CO2 (2): Moderately related because the application of binomial distribution can aid in decision-making and risk assessment in entrepreneurial contexts.

CO3 (2): Moderately related, as hyper-geometric distribution can be applied to market research and innovation strategies.

CO4 (2): Moderately related since Poisson distribution is used to model events that could affect business operations.

CO5 (2): Moderately related because geometric distribution can be applied in entrepreneurial risk management.

CO6 (2): Moderately related, as understanding distribution properties can aid in innovation and market analysis.

CO7 (2): Moderately related due to the application of standard probability distributions in business strategy formulation.

PO4: Specialized Skills and Competencies

CO1 (3): Strongly related, as applying joint distributions and independence requires specialized analytical skills.

CO2 (3): Strongly related because using binomial distribution involves problem-solving and technical expertise.

CO3 (3): Strongly related, as applying the hyper-geometric distribution demands proficiency in analytical techniques.

CO4 (3): Strongly related since applying Poisson distribution requires specialized knowledge.

CO5 (3): Strongly related because geometric distribution usage involves technical skills and problem-solving abilities.

CO6 (3): Strongly related, as studying distribution properties requires advanced analytical skills.

CO7 (3): Strongly related, as understanding the interrelation between distributions demonstrates high-level competence in the field.

PO5: Capacity for Application, Problem-Solving, and Analytical Reasoning

CO1 (3): Strongly related, as applying concepts like joint distributions and expectations involves problem-solving and analytical reasoning.

CO2 (3): Strongly related because applying binomial distribution requires critical thinking and problem-solving.

CO3 (3): Strongly related due to the need for analytical reasoning when applying hyper-geometric distribution.

CO4 (3): Strongly related, as applying Poisson distribution involves solving real-world problems.

CO5 (3): Strongly related because using geometric distribution involves significant problem-solving skills.

CO6 (3): Strongly related, as studying distribution properties enhances analytical reasoning and problem-solving capacity.

CO7 (3): Strongly related due to the analytical skills required to understand the relationships between distributions.

PO6: Communication Skills and Collaboration

CO1 (2): Moderately related, as explaining complex probability concepts involves communication skills.

CO2 (2): Moderately related because effectively applying binomial distribution often requires collaboration and clear communication.

CO3 (2): Moderately related, as hyper-geometric distribution problems may involve team collaboration.

CO4 (2): Moderately related since Poisson distribution applications often require communication of findings.

CO5 (2): Moderately related because communicating geometric distribution results is important in collaborative environments.

CO6 (2): Moderately related, as discussing distribution properties involves effective communication.

CO7 (2): Moderately related, as understanding and explaining interrelations between distributions requires collaboration and communication.

PO7: Research-related Skills

CO1 (3): Strongly related, as analyzing bivariate random variables and distributions is fundamental to research in probability and statistics.

CO2 (2): Moderately related because applying binomial distribution in research requires appropriate methodologies.

CO3 (2): Moderately related, as hyper-geometric distribution is often used in research settings.

CO4 (2): Moderately related since Poisson distribution is commonly applied in research to model real-world phenomena.

CO5 (2): Moderately related because geometric distribution can be used in research for analyzing certain types of data.

CO6 (3): Strongly related, as studying distribution properties is essential for conducting thorough research.

CO7 (3): Strongly related because understanding the relationships between distributions is crucial in research contexts.

PO8: Learning How to Learn Skills

CO1 (2): Moderately related, as understanding complex statistical concepts fosters independent learning skills.

CO2 (2): Moderately related because applying binomial distribution encourages continuous learning through practical application.

CO3 (2): Moderately related, as using hyper-geometric distribution requires adapting to new problems.

CO4 (2): Moderately related since applying Poisson distribution promotes self-directed learning in practical contexts.

CO5 (2): Moderately related because using geometric distribution involves learning new techniques for problem-solving.

CO6 (2): Moderately related, as studying distribution properties fosters the ability to learn and adapt.

CO7 (2): Moderately related, as understanding the interrelation between distributions promotes lifelong learning in statistics.

PO9: Digital and Technological Skills

CO1 (2): Moderately related, as analyzing distributions often involves using statistical software and digital tools.

CO2 (2): Moderately related because applying binomial distribution can require technological skills in data analysis.

CO3 (2): Moderately related, as hyper-geometric distribution is often analyzed using digital tools.

CO4 (2): Moderately related since Poisson distribution applications frequently involve technology.

CO5 (2): Moderately related because using geometric distribution in problem-solving may require digital tools.

CO6 (2): Moderately related, as studying distribution properties often involves technological tools.

CO7 (2): Moderately related, as understanding relationships between distributions is enhanced by digital analysis tools.

PO10: Multicultural Competence, Inclusive Spirit, and Empathy

CO1 (1): Partially related, as probability concepts can be applied in diverse cultural contexts to solve real-world problems.

CO2 (1): Partially related because binomial distribution may be used in studies involving multicultural populations.

CO3 (1): Partially related, as hyper-geometric distribution can apply to diverse settings, but its direct relation is limited.

CO4 (1): Partially related, since Poisson distribution applications might address multicultural issues.

CO5 (1): Partially related because geometric distribution has limited direct application to multicultural competence.

CO6 (1): Partially related, as distribution properties may apply to problems across diverse cultures.

CO7 (1): Partially related due to the limited direct application of distribution interrelations to multicultural competence.

PO11: Value Inculcation and Environmental Awareness

CO1 (1): Partially related, as understanding statistical principles can support responsible decision-making related to values and the environment.

CO2 (1): Partially related because applying binomial distribution could contribute to data-driven ethical decisions.

CO3 (1): Partially related, as hyper-geometric distribution can be used in studies promoting environmental awareness.

CO4 (1): Partially related since Poisson distribution might be used to model environmental phenomena.

CO5 (1): Partially related because geometric distribution can support decisions aligned with environmental conservation.

CO6 (1): Partially related, as studying distribution properties could support responsible citizenship.

CO7 (1): Partially related due to limited direct relevance to values and environmental issues.

PO12: Autonomy, Responsibility, and Accountability

CO1 (2): Moderately related, as applying statistical concepts independently fosters autonomy in problem-solving.

CO2 (2): Moderately related because applying binomial distribution requires responsibility in data analysis.

CO3 (2): Moderately related, as hyper-geometric distribution requires accountability in applying statistical methods.

CO4 (2): Moderately related since Poisson distribution fosters responsibility in analyzing real-world events.

CO5 (2): Moderately related because using geometric distribution encourages accountability in problem-solving.

CO6 (2): Moderately related, as studying distribution properties promotes independent learning and responsibility.

CO7 (2): Moderately related, as understanding relationships between distributions fosters accountability in research and practice.

PO13: Community Engagement and Service

CO1 (1): Partially related, as statistical knowledge can be applied to community-based research and service.

CO2 (1): Partially related because binomial distribution might be used in community-focused data analysis.

CO3 (1): Partially related, as hyper-geometric distribution can be applied to community engagement studies.

CO4 (1): Partially related since Poisson distribution might be used to model community events.

CO5 (1): Partially related because geometric distribution has limited direct application to community service.

CO6 (1): Partially related, as studying distribution properties may support community-based initiatives.

CO7 (1): Partially related due to limited direct application of distribution interrelations to community engagement.

CBCS Syllabus as per NEP 2020 for S.Y.B.Sc. Statistics (2023 Pattern)

Name of the Programme : B.Sc. Statistics
Programme Code : USST
Class : S.Y.B.Sc.
Semester : IV
Course Type : Minor Practical
Course Code : STA-262-MN
Course Title : Minor Statistics Practical – II
No. of Credits : 02
No. of Teaching Hours : 60

Course Objectives:

1. To provide a foundational understanding of raw and central moments for ungrouped data and their applications in statistical analysis
2. To introduce the concepts of skewness and kurtosis, their types, and various measures, and to apply these concepts in analyzing the symmetry and peakedness of data distributions.
3. To develop skills in handling bivariate discrete data, including understanding joint probability distributions, marginal, conditional distributions and independence.
4. To compute and interpret mathematical expectations, covariance, and correlation coefficients, and their application in the analysis of relationships between variables.
5. To equip students with the ability to fit regression, understand the significance of regression coefficients, and interpret the results in terms of slope and intercept.
6. To introduce non-linear regression techniques, including second-degree, exponential, and logistic curves, and to understand the implications of these models.
7. To develop problem-solving skills through practical examples and exercises, reinforcing theoretical knowledge and its application to real-world data.

Course Outcomes:

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

CO1. calculate raw and central moments for ungrouped data and understand the relationships between them, enabling deeper insight into data distribution characteristics.

CO2. compute and interpret various measures of skewness and kurtosis, and use these metrics to assess the symmetry and shape of data distributions.

CO3. handle bivariate discrete probability distributions, compute marginal and conditional probabilities, and assess the independence of random variables.

CO4. calculate and interpret covariance and correlation coefficients, including Karl Pearson's and Spearman's rank correlations, and understanding their significance in data analysis.

CO5. Fit and interpret lines of regression, understanding the relationship between variables, and analysing residuals to assess model accuracy.

CO6. apply non-linear regression techniques, fit appropriate curves to data, and interpret the implications of different models, such as exponential and logistic growth.

CO7. apply Poisson, Geometric, and Negative Binomial distribution to real-world situations.

Topics and Learning Points

Sr. No.	Title of Experiments
1	Moments, Skewness and Kurtosis-I
2	Moments, Skewness and Kurtosis-II
3	Bivariate discrete probability distribution-I
4	Bivariate discrete probability distribution-II
5	Simple Linear Regression
6	Non- Linear Regression
7	Introduction of R software – I
8	Introduction of R software – II
9	Introduction of R software – III
10	Fitting of Simple Linear Regression, second degree curve, and Exponential curve of type y = ab^x Using MS-Excel
11	Applications of Poisson Distribution
12	Applications of Geometric and Negative Binomial Distribution
13	Fitting of Poisson Distribution
14	Fitting of Negative Binomial Distribution
15	Model Sampling from Poisson and Negative Binomial Distribution

COs and POs Mapping:

Course Outcomes	Programme Outcomes (POs)

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PO13
CO1	3	2	1	2	3	1	2	2	1	1	1	2	1
CO2	3	2	1	2	3	1	2	2	1	1	1	2	1
CO3	3	3	1	3	3	2	2	2	2	1	1	3	1
CO4	3	3	1	3	3	2	3	2	2	1	1	3	1
CO5	3	3	1	3	3	2	3	2	2	1	1	3	1
CO6	3	3	1	3	3	2	3	2	2	1	1	3	1
CO7	3	3	1	3	3	2	3	2	2	1	1	3	1

Justifications for CO-PO Mappings

PO1: Comprehensive Knowledge and Understanding

CO1, CO2, CO3, CO4, CO5, CO6, CO7 (3): All these COs involve fundamental and advanced statistical concepts, providing students with a deep understanding of data analysis, probability distributions, regression techniques, and their applications in real-world scenarios.

PO2: Practical, Professional, and Procedural Knowledge

CO3, CO4, CO5, CO6, CO7 (3): These COs emphasize the practical application of statistical methods, such as correlation, regression, and probability distributions, which are essential for professional tasks in data analysis, economics, and business.

CO1, CO2 (2): These COs also contribute by laying the foundation for understanding data characteristics and distribution, which is crucial for professional data analysis.

PO3: Entrepreneurial Mindset and Knowledge

CO1 to CO7 (1): While these COs are more technically focused, they contribute to developing an entrepreneurial mindset by providing the analytical tools needed to assess market trends, forecast outcomes, and manage risks.

PO4: Specialized Skills and Competencies

CO3, CO4, CO5, CO6, CO7 (3): These COs focus on developing specialized skills in handling and interpreting complex data, which are critical for technical roles in statistics, economics, and business.

CO1, CO2 (2): These COs develop analytical abilities in understanding data distribution and moments, which are key for specialized analysis.

PO5: Capacity for Application, Problem-Solving, and Analytical Reasoning

CO1 to CO7 (3): All these COs strongly emphasize applying statistical methods to solve real-world problems, analyze data effectively, and interpret results, thus fostering strong problem-solving and analytical reasoning abilities.

PO6: Communication Skills and Collaboration

CO3, CO4, CO5, CO6, CO7 (2): These COs involve interpreting statistical results, which requires effective communication of findings. Collaborative exercises in applying these methods also enhance teamwork skills.

CO1, CO2 (1): These COs involve the communication of basic statistical concepts, which is foundational for effective collaboration in more complex tasks.

PO7: Research-related Skills

CO4, CO5, CO6, CO7 (3): These COs directly involve skills necessary for research, such as analyzing relationships between variables, fitting models, and applying probability distributions.

CO1, CO2, CO3 (2): These COs provide the foundational knowledge necessary for conducting research, such as understanding data distribution and calculating probabilities.

PO8: Learning How to Learn Skills

CO1 to CO7 (2): All these COs contribute to developing self-learning skills, as students must independently apply concepts, solve problems, and adapt their understanding to different contexts.

PO9: Digital and Technological Skills

CO3 to CO7 (2): These COs involve the use of software tools for statistical analysis, enhancing students' digital and technological proficiency.

CO1, CO2 (1): Basic calculations and understanding of statistical measures also require the use of digital tools, contributing to technological skills.

PO10: Multicultural Competence, Inclusive Spirit, and Empathy

CO1 to CO7 (1): Although primarily technical, these COs contribute to understanding diverse data contexts, which can indirectly foster a broader appreciation of multicultural perspectives.

PO11: Value Inculcation and Environmental Awareness

CO1 to CO7 (1): Statistical analysis can contribute to responsible decision-making, including in areas related to environmental sustainability, though this is not the primary focus of these COs.

PO12: Autonomy, Responsibility, and Accountability

CO3 to CO7 (3): These COs encourage students to independently apply statistical techniques, take responsibility for their analyses, and ensure accuracy in their work.

CO1, CO2 (2): Basic statistical calculations also require a degree of autonomy and responsibility, particularly in ensuring correct interpretations.

PO13: Community Engagement and Service

CO1 to CO7 (1): While these COs are technical, the analytical skills developed can be applied to community-focused projects and services, contributing to societal well-being

CBCS Syllabus as per NEP 2020 for S.Y.B. Sc. Statistics (2023 Pattern)

Name of the Programme : B.Sc.Statistics

Program Code : USST

Class : S.Y.B.Sc.

Semester : IV

Course Type : OE (Practical)

Course Code : STA-266-OE

Course Title : Practical based on Applied Statistical Techniques

No. of Credits : 2 credits

No. of Teaching Hours : 60

Course Objectives:

1. To provide students with the fundamental skills and knowledge necessary to effectively utilize MS-Excel for data analysis and management.
2. To formulate and interpret the equation of the regression plane involving multiple predictor variables.
3. To understand the concept of index numbers and their applications in economics.
4. To analyze time series data to identify trends, seasonal variations, and irregularities.
5. To understand the Fundamentals of Linear Programming (LP)
6. To formulate transportation models to minimize cost, time, or other resources in transportation and distribution.
7. To understand the assignment problem and its application in resource allocation, scheduling, and workforce assignment.
8. To understand the basics of decision theory, including decision-making under uncertainty and risk.

Course Outcomes:

Student will be able to

CO1. Understand the challenges and problems associated with the construction of index numbers.

CO2. Formulate and interpret regression equations involving multiple predictor variables, enhancing their capability to model and analyze relationships in data.

CO3. Develop an understanding of index numbers, their construction, and applications in economics, enabling them to analyze economic data and trends effectively.

CO4. To analyze time series data, identify patterns such as trends, seasonal variations, and irregularities, and make informed decisions based on their analysis.

CO5. Understand the basic principles of Linear Programming (LP), enabling them to Formulate and solve optimization problems in various contexts.

CO6. Formulate transportation models aimed at minimizing costs, time, or other resources, applying these skills to optimize logistics and distribution systems.

CO7. Understand and apply techniques to solve assignment problems, including resource allocation, scheduling, and workforce assignment, enhancing their decision-making capabilities.

Topics and Learning Points:

Sr. No.	Title of Experiment
1.	Multiple and Partial Correlation
2.	Multiple Linear Regression
3.	Index Number – I
4.	Index Number – II
5.	Time Series – I
6.	Time Series – II
7.	Linear Programming Problem – I
8.	Linear Programming Problem – II
9.	Transportation Problem – I
10.	Transportation Problem – II
11.	Assignment Problem – I
12.	Assignment Problem – II
13.	Demography – I
14.	Demography – II
15.	Life Table

COs and POs Mapping:

Course Outcomes	Programme Outcomes (POs)												
	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12	PO 13
CO1	3	2	1	2	3	1	2	1		1	1	1	1
CO2	2	2	1	3	3	2	3	2	2	1	1	2	1
CO3	3	2	1	2	3	1	2	1		1	1	1	1
CO4	2	2	1	2	3	1	2	2	2	1	1	2	1
CO5	3	2	1	3	3	1	2	2	2	1	1	2	1
CO6	2	3	1	3	3	1	2	2	3	1	1	2	2
CO7	2	3	1	3	3	1	2	2	3	1	1	2	2

Justifications for Mapping

PO1: Comprehensive Knowledge and Understanding

CO1, CO3, and CO5: Strongly related (3) because these COs emphasize understanding theoretical concepts and methodologies, which align with foundational knowledge.

CO2, CO4, CO6, and CO7: Related (2) as they involve applying foundational theories in practical scenarios.

PO2: Practical, Professional, and Procedural Knowledge

CO6, CO7: Strongly related (3) because these COs focus on practical skills relevant to professional and industry standards, like optimizing transportation and assignment problems.

CO1, CO2, CO3, CO4, CO5: Moderately related (2) due to their practical applications in analyzing data, formulating models, and solving real-world problems.

PO3: Entrepreneurial Mind-set and Knowledge

CO1 to CO7: Partially related (1) as they contribute indirectly by equipping students with analytical skills that can foster innovative thinking and problem-solving abilities.

PO4: Specialized Skills and Competencies

CO2, CO5, CO6, and CO7: Strongly related (3) because these COs directly develop technical and analytical skills critical for the field.

CO1, CO3, and CO4: Moderately related (2) as they contribute to building specialized competencies through data analysis and modelling.

PO5: Capacity for Application, Problem-Solving, and Analytical Reasoning

CO1, CO2, CO3, CO4, CO5, CO6, and CO7: Strongly related (3) as all COs emphasize practical application, problem-solving, and data analysis, aligning directly with this outcome.

PO6: Communication Skills and Collaboration

CO2: Moderately related (2) as interpreting and communicating regression analysis results can involve collaboration.

CO1, CO3 to CO7: Partially related (1) as they may require explanation or presentation of findings, though not directly emphasizing communication.

PO7: Research-related Skills

CO2: Strongly related (3) due to its emphasis on regression analysis, which involves research skills like hypothesis testing and data interpretation.

CO1, CO3, CO4, CO5, CO6, and CO7: Moderately related (2) as they involve research elements like data analysis, formulation of models, and drawing conclusions.

PO8: Learning How to Learn Skills

CO2, CO4, CO5, CO6, and CO7: Moderately related (2) because these COs encourage independent problem-solving and adapting to new information.

CO1, CO3: Partially related (1) since they involve learning new concepts, though not necessarily independently.

PO9: Digital and Technological Skills

CO6, CO7: Strongly related (3) as they specifically involve applying technological solutions to optimize logistics and resource allocation.

CO2, CO4, CO5: Moderately related (2) due to the use of digital tools in data analysis and modeling.

PO10: Multicultural Competence, Inclusive Spirit, and Empathy

CO1 to CO7: Partially related (1) as these COs do not directly address multicultural competence but involve decision-making that may consider diverse perspectives.

PO11: Value Inculcation and Environmental Awareness

CO1 to CO7: Partially related (1) as they indirectly support ethical decision-making in the context of data interpretation and problem-solving.

PO12: Autonomy, Responsibility, and Accountability

CO2, CO4, CO5, CO6, and CO7: Moderately related (2) since they involve independent problem-solving and taking responsibility for decisions.

CO1, CO3: Partially related (1) due to their focus on understanding foundational concepts with less emphasis on autonomy.

PO13: Community Engagement and Service

CO6, CO7: Moderately related (2) as optimizing logistics and resource allocation can have societal impacts.

CO1 to CO5: Partially related (1) since their community impact is more indirect.

CBCS Syllabus as per NEP 2020 for S.Y.B. Sc. Statistics (2023 Pattern)

Name of the Programme : B.Sc. Statistics
Program Code : USST
Class : S.Y.B.Sc.
Semester : IV
Course Type : SEC Practical
Course Code : STA-276-SEC
Course Title : Programming in R and Introduction to Tableau, Power BI
No. of Credits : 2 credits
No. of Teaching Hours : 60

Course Objectives:

1. To provide a thorough understanding of control structures (if-else, for loops, while loops, repeat loops) and their use in R programming.
2. The main focus of Tableau and Power BI software is for better understand datasets.
3. To implement the best design practices, and use the most appropriate chart for a particular situation.
4. To build interactive Tableau and Power BI dashboards and construct a data story using Tableau Story point.
5. To use data hierarchies, filters, groups, sets, and calculated fields.
6. To build data visualizations in Tableau and Power BI
7.

Course Outcomes:

Student will be able to

CO 1. implement control structures (if-else, for loops, while loops, repeat loops) in R to automate and streamline data analysis tasks.

CO 2. develop skills in writing R programs, including debugging and error handling.

CO 3. solve data analysis problems effectively.

CO 4. understand the basic concepts and components of Tableau and Power BI.

CO 5. connect to different data sources and prepare data for analysis using Tableau and Power BI.

CO 6. create and customize a variety of visualizations in Tableau and Power BI.

CO 7. build interactive visualizations and dashboards in Tableau and Power BI.

Topics and Learning Points

Unit 1. Programming in R: (20 L)

Statements: if and if…else, for loop, while loop, cat and print commands Writing programs in R.

1.1. Testing normality of number of samples

1.2. Verifying the assumptions in testing H_0: $\mu = \mu_0$ and then applying appropriate test.

1.3. Verifying the assumptions in testing H_0: $\mu_1 = \mu_2$ and then applying appropriate test.

1.4. Verifying the assumptions in testing H_0: $\mu_1 = \mu_2$ in paired data and then applying appropriate test.

1.5. Verifying the assumptions in testing H_0: $\sigma_1^2 = \sigma_2^2$ and then applying appropriate test.

1.6. Performing number of chi-square tests.

Unit 2. Introduction to Tablaue: (20 L)

2.1 Introduction to Tableau

Overview of Tableau and its role in data visualization and business intelligence, understanding the Tableau interface and key components, basic concepts: Dimensions, Measures, and Data Types.

2.2 Connecting to Data Sources

Introduction to data connections in Tableau, Importing data from Excel, CSV, databases, and other sources, working with live connections vs. data extracts, Preparing and cleaning data within Tableau.

2.3 Creating Basic Visualizations

Introduction to Tableau visualizations (Bar charts, Line charts, Pie charts, etc.), using the Show Me panel to create visualizations, understanding continuous vs. discrete fields, customizing visualizations: colors, labels, tooltips.

2.4 Advanced Visualizations and Calculations

Creating advanced charts: Maps, Scatter plots, Histograms, and Gantt charts, introduction to calculated fields and basic calculations, using Tableau's in-built functions for data manipulation, creating and using parameters in visualizations.

2.5 Building Interactive Dashboards

Introduction to dashboards in Tableau, Combining multiple visualizations into a single dashboard, Adding interactivity with filters, actions, and parameters, Best practices for dashboard design and layout.

Unit 3. Introduction to Power BI: (20 L)

3.1 Connecting to Data Sources

Introduction to data sources supported by Power BI, Importing data from Excel, CSV, and other common formats, Connecting to databases (SQL Server, Azure, etc.), Direct Query vs. Import Mode.

3.2 Data Transformation with Power Query

Introduction to Power Query Editor, Data cleaning and shaping, Transformations: filtering, merging, appending, and grouping data. Creating calculated columns and measures.

3.3 Data Modeling

Understanding data modeling concepts, Creating relationships between tables, Understanding DAX (Data Analysis Expressions) basics, Building calculated columns, measures, and tables.

3.4 Creating Visualizations

Introduction to Power BI visuals (charts, tables, maps, etc.), Customizing, visualizations (formatting, filtering, sorting), Creating slicers and filters, Using visual interaction features.

3.5 Building Interactive Dashboards

Designing effective dashboards, combining multiple visualizations in a single report, Using themes and layouts for dashboards, Best practices for dashboard design.

Sr. No.	Title of the Practical
1.	Conditional Statements with if and if…else
2.	Iteration with for Loops
3.	While Loops
4.	Combining Loops and Conditional Statements
5.	Exploring the Tableau Interface
6.	Connecting to Data Sources Using Tableau
7.	Creating Basic Visualizations Using Tableau
8.	Advanced Visualizations and Calculations Using Tableau
9.	Building an Interactive Dashboard Using Tableau
10.	Exploring the Power BI Interface
11.	Connecting to Data Sources Using Power BI
12.	Data Transformation with Power Query
13.	Data Modelling and Relationships Using Power BI
14.	Creating Visualizations Using Power BI
15.	Building a Dashboard Using Power BI

References:

1. Crawley, M. J. (2006). Statistics - An introduction using R. John Wiley, London 32
2. Purohit, S.G.; Gore, S.D. and Deshmukh, S.R. (2015). Statistics using R, second edition. Narosa Publishing House, New Delhi.

3. Pavagi, V. R. (2018).A Book of Statistical Computing using R Software, Nirali Prakashan.
4. Andy Nicholls, Richard Pugh, Aimee Gott (2016), Sams Teach Yourself R, Person India Education Services Pvt. Ltd.
5. Jared P. Lander (2018), R for Everyone, 2nd edition Person India Education Services Pvt. Ltd.
6. Chandraish Sinha (2022) Mater in Power BI, BPB Publications, New Delhi.
7. Alan Murray (2023). Power BI for Jobseekers, BPB Publications, New Delhi
8. Ryan Sleeper (2018), Practical Tableau, O'Reilly
9. George Peck (2020), Tableau 9: The Official Guide, McGraw Hill

COs and POs Mapping:

Course Outcomes	Programme Outcomes (POs)												
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PO13
CO1	3	3	1	3	3	2	2	3	3	1	1	2	1
CO2	3	3	1	3	3	2	2	3	3	1	1	2	1
CO3	3	3	1	3	3	2	2	3	3	1	1	2	1
CO4	2	2	1	2	2	2	2	2	3	1	1	2	1
CO5	2	3	1	2	2	2	2	2	3	1	1	2	1
CO6	2	3	1	2	2	2	2	2	3	1	1	2	1
CO7	2	3	1	2	2	2	2	2	3	1	1	2	1

PO1. Comprehensive Knowledge and Understanding

CO1, CO2, CO3 (Weightage: 3): These COs involve the fundamental understanding and application of programming and data analysis concepts, strongly aligning with PO1.

CO4, CO5, CO6, CO7 (Weightage: 2): These COs cover software tools (Tableau and Power BI) that are essential but supplementary to foundational programming knowledge, providing a moderate alignment with PO1.

PO2. Practical, Professional, and Procedural Knowledge

CO1, CO2, CO3, CO5, CO6, CO7 (Weightage: 3): These COs emphasize the practical application of programming and data visualization tools, which is essential for professional tasks, strongly aligning with PO2.

CO4 (Weightage: 2): Understanding the basics of Tableau and Power BI contributes to practical knowledge, but it is a preliminary step, making it moderately related to PO2.

PO3. Entrepreneurial Mindset and Knowledge

CO1 to CO7 (Weightage: 1): The COs focus more on technical and analytical skills rather than fostering an entrepreneurial mindset, making them only partially related to PO3.

PO4. Specialized Skills and Competencies

CO1, CO2, CO3 (Weightage: 3): These COs require a high level of technical proficiency in programming and problem-solving, strongly aligning with PO4.

CO4, CO5, CO6, CO7 (Weightage: 2): These COs involve skills in data visualization tools, which are important but supplementary, leading to a moderate alignment with PO4.

PO5. Capacity for Application, Problem-Solving, and Analytical Reasoning

CO1, CO2, CO3 (Weightage: 3): The course outcomes related to programming and data analysis are directly aligned with problem-solving and analytical reasoning, strongly relating to PO5.

CO4, CO5, CO6, CO7 (Weightage: 2): These COs involve applying data visualization tools, contributing to problem-solving but to a moderate extent.

PO6. Communication Skills and Collaboration

CO1 to CO7 (Weightage: 2): The course outcomes involve communication through code and visualizations, which contributes to communication skills, though not as directly as PO6 implies.

PO7. Research-related Skills

CO1 to CO7 (Weightage: 2): The COs involve data analysis and visualization, which are research-related skills, though not directly focused on research methodologies, making the alignment moderate.

PO8. Learning How to Learn Skills

CO1 to CO7 (Weightage: 3): The course outcomes encourage self-directed learning in programming and software tools, strongly aligning with PO8.

PO9. Digital and Technological Skills

CO1 to CO7 (Weightage: 3): All COs focus heavily on using digital tools and software (R, Tableau, Power BI), strongly aligning with PO9.

PO10. Multicultural Competence, Inclusive Spirit, and Empathy

CO1 to CO7 (Weightage: 1): The course outcomes do not directly address multicultural competence or empathy, leading to a partial alignment with PO10.

PO11. Value Inculcation and Environmental Awareness

CO1 to CO7 (Weightage: 1): The course is focused on technical skills and does not directly contribute to value inculcation or environmental awareness, resulting in a partial alignment with PO11.

PO12. Autonomy, Responsibility, and Accountability

CO1 to CO7 (Weightage: 2): The independent nature of programming and data visualization tasks requires a moderate degree of autonomy and responsibility, aligning moderately with PO12.

PO13. Community Engagement and Service

CO1 to CO7 (Weightage: 1): The course is not primarily focused on community engagement, resulting in a partial alignment with PO13.