KANNUR UNIVERSITY

(Abstract)

M.Sc Biostatistics in the Dept. of Statistical Sciences, Mangattuparamba Campus - Scheme (Credit distribution for courses) & Syllabus of the First Semester -Approved & Implemented w.e.f 2023 admission - Orders issued

---

ACADEMIC C SECTION

ACAD/ACAD C3/365/2023 Dated: 21.11.2023

Read:-1. GO (Ms) No 528/2022/HEDN dated 22/10/2022

2. Letter No Acad C3/20143/2022 dated 26/11/2022

3. Email dated 07/01/2023 from the Head, Dept of Statistical Sciences, Mangattuparamba Campus.

4. U.O of even number dated 07/02/2023

5. Minutes of the online meeting of the Expert Committee dated 21/03/2023.

6. UO of even number dated 22/06/2023

7. UO No ACAD C/ACAD C3/22373/2019 dated 12/09/2023

8. Email dated 05/11/2023 from the Head, Dept of Statistical Sciences, Mangattuparamba Campus.

9. Minutes of the meeting of the Department Council dated 18/09/2023

ORDER

1. As per paper read (1) above, Govt. of Kerala granted Administrative Sanction for starting Project Mode Programme *M.Sc Biostatistics with Specialization in Epidemiology and SAS Programming* in the Dept. of Statistical Sciences, Mangattuparamba campus, Kannur University during the Academic Year 2022-'23.

2. As per paper read (2) above, HoD, Dept. of Statistical Sciences was requested to prepare and submit the draft Scheme & Syllabus for the aforementioned Programme along with a panel of five-member Experts to constitute a committee to scrutinize the syllabus.

3. As per paper read (3) above, HoD, Dept. of Statistical Sciences submitted the Scheme and Syllabus for the Programme *MSc Biostatistics with Specialization in Epidemiology and SAS Programming* along with a panel of ten experts to scrutinize the syllabus.

4. As per paper read (4) above, a ten member Expert Committee was constituted, with the Head , Dept of Statistical Sciences as the Coordinator, to scrutinize the Scheme & Syllabus of the aforesaid programme. Head , Dept. of Statistical Sciences was authorised to submit the final Scheme & Syllabus of the Programme after incorporating the corrections/modifications , if any, suggested by the Expert Committee.

5 . As per paper read (5) above, the Expert Committee suggested to change the nomenclature of the Programme as ***MSc Biostatistics*** instead of *MSc Biostatistics with Specialization in Epidemiology and SAS Programming*.

6. Accordingly, as per paper read (6) above, nomenclature of the Programme is changed from *MSc Biostatistics with Specialization in Epidemiology and SAS Programming* to ***MSc***

Biostatistics.

7. As per paper read (7) above, Revised Regulations for PG Programmes under CBCSS in the University Teaching Depts./ Schools was implemented w.e.f 2023 admission.

8. As per paper read (8) above, the Head, Dept. of Statistical Sciences submitted the Scheme (Distribution of credits for courses) & Syllabus of the First Semester MSc Biostatistics Programme w.e.f 2023 admission, incorporating the suggestions of the Expert Committee and also in accordance with the revised Regulations implemented in the Teaching Departments / Schools of Kannur University w.e.f 2023 admissions. Department Council in its meeting held on 18/09/2023 vide paper read (9) above approved the aforesaid syllabus.

9. The Vice Chancellor, after considering the matter in detail and in exercise of the powers of the Academic Council conferred under section 11(1), Chapter III of Kannur University Act 1996, approved the **Scheme (distribution of credits) & Syllabus of the First Semester MSc Biostatistics Programme** subject to reporting to the Academic Council and accorded sanction to implement the syllabus of M.Sc Biostatistics Programme in the Department of Statistical Sciences, Mangattuparamba Campus w.e.f 2023 admission.

10.Scheme (Distribution of Credits) & Syllabus of the First Semester MSc Biostatistics Programme implemented in the Department of Statistical Sciences, Mangattuparamba Campus with effect from 2023 admission, is appended and uploaded in the University website (www.kannuruniversity.ac.in)

11. Orders are issued accordingly.

Sd/-

Narayanadas K

DEPUTY REGISTRAR (ACAD)

For REGISTRAR

To: 1.Head, Dept of Statistical Sciences, Mangattuparamba Campus
2.Convener, Curriculum Committee.

Copy To: 1. To Exam Branch (through PA to CE)
2. PS to VC/ PA to PVC/ PA to R/PA to CE
3. EP IV / EXCI/SWC 4 . DR/ AR I /AR II (Acad)
5 . To Webmanager (to publish in the website) 6 . Computer Programmer
7 . SF/DF/FC

Forwarded / By Order

SECTION OFFICER

KANNUR UNIVERSITY CIVIL STATION P.O. KANNUR Pin-670 002 THAV

KANNUR UNIVERSITY

(Abstract)

M. Sc Biostatistics Programme in the Dept of Statistical Sciences,Mangattuparamba Campus of Kannur University - Scheme & Syllabus of II, III & IV Semesters - Approved - Implemented w. e. f 2023 admission - Orders Issued.

ACADEMIC C SECTION

ACAD/ACAD C3/365/2023 Dated: 31.05.2024

Read:-1. U.O.s No ACAD C/ACAD C3/22373/2019 dated 12/09/2023, 08/11/2023 & 16/02/2023

2. U.O of even number dated 21/11/2023

3. Circulars of No ACAD C/ACAD C3/22373/2019 dated 01/02/2024 & 12/03/2024

4. Email dated 23/04/2024 from the Head , Dept of Statistical Sciences, Mangattuparamba Campus

5. Minutes of the meeting of the Department Council dated 22/04/2024

ORDER

1. The revised Regulations for PG Programmes under CBCSS in the University Teaching Depts/ Schools were implemented w. e. f 2023 admissions vide paper read (1) above.

2. As per paper read (2) above, Revised Scheme (Credit distribution of courses) & Syllabus (I Semester Only) of M. Sc Biostatistics Programme was approved and implemented in the Dept of Statistical Sciences, Mangattuparamba Campus w. e. f 2023 admission.

3. As per paper read (3) above, Heads of all Teaching Depts who had not submitted the syllabi in full, were requested to submit the syllabi of the remaining semesters in accordance with the approved Regulations and along with a copy of the Department Council Minutes.

4. As per paper read (4) above, the Head, Dept of Statistical Sciences submitted the Scheme & Syllabus (II, III & IV Semesters) of M. Sc Biostatistics Programme to be implemented in the University Teaching Dept w.e.f 2023 admission.

5. Dept Council vide paper read (5) above, recommended the aforementioned Syllabus of M. Sc Biostatistics Programme to be implemented in the Dept of Statistical Sciences, Mangattuparamba Campus w.e.f 2023 admission .

6. The Vice Chancellor, after considering the matter in detail and in exercise of the powers of the Academic Council conferred under Section 11(1), Chapter III of Kannur University Act 1996, ***approved the Scheme & Syllabus (II, III & IV Semesters) of M. Sc Biostatistics Programme and accorded sanction to implement the same in the Dept of Statistical Sciences, Mangattuparamba Campus w. e. f 2023 admission, subject to report to the Academic Council***.

7. The Scheme & Syllabus (Ist, II nd, III rd & IV th Semesters) of M. Sc Biostatistics Programme

under CBCSS, implemented in the Dept of Statistical sciences, Mangattuparamba Campus w. e. f 2023 admission, is appended and uploaded in the University website (www.kannuruniversity.ac.in).

8. Orders are issued accordingly.

Sd/-

Narayanadas K

DEPUTY REGISTRAR (ACAD)

For REGISTRAR

To: 1. Head, Dept of Statistical Sciences, Mangattuparamba Campus

2. Convenor, Curriculum Committee

Copy To: 1. PS to VC/ PA to R

2. PA to CE (to circulate among the sections of the Examination Branch concerned)

3. EP IV/ EX C1

4. Computer Programmer

5. Webmanager (to publish in the website)

6. SF/DF/FC

Forwarded / By Order

SECTION OFFICER

KANNUR UNIVERSITY

M.Sc. BIOSTATISTICS

SCHEME & SYLLABUS

(*Under Choice Based Credit & Semester System)*

2023 Admission Onwards

DEPARTMENT OF STATISTICAL SCIENCES

Mangattuparamba Campus

KANNUR UNIVERSITY

Post Graduate Programme in Biostatistics

M.Sc. Biostatistics programme is a two-year programme divided into four semesters. A student is required to complete at least 80 credits for the completion of the programme and the award of degree.

DURATION: 2 Years (4 semesters)

INTAKE: 15.

OBJECTIVES OF THE PROGRAMME

1. Gain sound knowledge in theoretical and practical aspects of Biostatistics.
2. Acquire the working knowledge of various statistical software and programming languages.
3. Acquire skills and competencies in Biostatistical computing methods and develop algorithms and computer programmes for analyzing complex datasets.
4. Communicate effectively complex statistical ideas to people working in diverse spheres of academics and organizational setups.
5. Handle and analyze large databases related to various biomedical research and make meaningful interpretations of the results.
6. Get wide range of job opportunities in industry as well as in government sector.
7. Make unique contribution for the development of discipline by addressing complex and challenging problems in emerging areas of the discipline.
8. Imbibe effective scientific and/or technical communication in both oral and writing.
9. Continue to acquire relevant knowledge and skills appropriate to professional activities and demonstrate highest standards of ethical issues in Biostatistical sciences.

ELIGIBILITIES:

The selection procedure will be based on an entrance examination by the University. The eligibility criteria for appearing entrance examination is any of the following degree with overall 50% marks:

1. B.Sc. Statistics/Biostatistics as core course.
2. B.A./B.Sc. Mathematics
3. B.Sc. Computer Science with Statistics/Mathematics as complementary course
4. B. Tech/B.E degree.
5. B.Sc. with Mathematics and Statistics as core courses.

ADMISSION:

- The selection of the candidate is mainly based on the marks secured in the Degree Course/Admission test.
- The admission test will cover statistics and mathematics at the undergraduate level.

Relaxation & Weightage

Relaxation and weightage will be as per Kannur University rule.

COURSE DETAILS:

A student must register for the required number of courses at the beginning of each semester. No students shall register for more than 28 credits and less than 16 credits per semester.

A total of 80 credits shall be the minimum for successful completion of the course in which minimum of 50% of credits have to be earned from Discipline Specific Courses including dissertation for any programme. Those who secure only minimum credit for core/ elective subjects has to supplement the deficiency for obtaining the minimum total credits required for successful completion of the program from the other divisions.

EVALUATION:

The faculty member who teaches the course shall do evaluation of the students for each course on the basis of Continuous Evaluation and End Semester Examination shall be evaluated by External Examiners. The proportion of the distribution of marks among the continuous evaluation and end semester examination shall be **40:60**.

Continuous Evaluation includes assignments, seminars, written examination and viva voce for each course. Weightage to the components of continuous evaluation shall be given for all theory papers of the course as follows:

Components of CE	Minimum Number	Weightage	Grade Points	Practical Weightage	Grade Points
Test paper	2	40	16	80	-
Assignments	1	20	08	--	--
Seminar presentation, Viva Voce, Discussion, Debate etc.	1	40	16		--
Record	--	--	--	20	-

Test Paper: For each course there shall be at least two class tests during a semester.

Assignments: Each student shall be required to do one assignment for each course.

Seminar: Students are required to present a seminar on a selected topic in each paper. The evaluation of the seminar shall be done by the concerned teacher handling the course.

Viva Voce – End semester theory Viva Voce examination will be conducted for each paper before the commencement of public examination.

Attendance: Minimum attendance required for each paper shall be 75% of the total number of classes conducted for that semester. Those who secured the minimum requirement of attendance only be allowed to register/appear for End Semester Examination.

Condonation of attendance to a maximum of 10 days in a semester subject to a maximum of two times during the whole period of the PG program may be granted by the university as per university rules.

Conduct of Examination:

The Vice Chancellor will approve the panel of examiners submitted by the Head of the Department. All the teachers of the Department will be the members of the Board of examiners with Head of the Department as the Chairperson, there shall be an external examiner.

Research Project:

The students have to complete a research project during IV Semester in collaboration with any of the authorized research institutions located within or outside the state or within their own Department.

KANNUR UNIVERSITY

DEPARTMENT OF STATISTICAL SCIENCES

VISION

Motivated by optimism and responsibility, the vision is to develop an exemplary centre for studies, practice and research in Statistics which will be beneficial to the stakeholders and the society.

MISSION

To develop an excellent centre of quality teaching and research in Statistics

To develop an international centre for advanced statistical computing and data analysis.

PROGRAMME OUTCOMES

PO 1 : **Critical Thinking:** Take informed actions after identifying the assumptions that frame our thinking and actions, checking out the degree to which these assumptions are accurate and valid, and looking at our ideas and decisions (intellectual, organizational, and personal) from different perspectives.

PO2 : **Problem Solving:** Identify, formulate, conduct investigations, and find solutions to problems based on in-depth knowledge of relevant domains.

PO 3 : **Communication:** Speak, read, write and listen clearly in person and through electronic media in English/language of the discipline, and make meaning of the world by connecting people, ideas, books, media and technology.

PO 4 : **Responsible Citizenship:** Demonstrate empathetic social concern, and the ability to act with an informed awareness of issues.

PO 5 : **Ethics**: Recognize different value systems including your own, understand the moral dimensions of your decisions, and accept responsibility for them.

PO 6 : **Self-directed and Life-long Learning:** Acquire the ability to engage in independent and life-long learning in the broadest context socio-technological changes.

PROGRAMME SPECIFIC OUTCOME

PSO 1: Expertise in the field of biostatistical theory and its practical applications.

PSO 2: Expertise to take up responsibilities as efficient Biostatisticians/Statistical Officers/Research Officers/ Statistical Analytics.

PSO 3: Expertise on techniques of biostatistics and in the field of data analysis.

PSO 4: Make Awareness on recent trends in biostatistical theory and applications.

PSO 5: Utilize statistical methods and tools to analyze data sets, draw meaningful conclusions, and make informed decisions based on biostatistical inferences.

PSO6: Demonstrate proficiency in using statistical software such as R and SPSS, to perform statistical computations, visualize data, and facilitate biostatistical analysis.

COURSE OUTCOME

CO1 : Demonstrate an in-depth understanding of Biostatistical concepts, including advanced clinical trials, statistical epidemiology, demography, sampling and design, statistical inference, regression analysis, probability and distribution theory.

CO 2 : Apply biostatistical techniques to analyze real life data using statistical packages such as SPSS, SAS and free software R and Python. .

CO 3 : Formulate suitable models for pharmaceutical research and drug development.

CO 4 : Apply statistical techniques to analyze medical data which enables the students to develop critical thinking skills and draw meaningful conclusions from complex datasets.

CO 5 : Develop research skills, including literature review, problem formulation, data collection, experimental design, and statistical analysis, to conduct independent biomedical research.

Distribution of Grades for the M. Sc. Biostatistics Programme with effect from 2023-24 Onwards									
	1	2	3	4	5	6	7	8	Total Credits
	Discipline Specific		Electives						
Semester	Core Courses (DSC)	Electives (DSE)	Interdisciplinary/ Multidisciplinary Elective	AEC 2 Credits	SEC(SEC) 2Credits	VAC /MOOC 2Credits	Internship /Field Visit /Minor Project /Institutional/Industrial Visit 2Credits	Dissertation / Major Project	
1	MSBST01DSC01 MSBST01DSC02 MSBST01DSC03 MSBST01DSC04	**Pool A** MSBST01DSE01 to 02 (any 1)							**19**
	4 Credits x 4 = 16 Credits	**3 Credits**							
2	MSBST02DSC05 MSBST02DSC06 MSBST02DSC07 MSBST02DSC08	**Pool B** MSBST02DSE03 to 04 (any 1)		**Pool C**	**Pool D**				**23**
	4 Credits x 4 = 16 Credits	**3 Credits**		**2 Credits**	**2 Credits**				
3	MSBST03DSC09 MSBST03DSC10	**Pool E** MSBST03DSE05 to 06 (any 1) **Pool F** MSBST03DSE07 to 12 (any 2)	**Pool G** To be obtained from other Departments			VAC/ MOOC	MSBST03DSC11		**23**
	4 Credits x 2 = 8 Credits	**3Credits x 3= 9 credits**	**4 Credits**			**2* Credits**	**2 Credits**		
4		**Pool H** MSBST04DSE13 to 14 (any 1) **Pool I** MSBST04DSE15 to 21 (any 1)						MSBST04DSC12	**18**
		3Credit x 2= 6 Credits						**12 Credits**	
	Total Credit for M. Sc. Biostatistics Program								**83**

*Credits are over and above the total credit requirement.

	FIRST SEMESTER								
Sl No	**Course Code**	**Title of Paper**	**Contact Hours/Week**			**Marks**			
			L	**T/S**	**P**	**ESE**	**CE**	**Total**	**Credits**
DISCIPLINE SPECIFIC CORE COURSES (DCE)									
1.1	MSBST01DSC01	Mathematical Methods for Biostatistics	4	1		60	40	100	4
1.2	MSBST01DSC02	Probability and Distribution Theory	4	2		60	40	100	4
1.3	MSBST01DSC03	Sampling Methods	4	1		60	40	100	4
1.4	MSBST01DSC04	Introduction to Biostatistics	4	2		60	40	100	4
DISCIPLINE SPECIFIC ELECTIVE COURSE (DSE)									
1.5	MSBST01DSExx	Elective-I-DSE (Pool A)		2	6	60	40	100	3
Total credits									**19**

L=Lecture, T/S=Tutorials/Seminar, P/I=Practical/Internship, CE =Continuous Evaluation, ESE = End Semester Evaluation

Sl No	**POOL A:- List of Courses for Elective -I DISCIPLINE SPECIFIC ELECTIVES (DSE)**								
1.5a	MSBST01DSE01	Biostatistical Computing Using R - I (**Practical**)		2	6	60	40	100	3
1.5b	MSBST01DSE02	Biostatistical Computing Using SPSS - I (**Practical**)		2	6	60	40	100	3

FIRST SEMESTER M.Sc. BIOSTATISTICS PROGRAMME

DISCIPLINE SPECIFIC CORE COURSE

Course Code & Title	**MSBST01DSC01-MATHEMATICAL METHODS FOR BIOSTATISTICS**	
Programme	**M.Sc. Biostatistics**	**Semester I**
Course Objectives	● To introduce the concept of sequence and series. ● To understand the improper integrals, beta and gamma functions. ● Learn Taylor's Theorem with applications. ● Describe optima of functions using examples. ● To achieve ideas on vector space, subspaces, independence of vectors, basis and dimension. ● Establish the relation between algebraic and geometric multiplicity. ● To achieve ideas on quadratic forms and reduction of quadratic forms and gets ability for solving problems in these areas.	

Modules	Content	Module Outcome
Module I: Sequence and series **(15 Hours)**	Sequences, series and their convergence, limit superior, limit inferior, limit of sequences, Cauchy sequence. Comparison test, D'Alembert's ratio test, Cauchy's root test, Raabi's test, Gauss test, Cauchy's integral test, Absolute convergence of series, Leibnitz's test for the convergence of alternating series, conditional convergence, indeterminate form, L'Hospital 's rule (problems only).	The students will be able to: ● Explain convergences of sequences and series. ● Solve problems using various tests to examine the convergences of series. ● Explain the concept of L Hospital's Rule

Module II: Special functions **(15 Hours)**	The beta and gamma functions, duplication formula for gamma function, incomplete beta and gamma functions, functions of several variables, Limits and continuity, Taylor's theorem and its applications, Conditions for the optima of multivariate functions, Lagrange's method for constrained optimum, examples (bivariate case only)	• Explain proper and improper beta and gamma functions. • Understand the calculus of multivariable functions • To find local and global optima of functions.
Module III: Vectors and Matrices **(15 Hours)**	Vector space, Subspaces, Linear dependence and independence, Basis and dimensions, Matrices and determinants, symmetric, orthogonal and idempotent matrices, Row and column space of matrix, Rank, inverse, Characteristic polynomial, Cayley-Hamilton Theorem (statement and problem).	• To be familiar with vector space, subspace and examples. • Explain linear dependence and independence. • State Cayley-Hamilton theorem and solve problems.
Module IV: Eigen values and spectral decomposition **(15 Hours)**	Eigen values and eigen vectors, Spectral decomposition, Algebraic and geometric multiplicities, Generalized inverse, Quadratic forms, Classification of quadratic forms, Properties and reductions.	• Determine the Eigen values and Eigen vectors of the given matrix • Write down the spectral decomposition of the given matrix • Explain different types of quadratic forms.
Reference	***Text Books*** 1. Malik, S.C & Arora, S. (2006***).*** *Mathematical Analysis, Second Edition***,** New-age international publishers. 2. Mathai, A. M. & Haubold, H. J. (2017***).*** *Linear Algebra – A course for Physicists and Engineers***,** De Gruyter, Germany.	

	Reference Books 1. Rudin, W. (2013). *Principles of Real Analysis (3rdEd.)*McGraw Hill. 2. Ramachandra Rao & Bhimasankaran (1992). *Linear Algebra.* Tata McGraw Hill, New Delhi. 3. Apsostol, T. M. (1974). *Mathematical Analysis, Second Edition*. Narosa, New Delhi. 4. Rao, C. R. (2002). Linear Statistical Inference and Its Applications, Second Edition, John Wiley and Sons, New York.
Course Outcomes	After successful completion of this course, student will be able to: 1. Understand the concepts of limit and continuity of functions and their properties 2. Understand convergence of sequences and series of real numbers and functions. 3. Understand the vector space, matrices and its properties 4. Understand the properties of quadratic forms and its reduction.

TEACHING LEARNING STRATEGIES

- **Lecturing, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	**Weightage**
End Semester Evaluation	**60**
Continuous Evaluation	
Tests	16
Assignment	08
Seminar	16
Total	**40**

Sample Questions to Test Outcomes:

1. Define limit of sequence of real numbers. Give an example of a sequence for which limit does not exist.

2. What is meant by absolute convergence of series?

3. What is incomplete gamma function?.

4. State conditions for the optima of a multivariate function.

5. State Cayley-Hamilton theorem.

6. Explain Gram-Schmidth orthogonalization process.

7. Write a short note on different types of quadratic forms.

DISCIPLINE SPECIFIC CORE COURSE

Course Code & Title	**MSBST01DSC02- PROBABILITY AND DISTRIBUTION THEORY**		
Programme	**M.Sc. Biostatistics**	**Semester**	I
Course Objectives	The Course aims • To introduce the basic concepts of probability • To understand the connection between three approaches of definitions of probability. • To get an idea on important theorems in probability using axiomatic definition of probability • To learn about various discrete and continuous probability distributions needed for biostatistical analysis.		

Modules	Content	Module Outcome
Module I: Probability and Random Variables **(15 Hours)**	Computation of probability based on classical and empirical definitions. Axiomatic approach to probability, probability space, conditional probability space, independence of events, Bayes' theorem and examples, random variable, distribution function, density function, expectation, variance and moments of a random variable and properties.	• Understand various definitions of probability • Conditional probability and Bayes' theorem • Concept of random variable and their distributions
Module II: Important large sample theorems **(15 Hours)**	Definition of moment generating function and its limitations, characteristic function, elementary properties, characteristic functions and moments. Sequence of random variables, various modes of convergence of sequence random variables (definition only), Weak law of large numbers, strong law of large numbers, central limit theorem, DeMoivre-Laplace and Lindbergh-Levy forms of CLT. Applications of CLT in biostatistics.	• Definition of characteristic function • Concept of weak and strong laws of large numbers • Concept of central limit theorem and its applications in biostatistics
Module III: Special Discrete Distributions **(15 Hours)**	Discrete Uniform, Bernoulli, Binomial, Poisson, Geometric, Negative binomial, Hyper geometric, Multinomial. Properties of these distributions. Sample simulation and fitting of discrete distributions.	• Explain different discrete distributions. • Properties of discrete distributions • Simulation of samples from standard discrete distributions

Module IV: Special Continuous Distributions **(15 Hours)**	Continuous Uniform, Exponential, Beta, Gamma, Normal, Weibull, Pareto, Laplace, Logistic, Cauchy and log-normal distributions. Properties of these distributions. Sample simulation and fitting of continuous distributions	• Explain different continuous distributions • Properties of continuous distributions • Simulation of samples from standard continuous distributions
References	***Text Books*** 1. Krishnamurthy, K.(2006). *Handbook of Statistical Distributions with Applications* .Chapman & Hall/CRC, New-York 2. Schinazi, R.B. (2010). *Probability with Statistical Applications-Second Ed* . Springer, New York. ***Reference Books*** 1. Bhat, B.R. (2004). *Modern Probability Theory*, New Age Publishers, New Delhi. 2. Rohatgi, V. K. (2020). *An Introduction to Probability Theory and Mathematical Statistics*, Wiley Eastern. 3. Johnson, N.L., Kotz, S.and Balakrishnan, N. (1995). *Continuous Univariate Distributions, Vol. I &Vol. II,* John Wiley and Sons, New-York. 4. Johnson, N.L., Kotz. S. and Kemp. A.W.(1992). *Univarite Discrete Distributions*, John Wiley and Sons, New York.	
Course Outcomes	After successful completion of this course, student will be able to: 1. Understand the concepts of probability and properties. 2. Understand characteristic function and its properties 3. Understand various laws of large numbers and central limit theorems. 4. Understand the concepts of discrete and continuous distributions. 5. Understand the normal distribution and various non-normal distributions, their properties and applications for scientific research.	

TEACHING LEARNING STRATEGIES

- **Lecturing, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	Weightage
End Semester Evaluation	**60**
Continuous Evaluation	
Tests	16
Assignment	08
Seminar	16
Total	**40**

Sample Questions to test Outcomes:

1. Define Poisson random variable. Find the moment generating function of a Poisson random variable.
2. Show that in the case of binomial distribution mean is always greater than variance, however, mean equal to variance in the case of Poisson distribution.
3. Define t-statistic and explain its important applications. Write down the probability density function of Student's t-distribution.
4. Obtain the characteristic generating function of a standard normal distribution.
5. Define bivariate normal distribution. Show that linear combination of independent normal variables is normally distributed.
6. Define chi-square distribution. Obtain the MGF of the Chi-square distribution. Use it to obtain the mean and variance.

DISCIPLINE SPECIFIC CORE COURSE

Course Code & Title	**MSBST01DSC03- SAMPLING METHODS**		
Programme	**M.Sc. Biostatistics**	**Semester**	I
Course Objectives	• Explain different types of sampling • Explain different errors in sampling • Difference between SRSWR and SRSWOR • Concept of stratified random sampling • Explain systematic sampling • Explain ratio and regression estimators		

Modules	Content	Module Outcome
Module I: Sampling theory and Simple random sampling **(15 Hours)**	Introduction to sampling theory, Errors in sampling, simple random sampling (with and without replacement)-estimation of population mean and population mean square, determination of sample size, comparing efficiency of SRSWOR with SRSWR, simple random sampling with attributes.	• Concept of sampling theory • Explain different types of errors • Differentiate between SRSWR and SRSWOR • Explain SRS with attributes
Module II: Stratified random sampling and allocations. **(15 Hours)**	Stratified random sampling- estimation of population mean and variance, methods of allocation of sample size to different strata, comparison of allocations.	• Concept of stratified random sampling • Explain methods of allocations.

Module III: Complex sampling schemes **(15 Hours)**	Systematic sampling, circular systematic sampling. Cluster sampling, multistage sampling, multiphase sampling.	• Explain circular systematic sampling • Explain cluster sampling • Explain two stage cluster sampling
Module IV: Auxiliary information based sampling **(15 Hours)**	Ratio and regression methods of estimation- bias and appropriate variances, unbiased ratio estimator, difference estimator, comparison of ratio estimator with regression estimator, Probability proportional to size sampling.	• Concept of ratio estimator • Explain regression estimator • Explain difference estimator • Concept of PPS sampling
References	***Text Books*** 1. Singh, D and Chowdhary, F.S. (1986). *Theory and Analysis of Sample Survey Designs*, New Age International, New Delhi. 2. Cochran. W.G. (2007). *Sampling Techniques*, John Wiley & Sons, New York. ***Reference Books*** 1. Des Raj, D. and Chandhok, P. (1998). *Sample Survey Theory*, Narosa Publishing House, New Delhi. 2. Gupta and Kapoor (2010). *Fundamentals of Applied Statistics*. Sulthan Chand & Sons. 3. Murthy, M.N. (1967). *Sampling Theory & Methods*. Statistical Publishing Society, Calcutta. 4. Parimal Mukopadhyay (2012). *Theory & Methods of Survey Sampling*, PHI Learning, New Delhi.	

Course Outcomes	After successful completion of this course, student will be able to: 1. Understand the concepts of probability and non-probability sampling. 2 Understand the estimation methods for population mean, total and proportion under various sampling schemes. 3 Understand the use of auxiliary information for the estimation various population parameters

TEACHING LEARNING STRATEGIES

- **Lecturing, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	**Weightage**
End Semester Evaluation	**60**
Continuous Evaluation	
Tests	16
Assignment	08
Seminar	16
Total	**40**

Sample Questions to test Outcomes:

1. Explain probability sampling and non probability sampling.
2. Define standard error of sample mean and explain its uses in the construction of confidence interval, testing hypothesis and to obtain p-value.
3. Explain circular systematic sampling.
4. Prove that sample mean is an unbiased estimate of population mean under stratified random sampling.
5. Distinguish between ratio estimators and regression estimators.
6. Explain the difference between the methods of SRS and varying probability scheme.

DISCIPLINE SPECIFIC CORE COURSE

Course Code & Title	**MSBST01DSC04– INTRODUCTION TO BIOSTATISTICS**		
Programme	**M.Sc. Biostatistics**	**Semester**	I
Course Objectives	• Introduce the concept of proportions, ratio and odds • Introduce the concept of risk, relative risk and their measurement • Explain applications of various probability models in medical research • Explain concept of estimation and its applications in biostatistics • Introduce the concept of hypothesis testing and applications in clinical research.		

Modules	Content	Module Outcome
Module I: Descriptive Methods for Categorical Data **(15 Hours)**	Proportions:- Comparative Studies, Screening Tests, Displaying Proportions. Rates:- Changes, Measures of Morbidity and Mortality, Standardization of Rates. Ratios:- Relative Risk, Odds and Odds Ratio, Generalized Odds for Ordered 2xk Tables, Mantel–Haenszel Method, Standardized Mortality Ratio.	• Explain proportions, ratios and rates • Explain risk, relative risk and odds ratio. • Explain Mantel–Haenszel Method and Standardized Mortality Ratio

Module II: Descriptive Methods for Continuous Data **(15 Hours)**	Tabular and Graphical Methods:- One-Way Scatter Plots, Frequency Distribution, Histogram and the Frequency Polygon, Cumulative Frequency Graph and Percentiles, Stem-and-Leaf Diagrams, Measures of Location, Measures of Dispersion, Box Plots, Special Case of Binary Data, Coefficients of Correlation, Pearson's Correlation coefficient, Nonparametric Correlation coefficients.	• Explain graphical and tabular methods • Explain measures of central tendency and dispersion • Explain parametric and non parametric correlations
Module III: Probability Models for data and estimation **(15 Hours)**	Practical applications of Normal, Binomial and Poisson distributions in bio medical research, Pair-Matched Case–Control Study, Introduction to Confidence interval Estimation, Estimation of Proportions, Estimation of Odds Ratios, Estimation of Correlation Coefficients.	• Explain various applications of probability models in medical research • Estimation of proportions, odds ratio and correlation coefficients • Give an introduction to Pair-Matched Case–Control Study and confidence interval estimation
Module IV: Introduction to Statistical Tests of Significance **(15 Hours)**	Hypothesis Tests, Statistical Evidence, Errors, Summaries and Conclusions, Rejection Region, P Values, Type I and Type II Errors, Relationship to Confidence Intervals. One-Sample Problem with Binary Data, Analysis of Pair-Matched Data, Comparison of Two Proportions, Mantel–Haenszel Method, Inferences for General Two-Way Table, Fisher's Exact Test, Ordered 2x k Contingency Tables.	• Explain basic concepts of hypothesis testing and P value • Comparison of Population proportions • Understand , Mantel–Haenszel Method, and Fisher's Exact Test.

References	***Text Book*** 1. Chap T.L. (2003). *Introductory Biostatistics*, John Wiley & Sons. ***Reference Books*** 1. Rosner, B. (2010). *Fundamentals of Biostatistics*, Cenage Learning, Harvard University. 2. Chernick, M.R. and Fris, R.H. (2003). *Introductory Biostatistics for the Health Sciences,* John Wiley & Sons. 3.Peter Armitage, Geoffrey Berry, J. N. S. Matthews (2008).· *Statistical Methods in Medical Research*. John Wiley & Sons 4. Daniel, Wayne W (2009). *Biostatistics: A Foundation for Analysis in the Health Sciences*.John Wiley & Sons.
Course Outcomes	After successful completion of this course, student will be able to: 1. Understand the descriptive methods for different types of data. 2. Understand the concepts of risk, odds and odds ratio. 3. Understand the concept of inferential procedures for medical research. 4. Understand different methods of comparison of proportions for biostatistical studies.

TEACHING LEARNING STRATEGIES

- **Lecturing, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	**Weightage**
End Semester Evaluation	**60**
Continuous Evaluation	
Tests	16
Assignment	08
Seminar	16
Total	**40**

Sample Questions to Test Outcomes:

1. Define odds and odds ratio.
2. Explain sensitivity and specificity.
3. Describe relative risk. How to quantify it.
4. Explain Mantel–Haenszel Method.
5. Explain contingency table.
6. Describe Pair-Matched Case–Control Study.

POOL A: DISCIPLINE SPECIFIC ELECTIVES

Course Code & Title	**MSBST01DSE01 - BIOSTATISTICAL COMPUTING USING R-I (PRACTICAL)**		
Programme	**M.Sc. Biostatistics**	**Semester**	**I**
Course Objectives	• Define the basic concepts of R software and R packages • Describe various concepts required for developing the R Language • Build our new functions in R • Illustrate different R-Graphics facilities • Find rank and inverse using R software • Describe different sampling methods using R software		

Modules	**Content**	**Module Outcome**
Module I: Basic Concepts of R Programming **(20 Hours)**	Introduction to R- Objects and their classes, operators, vectors and matrices, list and data frames, indexing and accessing data, importing and exporting data. Common built-in functions, R-Graphics.	• Define basics of statistical software R • Demonstrate the important data structures such as arrays, matrix, data frames, Class function etc. • Design an overview of the R Language such as Expressions, Objects, Symbols, Functions.

Module II: Matrices and Standard Probability Distributions **(25 Hours)**	Matrices, rank, determinants and inverse. Eigen values and vectors, power of matrices, g-inverse, system of linear equations, roots of algebraic and transcendental equations. Plotting of cdf and pdf of standard distributions. Generations of random samples from standard distributions, demonstrations of the sampling distributions	• How to find rank and inverse using R software. • How to solve system of linear equations using R software • Plotting pdf and cdf curve of different distributions
Module III: Biostatistical Sampling Methods **(25 Hours)**	Random samples elections, estimation of mean pro-portion, variance, confidence interval and efficiency under SRS, stratified random sampling, Various kind of allocation, stratification, estimators based on ratio and regression methods pps sampling, two stage cluster sampling, and systematic sampling.	• How to draw random samples using different sampling techniques • PPS sampling techniques using R softwares • Ratio and regression methods using R softwares.
Module IV: Biostatistical data analysis **(20 Hours)**	Measures of Morbidity and Mortality in R, Relative Risk, Odds and Odds Ratio, Generalized Odds for Ordered 2 x k Table, Mantel–Haenszel Method, Box Plots, Estimation of Proportions and Odds Ratios, testing of hypotheses.	• Compute Morbidity and Mortality in R • Computation of Odds and odds ratio using R • Data description using Box plot • Inference procedures in R

References	***Text Books*** 1. Maria D.U., Ana F.M. and Alan T.A. (2008*): Probability and Statistics with R.* CRC Press. 2. Dalgaard, P. (2008): *Introductory Statistics with R, (Second Edition),* Springer. ***Reference Books*** 1. Purohit, S.G, Ghore, S.D and Deshmukh, S.R. (2004): *Statistics Using R*. Narosa. ***2.*** Babak Shahbaba. (2012). *Biostatistics with R: An Introduction to Statistics through Biological Data.* Springer New York.
Course Outcomes	After successful completion of this course, student will be able to: 1. Understand various built in functions in R programming for biostatistical data analysis. 2. Understand different functions in R programming for writing compute r programmes and develop computer programmes for different problems 3. Understand the usage of packages in R for drawing various diagrams and computing descriptive statistics, comparison of means, ANOVA, non-parametric tests, simple correlation and regression procedures

TEACHING LEARNING STRATEGIES

- **Practical sessions through computers, statistical computations, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Hands on training**

ASSESSMENT RUBRICS

Components	**Weightage**
End Semester Evaluation	**60**
Continuous Evaluation	
Practical Tests	32

Record	08
Total	**40**

Sample Questions to Test Outcomes:

1. Write an R program to create a matrix taking a given vector of numbers as input. Display the matrix.
2. Import a given dataset in R, and conduct its descriptive analysis.
3. Select a simple random sample of 50 numbers without replacement from the numbers 1 to 2000.
4. Generate a random sample of size 100 from a standard normal distribution.
5. Illustrate the law of large numbers using R.
6. Enter the given 2 matrices, and find their product.

POOL A: DISCIPLINE SPECIFIC ELECTIVE

Course Code & Title	**MSBST01DSE02-BIOSTATISTICAL COMPUTING USING SPSS -I (PRACTICAL)**		
Programme	**M.Sc. Biostatistics**	**Semester**	I
Course Objectives	• The main focus of the course will be on to solve biostatistical research question using SPSS • Illustrate different toolboxes in SPSS • Data definition and access and data analysis and presentation. • Apply SPSS software to develop different statistical tools • Students get awareness to chose appropriate statistical technique and interpret results using SPSS.		

Modules	Content	Module Outcome
Module I: SPSS Environment, Basic Concepts of SPSS Programming **(20 Hours)**	Introduction to SPSS- Starting SPSS, Working with data file, SPSS windows, Menus, Dialogue boxes. Preparing the Data file, Creating data file and entering data, Defining the variables, Entering data, modifying data file, import file. Variable types in SPSS and Defining variables – Creating a Codebook in SPSS. Screening and cleaning data, Manipulation of data.	▪ Understand the installation and familiar with toolboxes of SPSS. ▪ Data management and modifications of data.
Module II: Preliminary Analysis in SPSS **(25 Hours)**	Computing Variables- Recoding (Transforming) Variables: Recoding Categorical String Variables using Automatic Recode - Sorting Data - Grouping or Splitting Data. Categorical variables, continuous variables. The Explore procedure - Frequencies Procedure – Descriptive - Compare Means - Frequencies for Categorical Data, different statistical distributions	▪ Working with Data types ▪ Recoding and sorting ▪ Descriptive statistics ▪ Explore procedure, graphics in SPSS
Module III: Inferential Statistics **(25 Hours)**	Pearson Correlation, Chi-square Test of Independence – Inferential Statistics for Comparing Means: One Sample t Test, Paired Samples T Test, Independent Samples T Test, One-Way ANOVA. Two way ANOVA, Multivariate ANOVA.	▪ Compute and interpret correlation coefficients ▪ Learn how to conduct various statistical tests using SPSS ▪ Preparing ANOVA

Module IV: Non-Parametric statistics **(20 Hours)**	Independent Chi square Test, Mann-Whitney test , Wilcoxon signed rank test, Kruskal- Wallis test. Interpreting the output of tests, p-value computation.	▪ Learn how to perform non parametric tests ▪ Get p value of various tests ▪ Interpretation of test results
References	***Text Books*** 1. Hinton, P. R., Brownlow, C, Mc Murray, I. and Cozens, B. (2004): *SPSS Explained,* Routledge, Taylor and Francis group, New York. ***Reference Books*** 1 Field, A. (2011); *Discovering Statistics Using SPSS,* Sage Publications. 2 William E. Wagner. (2015). *Using IBM SPSS statistics for research methods and social science statistics*, Fifth edition, SAGE Publications, Inc.	
Course Outcomes	After successful completion of this course, student will be able to: 1. Build capacity to analyzing complex information with the help of SPSS. 2. Understand with the tool box of statistical software SPSS 3. Summarize variables using frequencies and descriptive analysis. 4. Understand to producing cross tabulation tables and testing for significant relationships with chi square test. 5. Understand the usage of assessing relationships between continuous variables through plots and correlations.	

TEACHING LEARNING STRATEGIES

- **Practical sessions through computers, statistical computations, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Hands on training**

ASSESSMENT RUBRICS

Components	Weightage
End Semester Evaluation	**60**
Continuous Evaluation	
Practical Tests	32
Record	08
Total	**40**

Sample Questions to Test Outcomes:

1. Compute t-test for difference of means on a given continuous variable, based on a categorical variable.
2. Compute descriptive statistics of a given continuous variable of a dataset.
3. Perform an appropriate ANOVA for the given data.
4. Perform appropriate non parametric test for the given data.
5. Load the given external spreadsheet data into SPSS, and obtain the pie chart, histogram of the variables.
6. Perform MANOVA for the given data set.

	SECOND SEMESTER								
Sl No	**Course Code**	**Title of Paper**	**Contact Hours/Week**			**Marks**			
			L	**T/S**	**P**	**ESE**	**CE**	**Total**	**Credits**
DISCIPLINE SPECIFIC CORE COURSES (DCE)									
2.1	MSBST02DSC05	Biostatistical Inference	4	1		60	40	100	4
2.2	MSBST02DSC06	Applied Regression Analysis	4	1		60	40	100	4
2.3	MSBST02DSC07	Statistical Epidemiology	4	1		60	40	100	4
2.4	MSBST02DSC08	Survival Analysis	4	1		60	40	100	4
DISCIPLINE SPECIFIC ELECTIVES (DSE)									
2.5	MSBST02DSExx	Elective-II-DSE (One course has to be chosen from **Pool B**)		2	6	60	40	100	3
ABILITY ENHANCEMENT COURSE(AEC)									
2.6	MSBST02AECxx	Offered to other Departments. (One course has to be chosen from **Pool C**)	2			--	100	100	2
	----------	(To be obtained from other Departments)	--	--	--	--	--	--*	2
SKILL ENHANCEMENT COURSE(SEC)									
2.7	MSBST02SECxx	Offered to other Departments. (One course has to be chosen from **Pool D)**	2			60	40	100	2
	----------	(To be obtained from other Departments)	--	--	--	--	--	--*	2
Total Credits									**23**

L=Lecture, T/S=Tutorials/Seminar, P/I=Practical/Internship, CE =Continuous Evaluation, ESE = End Semester Evaluation.

(*Note: Evaluation is determined by respective Department)

Sl No	POOL B DISCIPLINE SPECIFIC ELECTIVE (DSE)								
2.5.1	MSBST02DSE03	Biostatistical Computing Using SPSS-II (**Practical**)		2	6	60	40	100	3
2.5.2	MSBST02DSE04	Biostatistical Computing Using SAS-I (**Practical**)		2	6	60	40	100	3
Sl No	**POOL C** **ABILITY ENHANCEMENT COURSE (AEC)**								
2.6.1	MSBST02AEC01	A foundation course in LaTex for scientific documentation		2	6	60	40	100	2
2.6.2	MSBST02AEC02	Basic Statistical data analysis using EXCEL		2	6	60	40	100	2
Sl No	**POOL D** **SKILL ENHANCEMENT COURSE (SEC)**								
2.7.1	MSBST02SEC01	Exploratory Data Analysis Using SPSS		2	6	60	40	100	2
2.7.2	MSBST02SEC02	Regression Analysis Using SPSS		2	6	60	40	100	2

SECOND SEMESTER M.Sc. BIOSTATISTICS PROGRAMME

DISCIPLINE SPECIFIC CORE COURSE

Course Code & Title	**MSBST02DSC05-BIOSTATISTICAL INFERENCE**	
Programme	**M.Sc. Biostatistics**	**Semester II**
Course Objectives	• A thorough understanding of important properties of estimators of a parameter such as sufficiency, consistency, unbiasedness and efficiency. • Understanding the notion of Fisher-Neymann factorization theorem for sufficiency,minimal and complete sufficient statistics. • Derivation of the Cramer-Rao lower bound and the conditions for the existence of MVB estimator, apply the concept of Rao-Blackwell and Lehmann-Scheffe theorems to obtain UMVUE of a parameter. • To introduce the concept of testing of hypothesis, critical region, significance level, power of the test and p-value. • To introduce the concept of uniformly most powerful test for testing simple hypothesis against simple alternative and obtain sequential probability ratio test for testing the hypothesis. • To introduce the notion of likelihood ratio test and confidence interval estimation and its applications.	

Modules	Content	Module Outcome
Module I: Properties of estimators. **(15 hours)**	Unbiasedness, consistency, consistent asymptotically normal (CAN) estimators, efficiency, sufficiency, invariance property of consistent estimators, Fisher-Neymann factorization theorem for sufficiency (proof for discrete distributions only), joint sufficient statistics, minimal and complete sufficient statistic.	• Derive the important properties of estimators • Determine the sufficient statistic • Determine consistent estimators and consistent and asymptotically normally distributed estimators. • Identify efficient estimators.

Module II: Minimum Variance Unbiased Estimation. (**15 hours**)	Minimum variance unbiased estimator (MVUE), Likelihood and score functions. Fisher information, Cramer-Rao inequality and its applications, Cramer–Rao Lower Bound (CRLB), Minimum variance bound unbiased estimator (MVB). Rao-Blackwell and Lehmann-Scheffe theorems. Method of moments and method of MLE and their properties.	• State and prove Cramer-Rao inequality. • Examples of Minimum variance bound estimator. • Apply Rao-Blackwell and Lehmann-Scheffe theorems to find UMVUE. • Understand the concept of MLE and its properties. • Examples of moment estimators.
Module III: Tests of Hypotheses and Most Powerful Tests. (**15 hours**)	Null and alternative hypotheses, simple and composite hypotheses, two types of errors in testing of hypothesis, p-value, level of significance and size of test, power function, Neymann-Pearson lemma, most powerful and uniformly most powerful tests. Sequential Probability Ratio Test (SPRT)(Concept only).	• Identify null, alternative, simple and composite hypothesis • Find critical region, size and power of the test . • Apply Neymann-Pearson lemma to find most powerful test. • Derive SPRT for test the parameters of normal distribution, exponential and Poisson distributions, binomial distribution.
Module IV: Parametric test, Likelihood ratio test and SPRT (**15 hours**)	Different parametric test (Z ,t and F tests), large sample tests. Likelihood ratio test, monotone likelihood ratio property. Confidence interval estimation of mean and variance, difference of mean and population proportion, and difference of population proportion.	• Apply different parametric test for various real life applications. • Understand likelihood ratio property and its applications. • Compute the confidence interval for mean, population proportional and variance.

References	***Text Books*** **1.** Hogg, R. V., McKean, J. W., & Craig, A. T. (2013). *Introduction to Mathematical Statistics*. Pearson Education India. **2.** Vijay K. Rohatgi, A. K. Md. Ehsanes Saleh (2015). *An Introduction to Probability and Statistics,3rd Edition,* John Wiley and Sons, NewYork**.** 3. Mood, A. M., & Graybill, F. A. (6). Boes, DC (1974). Introduction to the Theory of Statistics**.**Third edition. McGraw Hill. ***Reference books*** 1. Casella,G.andBerger,R.L.(2002).*StatisticalInference*,*SecondEdition,*Dux bury,Australia. 2. Lehman, E. L. (1986):*Testing of Statistical Hypotheses*. John Wiley, New York. 3. Lehmann, E. L(1983). *Theory of Point Estimation*, John Wiley and Sons, New York.
Course Outcomes	After successful completion of this course, student will be able to: 1. Understand the concepts of Sufficiency and Completeness 2. Understand the concepts of Minimum Variance Unbiased Estimation. 3. Understand various estimation methods and applications in real life problems 4. Apply various parametric and sequential testing procedures to deal with real life problems. 5. Understand Most Powerful Tests for testing simple null hypothesis and developing MP tests for different problems.

TEACHING LEARNING STRATEGIES

- **Lecturing, Hands on Training, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	Weightage
End Semester Evaluation(ESE)	**60**
Continuous Evaluation	
Tests	16
Assignment	08
Seminar	16
Total	**40**

Sample Questions to Test Outcomes:

1. State factorization theorem for sufficiency.
2. Let $X_1, X_2, ..., X_n$ be a random sample of n observations from Gamma distribution with shape parameter θ. Find sufficient statistic for θ.
3. State and prove Basu's theorem.
4. Let $X_1, X_2, ..., X_n$ be random sample from a Poisson population with parameter λ. Show that the $\overline{X}^2$ is a UMVUE for λ^2.
5. Let $X_1, X_2, ..., X_n$ be a random sample of size n observations from beta first kind distribution with parameter α and β. Find the estimators of α and β by method of MLE.
6. Let p be the probability that a coin will fall head in a single toss in order to test$H_0 : p = 1/2$against $H_1 : p = 3/4$. The coin is tossed 5 times and H_0 is rejected if more than 3 heads are obtained. Find the probability of type-I error and power of the test.
7. Define a most powerful test and explain the utility of Neyman-Pearson lemma.
8. Given the nine sample values 4.5, 6.5, 3.8, 4.2, 7.7, 8.5, 9.4, 5.3, 3.9 from a normal distribution with mean μ and variance 4. Find the best critical region for testing$H_0 : \mu = 4$ versus $H_1 : \mu = 5$ of size 0.05. Also calculate the power of the test.
9. Obtain OC function for testing $H_0 : p = p_0$ versus $H_1 : p = p_1$ using SPRT with strength (α, β) based on observations from b(n, p).
10. Show that the likelihood ratio test for testing the equality of variances of two normal distributions is the usual F-test.

DISCIPLINE SPECIFIC CORE COURSE

Course Code & Title	**MSBST02DSC06-APPLIED REGRESSION ANALYSIS**	
Programme	**M.Sc. Biostatistics**	**Semester II**
Course Objectives	• Describe simple linear regression models and its properties • Apply principle of least square method to estimate the parameters in simple linear regression models. • Describe multiple linear regression models and its properties and model adequacy. • Identify multicollinearity problem, its consequences, discuss the problem of estimation of parameters when multicollinearity occurs and also Identify auto correlation and its consequences. • Explain polynomial regression in one and several variables. • Understand the notion of nonlinear regression. • Explain generalized linear model. • Describe the logistic regression, Poisson regression.	

Modules	Content	Module Outcome
Module I: Linear Regression Models **(15 Hours)**	The simple linear regression models, least square estimation, statistical assumptions and properties of estimators, standard error of estimates, tests of significance and confidence intervals for the parameters, error and residual plots.	• Explain simple linear regression model • Describe least square estimators. • Articulate to inference regarding regression parameters. • Explain ANOVA.
Module II: Regression Diagnostics **(15 Hours)**	Multiple regression models, OLS and ML estimators, testing and prediction. Multicollinearity, heteroscedasticity, autocorrelation: their nature, consequences,	• Explain multiple linear regression models. • Explain multicollinearity. • Discuss detection and remedial measures of

	detection, remedial measures and estimation in the presence of them.	multicollinearity. • Explain heteroscedasticity and remedial measures of heteroscedasticity.
Module III: Polynomial and nonlinear regression **(15 Hours)**	Polynomial regression in one and several variables. Linearization transforms, Diagnostic checks and correction. Nonparametric regression and concept of spline smoothing.	• Explain polynomial regression. • Discuss non-parametric regression. • Explain concept of spline smoothing.
Module IV: Non-Linear Regression **(15 Hours)**	Generalized linear models. Logistic regression, Poisson Regression. Estimation, model adequacy of GLM and diagonostic tests.	• Discuss generalised linear model. • Explain logistic regression. • Introduce the concept of Poisson regression.
References	***Text Books*** 1. McCullagh, P. (2019). Generalized Linear Models. United Kingdom: CRC Press. 2. Hosmer, D.W. and Lemeshow, S. (1989): *Applied Logistic Regression*, John Wiley. 3. Montgomery, D. C, Peek, E.A. and Vining, G.G. (2006): *Introduction to Linear Regression Analysis*, John Wiley. ***Reference books*** 1. Seber, G. A. F. and Lee, A. J. (2003): *Linear Regression Analysis*, Wiley 2. Draper ,N.R. and Smith,H.(1998): *Applied Regression Analysis, 3rd Ed.* John Wiley. 3. Goon,Gupta,DasGupta(2001):*An Outline Series in Statistics Vol1I,* World Press. 4. Chatterjee, S., & Hadi, A. S. (2013). *Regression Analysis by Example*. John Wiley & Sons.	

Course Outcomes	After successful completion of this course, student will be able to: 1. Understand various regression models including logistic regression models. 2. Understand consequences of multicollinearity, heteroscedasticity, autocorrelation, their detection and remedial measures. 3. Apply statistical techniques to model relationships between variables and make predictions. 4. Acquire knowledge of various advanced regression models with applications in biostatistics.

TEACHING LEARNING STRATEGIES

- **Lecturing, Hands on Training, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	Weightage
End Semester Evaluation(ESE)	**60**
Continuous Evaluation	
Tests	16
Assignment	08
Seminar	16
Total	**40**

Sample Questions to Test Outcomes:

1. Derive the OLS estimators of intercept and slop coefficients of a simple linear regression model.
2. Explain measure of goodness of fit in regression analysis. obtain its relationship with correlation coefficient.
3. Define multicollinearity.
4. How to detect heteroscedasticity using Spearman's rank correlation test.
5. Discuss Durbin-Watson test for autocorrelation.
6. Explain the difference between R^2 and adjusted R^2 in multiple regression. Mention its uses

7. What is logistic regression model?
8. Explain link function and linear predictor?
9. Explain the parameter estimation and inference of GLM.
10. Explain orthogonal polynomial regression.

DISCIPLINE SPECIFIC CORE COURSE

Course Code & Title	**MSBST02DSC07-STATISTICAL EPIDEMIOLOGY**	
Programme	**M.Sc. Biostatistics**	**Semester II**
Course Objectives	• To understand basic concepts of epidemiology and explain different study designs in epidemiology. • Identifying achievements in epidemiology and measuring health and disease. • To aquire the knowledge of cohort study designs, case control study designs. • Develop the knowledge of various matched case control studies and cross over study designs. • To know about statistical concepts and inference. • Understand relationship between variables. • State Mendal's law and estimation of allele frequency, estimation of allele frequencies, Hardy-Weinberg law. • Introduce the concept of detection and estimation of linkage, inheritance of quantitative traits, stochastic models of carcinogenesis.	

Modules	Content	Module Outcome
Module I: Basic concepts of Epidemiology **(15 Hours)**	Basic concepts of epidemiology: definition and scope of epidemiology, achievements in epidemiology, measuring health and disease, definition of health and disease, measures of disease frequency, comparing disease occurrence.	• Explain the basic concepts of epidemiology. • Measuring health and disease • Comparing disease occurrence.

Module II: Study designs in epidemiology **(15 Hours)**	Types of study: observations and experiments, observational epidemiology, cohort study designs, case control study designs, randomized, field trails, control trials, potentials errors in epidemiological studies, ethical issues	• Explain different study designs in epidemiology. • Understand various cohort and case control studies in biostatistics.
Module III: Exploratory analysis **(15 Hours)**	Distribution and summary measures: distribution, measures of central tendency, measures of variability, normal and log normal distributions, estimation, testing, CI, ANOVA, relationship between two variables: chi-square test, correlation, regression, logistic regression.	• To understand statistical concepts. • To know about relationship between two variables.
Module IV: Clinical epidemiology **(15 Hours)**	Concept of cause, establishing the cause of disease, scope of prevention, levels of prevention, Introduction to clinical epidemiology.	• Understand the concept of Concept of cause. • Identifying the scope of prevention. • Understand the basic concepts of clinical epidemiology.
References	***Text Books*** 1. Beaglehole, R., Bonita, R. and Kjellstorm, T. (1993). *Basic Epidemiology*. World Health Organization, Geneva. 2. Newman,S.C.(2001).*Biostatistical Methods in Epidemiology*. John Wiley & Sons, New York. 3. Virasakdi, C. (2010). *Analysis of Epidemiological Data Using R and Epicalc*. Epidemiological Unit, Songla University,Thailand ***Reference books*** 1. Rothan, K. J., Greenland, S. and Lash, T. L. (2008). *Modern Epidemiology, 3rd Edition.* Wokers Kluver.	

	2. Clayton, D.and Hills, M.(1993). *Statistical Methods in Epidemiology.* Oxford University Press.
Course Outcomes	After successful completion of this course, student will be able to: 1. Understand Basics of epidemiology. 2. Understandtypes of study used in epidemiology 3. Understand the concept of clinical epidemiology. 4. Identify the scope of cause and prevention of disease.

TEACHING LEARNING STRATEGIES

- **Lecturing, Hands on Training, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	**Weightage**
End Semester Evaluation(ESE)	**60**
Continuous Evaluation	
Tests	16
Assignment	08
Seminar	16
Total	**40**

Sample Questions to Test Outcomes:

1. Explain Basic concepts of epidemiology along with its scope.
2. Narrate the notion of measures of disease frequency.
3. Differentiate between health and disease.
4. Elaborate various types of study in epidemiology.
5. Differentiate between cohort study and case control study.
6. Describe the potential error in epidemiological studies.
7. Define ANOVA and mention its application.

8. Why we need chi-square test? Elaborate the concept.

9. Explain the concept of clinical epidemiology.

10. Differentiate between scope of prevention and levels of prevention in epidemiological study.

DISCIPLINE SPECIFIC CORE COURSE

Course Code & Title	**MSBST02DSC08-SURVIVAL ANALYSIS**	
Programme	**M.Sc. Biostatistics**	**Semester II**
Course Objectives	• Understand the basic notion of survival analysis • Understanding applications of survival function, hazard function, mean residual life functions and other aging concepts . • Understand and examine the properties of standard lifetime distributions. • Develop the concept of nonparametric tests like KS test, sign test, Mann-Whitney and Wilcoxon U tests. • Introduce the concepts of censoring and truncation and its various classifications. • Introduce the notion of Estimating survival rates using large scale data like DHS, NFHS, DLHS, etc. Comparing survival curves. • Introduce the notion of Kaplan-Meier estimation technique, life tables, Mantel-Haenszel test.	

Modules	Content	Module Outcome
Module I Basics of survival analysis **(15 Hours)**	Basics of survival analysis- discrete and continuous time models, survival function, hazard rate function, probability density function, mean residual life time. Aging classes-IFR, IFRA and their duals, Bathtub failure rate.	• Understand the basic concepts and ideas of survival analysis. • Understand the basic concepts of ageing classes.
Module II Life	Life distributions-exponential, Weibull, lognormal and gamma distributions,	• Examine the properties and methods for standard

distributions **(15 Hours)**	characterizations. Nonparametric tests- Kolmogorov-Smirnov test, sign test and signed-rank test, Mann-Whitney, Wilcoxon U test, chi- square test for goodness of fit, test for independence of attributes.	survival time distributions. • Understand the basic concepts of nonparametric tests.
Module III Censoring and Truncation **(15 Hours)**	Concepts of censoring Mechanism -Type-I, Type-II and random censoring, Progressive censoring, Truncation, Methods for truncated and interval censored data. Likelihood construction and estimation of Censored and Truncated Data. Estimating survival rates using large scale data like DHS, NFHS, DLHS, etc. Comparing survival curves.	• Estimate survival functions using parametric and non-parametric methods. • Understand the basic concepts of censoring and truncations.
Module IV Estimation of Survival function **(15 Hours)**	Kaplan-Meier estimation technique, life tables, Mantel-Haenszel test. Interval estimation of survival probabilities. Introduction to survival regression. Cox proposional hazard model.	• Apply and interpret regression models for survival data . • Understand the concept of Cox-Proportional hazard model.
References	***Text Books*** 1. Lawless, J.F. (2003): *Statistical Methods for Lifetime (Second Edition)* ,John Wiley & Sons Inc., New Jersey. 2. Kalbfleisch, J. D. and Prentice, R.L. (1980): *The Statistical Analysis of Failure Time Data,* John Wiley & Sons Inc. New Jersey. 3. Moore, D.F. (2016): *Applied Survival Analysis Using R*, Springer. ***Reference Books*** 1. Klein J.P. and Moeschberger M.L. (2003) *Survival Analysis - Techniques for Censored and Truncated Data*, Second Edition,	

	Springer-Verlag, New York. 2. Miller, R. G.(1981): *Survival Analysis*, John Wiley & Sons Inc. 3. Bain, L. G.(1978*): Statistical Analysis of Reliability and Life testing Models*, Marcel Decker. 4. Cox, D.R and Oakes, D.(1984): *Analysis of Survival Data*. Chapman and Hall. 5. Fraser, D.A.S.(1957): *Non-parametric Methods in Statistics*, Wiley, New York.
Course Outcomes	After successful completion of this course, student will be able to: 1. Understand various lifetime probability distributions and their structural properties 2. Understand different methods for the estimation of survival function. 3. Conduct analysis of life time data. 4. Apply statistical techniques to model lifetime data and make predictions.

TEACHING LEARNING STRATEGIES

- **Lecturing, Hands on Training, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	Weightage
End Semester Evaluation(ESE)	**60**
Continuous Evaluation	
Tests	16
Assignment	08
Seminar	16
Total	**40**

Sample Questions to Test Outcomes:

1. Define survival function and hazard function.
2. Explain aging classes and their dual.
3. What do you mean by bathtub failure rate?
4. Explain any three life distributions.
5. What is goodness of fit test? Explain any two goodness of fit test.
6. Define censoring. Elaborate various types of censoring.
7. Explain truncation. Differentiate between censoring and truncation.
8. Explain Kaplan-Meier estimation technique.

POOL B: DISCIPLINE SPECIFIC ELECTIVE COURSE (Practical)

Course Code & Title	**MSBST02DSE03- BIOSTATISTICAL COMPUTING USING SPSS – II (Practical)**		
Programme	**M.Sc. Biostatistics**	**Semester**	II
Course Objectives	• To introduce some advanced statistical computing techniques in applied statistics to extract information and visualization thereby enabling them to perform data analysis effectively and efficiently in SPSS. • Illustrate different statistical techniques based on all the elective course in second semester.		

Modules	**Content**	**Module Outcome**
	Statistical Computing II is a practical course. The practical is based on all the elective courses in the second semester.	• Describe different statistical technique to solve problems coming under all the elective courses in second semester.
Course Outcomes	After successful completion of this course, student will be able to: 1. Equipped with different theoretical methods in biostatistics to achieve the objectives. 2. Enhanced with the basic concepts of biostatistical theories besides developing their ability to handle real world problems with large scale data.	

TEACHING LEARNING STRATEGIES

- **Practical sessions through computers, statistical computations, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Hands on training**

ASSESSMENT RUBRICS

Components	Weightage
End Semester Evaluation	**60**
Continuous Evaluation	
Practical Tests	32
Record	08
Total	**40**

POOL B: DISCIPLINE SPECIFIC ELECTIVE COURSE (Practical)

Course Code & Title	**MSBST02DSE04- BIOSTATISTICAL COMPUTING USING SAS – I (Practical)**		
Programme	**M.Sc. Biostatistics**	**Semester**	II
Course Objectives	• To introduce some advanced statistical computing techniques in applied statistics to extract information and visualization thereby enabling them to perform data analysis effectively and efficiently in SAS programming. • Illustrate different statistical techniques based on all the elective course in second semester.		

Modules	Content	Module Outcome
	Statistical Computing II is a practical course. The practical is based on all the elective courses in the second semester.	• Describe different statistical technique to solve problems coming under all the elective courses in second semester.

Course Outcomes	After successful completion of this course, student will be able to: 1.Equipped with different theoretical methods in biostatistics to achieve the objectives. 2. Enhanced with the basic concepts of biostatistical theories besides developing their ability to handle real world problems with large scale data.

TEACHING LEARNING STRATEGIES

- **Practical sessions through computers, statistical computations, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Hands on training**

ASSESSMENT RUBRICS

Components	Weightage
End Semester Evaluation	**60**
Continuous Evaluation	
Practical Tests	32
Record	08
Total	**40**

POOL C:- List of Ability Enhancement Courses (Offered to other Departments)

ABILITY ENHANCEMENT COURSE (AEC)

Course code**: MSBST02AEC01**

Name of the Course: **A foundation course in *LaTex* for scientific documentation.**

Department Offering the Course: **Department of Statistical Sciences**

Mode of Delivery: **Hybrid**

Credit Distribution, Eligibility and Pre-Requisites of the Course

Credits	Contact hours per week			Eligibility Criteria	Pre-requisite (if any)
	Lecture	Tutorial	Practical/ Internship		
2	1	1	2		

Skill Outcomes: To introduce students with a software that is being widely used for scientific typesetting, To make students know importance of this software for publishing research articles, letters, project reports, books and beamer/slide presentation and thereby help them to be comfortable with the software .

Course Contents:

Module 1: Installation of Kile and MikeTeX. Class and packages. Latex programming and commands, sample packages. Error messages, Some sample errors, list of LaTeX error messages.

Module 2: Fonts, symbols, Indenting, paragraphs, line spacing, word spacing, titles and subtitles. Document class, page style, parts of the documents, table of contents. Command names and arguments, environments, declarations. Theorem like declarations, comments within text.

Module 3: Mathematical environments, math mode, mathematical symbols. Graphic package, multivalued functions, drawing matrices. Tables, tables with captions. References to figures and tables in text.

Module 4: Picture environments. Extended pictures, other drawing packages. Preparing book, project report in LaTeX, LaTeX Beamer for Technical Presentations.

Suggested Readings:

1. Kottwitz, S. (2021). *LaTeX Beginner's Guide: Create Visually Appealing Texts, Articles, and Books for Business and Science Using LaTeX*. United Kingdom: Packt Publishing**.**
2. Lamport **(**1994**).** *Latex: A Document Preparation System, 2/E*. India: Pearson Education**.**
3. Kopka, H., Daly, P. W. (2003). *Guide to LaTeX*. United Kingdom: Pearson Education.

TEACHING LEARNING STRATEGIES

- **Hands on training, Lecturing, Visualization, Team Learning.**

MODE OF TRANSACTION

- **Lab session, Lecture, Seminar, Discussion, Questioning and Answering**

Assessment Rubrics: Evaluation by Department

Sample Questions to Test Outcomes:

1. How do you install Kile and MiKTeX for LaTeX editing and compilation, and what are their roles in the LaTeX ecosystem?
2. Explain the concept of classes and packages in LaTeX, and how they contribute to document formatting and customization.
3. Discuss LaTeX programming and commands, including sample packages, and common error messages encountered during compilation.
4. What are some sample errors in LaTeX, and how can they be addressed? Provide a list of common LaTeX error messages.
5. Describe the role of fonts and symbols in LaTeX, and how they can be customized for document appearance.
6. How do you manage indentation, paragraphs, line spacing, and word spacing in LaTeX documents?
7. Discuss titles, subtitles, document classes, page styles, and the creation of table of contents in LaTeX.
8. Explain command names and arguments, environments, and declarations in LaTeX, including theorem-like declarations and comments within text.
9. What are mathematical environments and math mode in LaTeX, and how can mathematical symbols be utilized?
10. Describe the usage of the graphic package in LaTeX, including drawing matrices, creating tables with captions, and referencing figures and tables within text.

POOL C: ABILITY ENHANCEMENT COURSE –II

Course code: **MSBST02AEC02**

Name of the Course: **Basic Statistical data analysis using EXCEL**.

Department Offering the Course: **Department of Statistical Sciences**

Mode of Delivery: **Hybrid**

Credit Distribution, Eligibility and Pre-Requisites of the Course

Credits	Contact hours per week	Eligibility	Pre-requisite

	Lecture	Tutorial	Practical/ Internship	Criteria	(if any)
2	1	1	2		Basic knowledge of statistics

Skill Outcomes:

To build a strong understanding on the Basics of Microsoft Excel, To understand data crunching, Understand core analytic techniques that work in Excel, Data visualization in Excel.

Course Contents:

Module 1: Excel Introduction, Basic Navigation Tab, Concept of Cell and Cell address , row Column concept, Basic mathematical and statistical functions in Excel.

Module 2: Min, Max, Trim, Lower, Upper, Proper, Left, Right, Mid Exact, Randbetween, Rand, Len (Length of character) Paste special, SQRT, If function with Example of IF, More function like And, OR with their example, Conditional Formatting basic and advance level with OR, AND, Nested IF function, Index, Offset, Match.

Module 3: Graphics in excel-pie chart, bar chart, multiple bar diagram, sub-divided bar diagram, histogram, line chart, scatter diagram, box plot.

Module 4: Median, Mode, Standard Deviation (SD), Correlation, Large, Small, Pivot Table, Pivot Charts, Slicing, Sparkling.

Suggested Readings:

1. Linoff, Gordon S (2015). *Data analysis using SQL and Excel*. John Wiley & Sons.
2. Guerrero, Hector, Rauscher Guerrero, and Rauscher (2019). *Excel data analysis*. Springer International Publishing.

TEACHING LEARNING STRATEGIES

- **Hands on training, Lecturing, Visualization, Team Learning.**

MODE OF TRANSACTION

- **Lab session, Lecture, Seminar, Discussion, Questioning and Answering**

Assessment Rubrics: Evaluation by Department

Sample Questions to Test Outcomes:

1. What is the purpose of Excel and how is it commonly used in data management and analysis?
2. Describe the basic navigation tab in Excel and explain the concept of cells, cell addresses, rows, and columns.
3. Discuss basic mathematical and statistical functions in Excel, including examples of their usage.
4. Explain the functions MIN, MAX, TRIM, LOWER, UPPER, PROPER, LEFT, RIGHT, MID, EXACT, RANDBETWEEN, RAND, and LEN, and how they are used in Excel.
5. How does the Paste Special function work in Excel, and what are its common applications?
6. Walk through the usage of the SQRT function and demonstrate its application in Excel.
7. Provide examples of the IF function in Excel and discuss its importance in conditional logic.
8. Explain the functions AND and OR in Excel with examples, and discuss their usage in conjunction with conditional formatting.
9. What are nested IF functions in Excel, and how are they implemented? Provide examples.
10. Discuss advanced functions such as INDEX, OFFSET, MATCH, and their applications in Excel data analysis and manipulation.
11.

POOL D:- List of Skill Enhancement Courses (Offered to other Departments)

SKILL ENHANCEMENT COURSE (SEC)-I

Course code: **MSBST02SEC01**

Name of the Course: **Exploratory Data Analysis Using SPSS**

Department Offering the Course: **Department of Statistical Sciences**

Mode of Delivery: **Hybrid**

Credit Distribution, Eligibility and Pre-Requisites of the Course

Credits	Contact hours per week			Eligibility Criteria	Pre-requisite (if any)
	Lecture	Tutorial	Practical/ Internship		
2	1	1	2		Knowledge of basic statistics

Skill Outcomes: To introduce students with a software that is being widely used for Statistical data analysis.To make students know importance of this software for data analysis in research articles and thereby help them to be comfortable with the software.

Course Contents:

Module 1: What is SPSS?, Opening SPSS, Layout of SPSS,Structure of SPSS Exiting SPSS, inputting data, An overview of SPSS.

Module 2: Exploring data distributions using descriptive statistics, Creating frequency distributions and summary tables, Generating basic visualizations (e.g., histograms, box plots) in SPSS.

Module 3: Understanding correlation and covariance, Performing correlation analysis in SPSS, Introduction to linear regression and its application in SPSS,Understanding hypothesis testing principles, Conducting hypothesis tests in SPSS, Interpreting SPSS output for hypothesis testing.

Module 4: Generating various types of charts and graphs in SPSS, Customizing visualizations for clarity and impact, exploring the SPSS Chart Builder tool.

Suggested Readings:

1. Landau, S., & Everitt, B. S. (2003). *A handbook of statistical analyses using SPSS*. Chapman and Hall/CRC.
2. Tukey, J. W. (1977). *Exploratory data analysis* (Vol. 2).
3. Aldrich, J. O. (2018). *Using IBM SPSS Statistics: An interactive hands-on approach*. Sage Publications.

TEACHING LEARNING STRATEGIES

- **Hands on training, Lecturing, Visualization, Team Learning.**

MODE OF TRANSACTION

- **Lab session, Lecture, Seminar, Discussion, Questioning and Answering**

Assessment Rubrics: Evaluation by Department

Sample Questions to Test Outcomes:

1. What is the purpose of SPSS and how does it contribute to statistical analysis?
2. Describe the layout and structure of SPSS interface, including its main components and functions.
3. How do you open and exit SPSS, and what considerations should be taken into account?
4. Explain the process of inputting data into SPSS and discuss common formats accepted.
5. What are descriptive statistics, and how can they be used to explore data distributions in SPSS?
6. Describe the steps involved in creating frequency distributions and summary tables using SPSS.
7. How can basic visualizations such as histograms and box plots be generated in SPSS, and what insights can they provide?
8. What are correlation and covariance, and how are they calculated and interpreted in SPSS?
9. Discuss the principles of hypothesis testing and how it is conducted in SPSS.
10. Explain how to customize visualizations for clarity and impact in SPSS, and explore the functionalities of the SPSS Chart Builder tool.

POOL D: SKILL ENHANCEMENT COURSE (SEC)-II

Course code: **MSBST02SEC02**

Name of the Course: **Regression Analysis Using SPSS**

Department Offering the Course: **Department of Statistical Sciences**

Mode of Delivery: **Hybrid**

Credit Distribution, Eligibility and Pre-Requisites of the Course

Credits	Contact hours per week			Eligibility Criteria	Pre-requisite (if any)
	Lecture	Tutorial	Practical/ Internship		

2	1	1	2		Basic knowledge in regression.

Skill Outcomes: To introduce students with software that is being widely used for regression analysis.To make students know importance of this software for data analysis in research articles and thereby help them to be comfortable with the software.

Course Contents:

Module 1: What is SPSS?Opening SPSS, Layout of SPSS, Structure of SPSS Exiting SPSS, inputting data, an overview of SPSS.

Module 2: Introduction to SPSS regression procedures, Importing and preparing data for regression analysis, Overview of the SPSS regression dialog box, Conducting simple linear regression in SPSS, Interpreting regression output in SPSS, Assumptions and diagnostics in simple linear regression

Module 3: Understanding and formulating multiple linear regression models, Conducting multiple linear regression analysis in SPSS, Interpreting output and assessing model fit

Module 4: Assumption testing for multiple regression, Dealing with multicollinearity in SPSS. Interpreting diagnostic plots and statistic, Organizing and documenting regression analysis in SPSS, Creating comprehensive reports with SPSS output, Best practices for presenting regression findings to diverse audiences

Suggested Readings:

1. Landau, S., & Everitt, B. S. (2003). *A handbook of statistical analyses using SPSS*. Chapman and Hall/CRC.
2. Aldrich, J. O. (2018). *Using IBM SPSS Statistics: An interactive hands-on approach*. Sage Publications.
3. Chatterjee, S., & Hadi, A. S. (2013). *Regression analysis by example*. John Wiley & Sons.

TEACHING LEARNING STRATEGIES

- **Hands on training, Lecturing, Visualization, Team Learning.**

MODE OF TRANSACTION

- **Lab session, Lecture, Seminar, Discussion, Questioning and Answering**

Assessment Rubrics: Evaluation by Department

Sample Questions to Test Outcomes:

1. What are the key components of SPSS and how do they contribute to statistical analysis?
2. How do you input data into SPSS and what are the common formats accepted?
3. Describe the process of exiting SPSS and any considerations to keep in mind.
4. What are the steps involved in importing and preparing data for regression analysis in SPSS?
5. Explain the regression dialog box in SPSS and its various options for analysis.
6. Walk through the process of conducting simple linear regression in SPSS, including data input and interpretation of results.
7. What are the assumptions and diagnostics involved in simple linear regression analysis, and how can they be assessed in SPSS?

	THIRD SEMESTER								
Sl No	**Course Code**	**Title of Paper**	**Contact Hours/Week**			**Marks**			
			L	**T/S**	**P**	**ESE**	**CE**	**Total**	**Credits**
DISCIPLINE SPECIFIC CORE COURSES (DSC)									
3.1	MSBST03DSC09	Design of Experiments and Quality Control	4	1		60	40	100	4
3.2	MSBST03DSC10	Analysis of Clinical Trials	4	1		60	40	100	4
DISCIPLINE SPECIFIC ELECTIVES (DSE)									
3.3	MSBST03DSExx	(One course has to be chosen from **Pool E**)	3	2		60	40	100	3
3.4	MSBST03DSExx	Elective-II (DSE) (Any two courses have to be chosen from **Pool F**)	3	2		60	40	100	3
INTERDISCIPLINARY ELECTIVE COURSE (IDC)*									
3.5	MSBST03IDCxx	IDC Elective (One course has to be chosen from **Pool G**) (Offered to other Departments)	2	2	4	60	40	100	4
	----------	To be obtained from other departments	--	--	--	--	--	--	--
*Note: Compulsory course									
VAC/MOOC COURSE**									
3.6	MOOC	Offered by external agencies				60	40	100	At least 2
	MSBST03VACxx	VAC (To be decided by Department as per requirements)	2	2	2	60	40	100	2

**Note: The course shall be considered as additional credits and shall not be considered for computation of CGPA.

INTERSHIP/FIELD VISIT/ MINOR PROJECT/ INDUSTRIAL VISIT									
3.7	MSBST03DSC11	Any one of: internship, field visit, minor project or industrial visit.	2	-	-	60	40	100	2
Total			**19**						**23**

L=Lecture, T/S=Tutorials/Seminar, P/I=Practical/Internship, CE =Continuous Evaluation, ESE = End Semester Evaluation

Sl No	POOL E: DISCIPLINE SPECIFIC ELECTIVES (DSE) (Practical)								
3.3.1	MSBST03DSE05	Biostatistical Computing Using R -II		2	6	60	40	100	3
3.3.2	MSBST03DSE06	Biostatistical Computing Using SAS -II		2	6	60	40	100	3
	POOL F : DISCIPLINE SPECIFIC ELECTIVES (DSE)								
3.4.1	MSBST03DSE07	Stochastic Process and Time Series Analysis	4	2	-	60	40	100	3
3.4.2	MSBST03DSE08	Applied Multivariate Analysis	4	2	-	60	40	100	3
3.4.3	MSBST03DSE09	Machine Learning Techniques for Biostatistics	4	2	-	60	40	100	3
3.4.4	MSBST03DSE10	Categorical Data Analysis	4	2	-	60	40	100	3
3.4.5	MSBST03DSE11	Operations Research	4	1	-	60	40	100	3
3.4.6	MSBST03DSE12	Data Visualisation and analysis using Python	2	1	4	60	40	100	3

Sl No	POOL G: INTERDISCIPLINARY ELECTIVES (for other Departments)(IDC)								
3.5.1	MSBST03IDC01	Statistical Data Analysis using SPSS		2	6	60	40	100	3
3.5.2	MSBST03IDC02	Statistical Data Analysis using R		2	6	60	40	100	3

DISCIPLINE SPECIFIC CORE COURSE

Course Code & Title	**MSBST03DSC09- DESIGN OF EXPERIMENTS AND QUALITY CONTROL**	
Programme	**M.Sc. Biostatistics**	**Semester III**
Course Objectives	• This course provides the students the ability to understand the design and conduct experiments, as well as to analyze and interpret data. After successful completion of this course, student will be able to: • Apply ANOVA for one way and two-way classification, fixed effect models with equal and unequal number of observations percell, Random and Mixed effect models. • Design and analyse incomplete block designs, understand the concepts of orthogonality, connectedness and balance. • Identify the effects of different factors and their interactions and analyse factorial experiments. • Understand basics of statistical quality control and various control charts.	

Modules	Content	Module Outcome
Module I Basic concepts of design of experiments **(15 Hours)**	Basic terminology and definitions, Randomization, Replication and local control, Fixed, mixed and random effect models, Gauss Markov theorem, fundamental principles of design of experiments, Analysis of variance-one way and two ways.	• Understand the basic concepts of design of experiments. • Differentiate between fixed, mixed, and random effect models. • Explain the Gauss-Markov theorem. • Perform analysis of variance for one-way and two-way experimental designs.

Module II Randomized designs **(15 Hours)**	Completely randomized design analysis, randomized block design-analysis, Latin square design-analysis, Graeco-Latin square designs, analysis of missing data. Analysis of covariance for RBD.	• Design and analyze CRD, RBD, LSD. • Design and analyze experiments using Graeco-Latin square designs. • Understand methods for analyzing and handling missing data.
Module III Incomplete block and factorial designs **(15 Hours)**	Incomplete block design-Balanced incomplete block design, construction of BIBD design, intra block analysis of BIBD, Factorialexperiments-2^n, concept of confounding.	• Design and analyze experiments using incomplete block designs. • Perform intra block analyses for balanced incomplete block designs. • Design and analyze factorial experiments with 2^n levels. • Identify and understand the presence of confounding in factorial experiments.
Module IV Introduction to statistical quality control **(15 Hours)**	Quality and quality assurance, Methods of quality assurance, statistical quality control Control charts, Basic ideas, designing of control charts for the number of non-conformities and fraction non-conformities, mean charts, Median charts, R-charts, and S-charts, ARL, Economic design of Shewarts control charts.	• Define and differentiate between quality and quality assurance.. • Design and interpret mean charts, median charts, range (R)-charts, and standard deviation (S)-charts for process monitoring. • Understand the economic implications of implementing Shewhart's control charts.
References	***Text Books*** 1. Das, M.N. and Giri, N.S. (2002): *Design and Analysis of*	

	Experiments, 2nd Edition, New Age International (P) Ltd., New Delhi. 2. Joshi,D.D.(1987):*Linear Estimation and Design of Experiments*. Wiley Eastern Ltd., New Delhi. 3. Montgomery, D.C. (2001*): Design and Analysis of Experiments. 5th Edition,* John Wiley & Sons- New York. 4. Montgomery, R.C. (1985). Introduction to Statistical Quality Control, Fourth edition, Wiley. ***Reference books*** 1. Gupta, S. C and Kapoor, V. K.(2010). *Fundamentals of Applied Statistics.* Sulthan Chand & Co,NewDelhi. 2. Amitava Mitra - Fundamentals of Quality Control and Improvement – Pearson Education Asia 2001. 3. The ISO 9000 book, Second Edition, Rabbit, J T and Bergle, PA Quality resources.
Course Outcomes	After successful completion of this course, student will be able to: 1. Demonstrate a mastery of fundamental principles guiding the design of experiments. 2. Able to design and analyze experiments using Latin square and Graeco-Latin square designs. 3. Demonstrate expertise in designing and analyzing factorial experiments with 2^n levels, and will be able to identify and control confounding factors within factorial designs to ensure the accuracy of results. 4. Identify different types of control charts and their applications in monitoring processes. Optimize the design of control charts to minimize costs while ensuring effective quality control.

TEACHING LEARNING STRATEGIES

- **Lecturing, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	Weightage
End Semester Evaluation(ESE)	**60**
Continuous Evaluation	
Tests	16
Assignment	08
Seminar	16
Total	**40**

Sample Questions to Test Outcomes:

1. Explain Gauss-Markov theorem.
2. Explain Cochran's theorem.
3. Describe the three fundamental principles of experimentation and explain the importance of these principles with respect to designing statistical experiments.
4. Briefly explain contrasts and orthogonal contrasts.
5. Explain in detail the analysis of two-way classified data with one observation per cell.
6. Explain the efficiency of RBD relating to CRD.
7. Distinguish between ANOVA and ANCOVA
8. What do you understand by "Analysis of Covariance"? Illustrate with suitable example.
9. Explain Greaco Latin Square Design and orthogonal latin square design.
10. In a LSD a single observation is missing. How will you estimate the missing value and carry out the analysis of the design?

DISCIPLINE SPECIFIC CORE COURSE

Course Code & Title	**MSBST03DSC10- ANALYSIS OF CLINICAL TRIALS**	
Programme	**M.Sc. Biostatistics**	**Semester III**

Course Objectives	The objective of this course is to study more advanced topics in design and analysis of clinical trials. After successful completion of this course, student will be able to: **1.** Understand Basics of Clinical Trails **2.** Understand design of clinical trials **3.** Understand Sample size determination in clinical trials **4.** Understand the concept of meta analysis in clinical trials.

Modules	Content	Module Outcome
Module I Introduction to clinical trails **(15 Hours)**	Introduction to clinical trails, the need and ethics of clinical trials, bias and random error in clinical studies, Protocols, conduct of clinical trials, over view of Phase I-IV trials, Data management-data definitions, standard operating procedure, informed consent form, case report forms, database design, data collection systems for good clinical practice.	• Understand the need and ethics of clinical trial. • Describe different types of forms used in clinical trial. • Understand about different types of errors that occur during the conduct of clinical trial. • Have a clear idea on clinical practice and clinical data management.
Module II Design of clinical trials **(15 Hours)**	Design of clinical trials- Different phases, Comparative and controlled trials, Random allocation, Randomization, response adaptive methods and restricted randomization. Methods of Blinding, Parallel group designs, Crossover designs, Symmetric designs, Adaptive designs, Group sequential designs, Zelen's designs, design of bioequivalence trials. Outcome measures.	• Understand about different types of randomization procedures used in clinical trials. • Have an idea on different types of bliniding. • Describe different types of designs used in clinical trials.
Module III		• Determine the sample

Statistical analysis in clinical trails **(15 Hours)**	Sample size determination in one and two sample cases, comparative trials, activity studies, testing and other purposes, unequal sample sizes and case of ANOVA. Surrogate endpoints-selection and design of trials with surrogate, analysis of surrogate end point data. Reporting and analysis-Interpretation of result, multi-center trials.	size for different types of response variables. • Describe multi-centre trials. • Have a clear idea on reporting and interpreting of results in a clinical trial. • Describe surrogate endpoints.
Module IV Meta analysis in clinical trials **(15 Hours)**	Meta analysis in clinical trials-concept and goals, fixed and random effect approaches. Bioassay: Direct and indirect assays, Quantal and quantitative assays, Parallel line and slope ratio assays, Design of bioassays.	• Understand different types of bioassay and their analysis. • Describe meta-analysis.
References	***Text Books*** 1. Chen, D.G. and Peace, K.E. (2011). *Clinical Trial Data Analysis Using R*. Chapman & Hall 2. Friedman, L. M., Furburg, C. D. Demets, L. (1998): *Fundamentals of Clinical Trials*, Springer Verlag. 3. Kulinskaya E, Morgeathaler S, Staudte R G(2008). *Meta analysis*, Wiley. ***Reference Books*** 1. Das, M. N. and Giri(2008). *Design of Experiments*, New Age, India 2. Jennison and B.W. Turnbull (1999*): Group Sequential Methods with Applications to Clinical Trials,* CRC Press.	
Course Outcomes	After successful completion of this course, student will be able to: 1. Understand Basics of Clinical Trials. 2. Understand design of clinical trials. 3. Understand Sample size determination in clinical trials. 4. Understand the concept of meta-analysis in clinical trials.	

TEACHING LEARNING STRATEGIES

- **Lecturing, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	Weightage
End Semester Evaluation(ESE)	**60**
Continuous Evaluation	
Tests	16
Assignment	08
Seminar	16
Total	**40**

Sample Questions to Test Outcomes:

1. What is a clinical trial?
2. Differentiate between a prospective an retrospective study.
3. Define response adaptive randomization.
4. Distinguish between blinding and masking
5. Define sample size calculation for independent continuous response variables.
6. Differentiate between analytical dilution assay and comparative dilution assay.
7. Distinguish between adaptive and non-adaptive randomization.
8. Explain about blocked randomization and stratified randomization.
9. Explain in detail about blindness in a clinical trial.
10. Describe the design of adaptive trials and group sequential design.

POOL E: DISCIPLINE SPECIFIC ELECTIVE COURSE (Practical)

Course Code & Title	**MSBST03DSE05- BIOSTATISTICAL COMPUTING USING R – II (Practical)**		
Programme	**M.Sc. Biostatistics**	**Semester**	III

Course Objectives	• To introduce some advanced statistical computing techniques in applied statistics to extract information and visualization thereby enabling them to perform data analysis effectively and efficiently in R programming. • Illustrate different statistical techniques based on all the elective course in third semester.

Modules	Content	Module Outcome
	Statistical Computing III is a practical course. The practical is based on all the elective courses in the third semester.	• Describe different statistical technique to solve problems coming under all the elective courses in third semester.
Course Outcomes	After successful completion of this course, student will be able to: 3. Equipped with different theoretical methods in biostatistics to achieve the objectives. 4. Enhanced with the basic concepts of biostatistical theories besides developing their ability to handle real world problems with large scale data.	

TEACHING LEARNING STRATEGIES

- **Practical sessions through computers, statistical computations, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Hands on training**

ASSESSMENT RUBRICS

Components	Weightage
End Semester Evaluation	**60**
Continuous Evaluation	
Practical Tests	32
Record	8
Total	**40**

POOL E: DISCIPLINE SPECIFIC ELECTIVE COURSE (Practical)

Course Code & Title	**MSBST03DSE06-BIOSTATISTICAL COMPUTING USING SAS- II (PRACTICAL)**		
Programme	**M.Sc. Biostatistics**	**Semester**	III
Course Objectives	• To introduce some advanced biostatistical computing techniques in applied statistics to extract information and visualization thereby enabling them to perform data analysis effectively and efficiently in SAS programming. • Illustrate different biostatistical techniques based on all the elective course in third semester.		

Modules	Content	Module Outcome
	Biostatistical Computing III is a practical course. The practical is based on all the elective courses in the third semester.	• Describe different statistical technique to solve problems coming under all the elective courses in third semester.
Course Outcomes	After successful completion of this course, student will be able to: 1.Equipped with different theoretical methods in biostatistics to achieve the objectives. 2. Enhanced with the basic concepts of biostatistical theories besides developing their ability to handle real world problems with large scale data.	

TEACHING LEARNING STRATEGIES

- **Practical sessions through computers, statistical computations, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Hands on training**

ASSESSMENT RUBRICS

Components	Weightage
End Semester Evaluation	**60**
Continuous Evaluation	
Practical Tests	32
Record	8
Total	**40**

POOL F: DISCIPLINE SPECIFIC ELECTIVE COURSE

Course Code & Title	**MSBST03DSE07-STOCHASTIC PROCESS AND TIME SERIES ANALYSIS**	
Programme	**M.Sc. Biostatistics**	**Semester III**
Course Objectives	The course will help the Students to develop a comprehensive understanding of stochastic processes, particularly focusing on Markov chains, including their definition, classification, and real-world examples. It also help the student to delve in to basic concepts of time series analysis and time series modelling. Also Students will develop skills in time series analysis, including autocorrelation, stationarity, and regression modeling.	

Modules	Content	Module Outcome
Module I: Introduction to stochastic process **(16 Hours)**	Introduction to stochastic process, Markov Chains: Definition, Examples and classification, Discrete renewal equation and basic limit theorem, Absorption probabilities, Criteria for recurrence.	• Define Markov chains and classify them based on various criteria. • Understand the discrete renewal equation and its significance in the context of Markov chains. • Calculate absorption probabilities for absorbing

		Markov chains. • Understand the implications of recurrence on the long-term behavior of Markov chains.
Module II: Continuous time Markov chains **(14 Hours)**	Continuous time Markov chains, Examples, General pure birth process, Poisson process, Birth and death process, Finite state continuous time Markov chains, Applications to queuing models.	• Provide examples of continuous-time Markov chains in various domains. • Understand the concept of pure birth processes and Poisson processes. • Analyze birth and death processes in continuous time Markov chains. • Analyze finite-state continuous-time Markov chains using mathematical tools.
Module III: Introduction to time series analysis **(16 Hours)**	Characteristics of time series: Time series as a discrete parameter stochastic process, Autocorrelation (ACF) and cross correlations, Stationary time series, Estimation of autocorrelations. Classical regression in time series context, exploratory data analysis, smoothing methods for time series. Wold representation of linear stationary processes.	• Understand time series as a discrete parameter stochastic process. • Calculate autocorrelation functions (ACF) and cross-correlations for time series data. • Define stationary time series and understand its importance in time series analysis. • Estimate autocorrelations and conduct classical regression analysis for time series data.
Module IV: Linear time series model **(14 Hours)**	Linear time series models : Autoregressive (AR), Moving Average (MA), Autoregressive Moving Average (ARMA) and	• Understand the concepts and properties of autoregressive (AR) and moving average (MA) models.

	Autoregressive Integrated Moving Average (ARIMA) models. Forecasting and estimation of ARMA models. Seasonal ARIMA models, Residual analysis and diagnostic checking.	• Understand the structure and parameters of autoregressive moving average (ARMA) models. • Apply ARIMA models to analyze and forecast time series data with trend and seasonality. • Conduct residual analysis and diagnostic checking to assess the adequacy of time series models.
References	***Text Books*** 1. Karlin.S. and Taylor, H.M. (1975) A First Course in Stochastic Processes, second edition, Academic Press. 2. Bhat, B.R. (2002) Stochastic Processes, second edition, New Age Publication. 3. Shumway, R. H and Stoffer, D. S. (2006). Time series Analysis and its Applications. Springer. 4. Box, G. E. P. Jenkins, G. M. and Reinsel, G. C. (1994). Time Series Analysis: Forecasting and Control, Pearson Education. ***Reference Books*** 1. Feller, W. (1965, 1968), An Introduction to Probability Theory and its Applications, Volume I and II, Wiley Eastern. 2. Bhat, U.N. (1984) Elements of Applied Stochastic Processes, John Wiley. 3. Cinlar, E. (1975) Introduction to Stochastic Processes, Prentice Hall. 4. Brockwell, P.J and Davis R.A. (2006) Time Series: Theory and Methods, 2ndedn. Springer-Verlag 5. Chatfield, C. (2004) The Analysis of Time Series - An Introduction, Sixth edition, Chapman and Hall. 6. Anderson, T.W (1971) Statistical Analysis of Time Series, Wiley.	
Course Outcomes	After successful completion of this course, student will be able to: 1. Understand the implications of recurrence on the long-term behavior of Markov chains.	

	2. Determine key characteristics and parameters of birth and death processes. 3. Apply AR and MA models to analyze and forecast time series data. 4. Apply ARIMA models to analyze and forecast time series data with trend and seasonality.

TEACHING LEARNING STRATEGIES

- **Lecturing, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	Weightage
End Semester Evaluation(ESE)	**60**
Continuous Evaluation	
Tests	16
Assignment	08
Seminar	16
Total	**40**

Sample Questions to Test Outcomes:

1. Define Markov chain.
2. Prove that a Markov chain is completely determined by its initial distribution and one step TPM.
3. Define periodic and aperiodic Markov chains.
4. Consider Markov chain whose TPM is

$$\begin{pmatrix} 0 & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

(i)Is the chain is irreducible.

(ii)Is the chain is ergodic? Explain.

(iii)Find the stationary distribution of the chain.

5. Show that in an irreducible Markov chain all the states are of same type.
6. Show that probability of extinction of Galton Watson branching process is the smallest positive root of the equation s = P(s), where P(s) is the PGF of the offspring distribution.
7. Suppose that customers arrive at a bank according to a Poisson process with mean rate of 3 per minute. Find the probability that during a time interval of 2 minutes (i)exactly 4 customers arrive (ii) more than 4 customers arrive and (iii) at least one customer arrive.
8. Describe the relation between Poisson process and binomial distribution.
9. Explain renewal process. Give examples
10. Describe the stationary behavior of a birth and death processes and hence obtain the system size distribution of M/M/1 and M/M/c queueing models.

POOL F: DISCIPLINE SPECIFIC ELECTIVE COURSE

Course Code & Title	**MSBST03DSE08-APPLIED MULTIVARIATE ANALYSIS**	
Programme	**M.Sc. Biostatistics**	**Semester III**

Course Objectives	• Understand the fundamentals of multivariate data analysis and the notion of multivariate distributions. • Demonstrate proficiency in analyzing multivariate data using the multivariate normal distribution, including understanding marginal and conditional distributions. • Master the concept of characteristic functions and their application in multivariate data analysis. • Develop skills in estimating the mean vector and covariance matrix of multivariate datasets. • Interpret canonical variates and canonical correlations, both in population and sample contexts, as part of Canonical Correlation Analysis (CCA). • Apply orthogonal factor models in Factor Analysis, including methods of estimation, factor rotation, and computation of factor scores. • Utilize various similarity measures and hierarchical and non-hierarchical clustering methods in Cluster Analysis.

Modules	Content	Module Outcome
Module I Multivariate normal distribution **(15 Hours)**	Multivariate data, preliminary analysis, notion of multivariate distributions, multivariate normal distribution, marginal and conditional distributions, characteristic function, estimation of mean vector and covariance matrix.	• Able to identify and describe the notion of multivariate distributions. • Learn about marginal and conditional distributions in the context of multivariate data. • Derive characteristic functions and their importance in the analysis of multivariate data.
Module II Principal component and canonical correlation	Principal components Analysis: - population principal components, summarizing sample variation by principal components, graphing the principal components; Canonical correlation analysis: - canonical variates	• Students will demonstrate a comprehensive understanding of PCA. • Summarize sample variation effectively using

analysis **(15 Hours)**	and canonical correlations, interpreting the population canonical variables, the sample canonical variates and sample canonical correlations.	principal components. • Master the principles of Canonical Correlation Analysis,
Module III Factor and cluster analysis **(15 Hours)**	Factor analysis: - orthogonal factor model; methods of estimation, factor rotation, factor scores; Cluster analysis: - similarity measures, hierarchical clustering methods, non-hierarchical clustering methods.	• Demonstrate a comprehensive understanding of factor analysis including the orthogonal factor model, methods of estimation. • Develop proficiency in applying factor analysis techniques, including estimating factor loadings, conducting factor rotation to simplify interpretation. • Demonstrate the ability to apply cluster analysis techniques to real-world datasets.
Module IV: MANOVA and Multidimensional scaling **(15 Hours)**	Comparison of several multivariate population means (one-way MANOVA), simultaneous confidence intervals for treatment effects, two-way multivariate analysis of variance; Distance methods: - multidimensional scaling, correspondence analysis.	• Demonstrate proficiency in comparing several multivariate population means using one-way Multivariate Analysis of Variance (MANOVA). • Construct simultaneous confidence intervals for treatment effects in multivariate data analysis settings. • Proficiency in distance-based multivariate analysis methods.

References	***Text Books*** 1. Johnson, R.A. and Wichern, D.W. (2007) Applied Multivariate Statistical Analysis, PHI Learning Private Ltd, New Delhi, Sixth edition. 2. Rencher, A.C. (1995) Methods of Multivariate Analysis, John Wiley. 3. Dillon, W.R. and Goldstein, M (1984) Multivariate Analysis, John Wiley. ***Reference Books*** 1. Anderson, T.W. (1984) An Introduction to Multivariate Statistical Analysis, John Wiley. 2. Seber G.A.F. (1983) Multivariate Observations, Wiley. 3. Tabachnick, B.G. and Fidell, L.S. (2018) Using multivariate statistics, Sixth edition, Pearson India Education Services Pvt Ltd, India.
Course Outcomes	After successful completion of this course, student will be able to: 1. Demonstrate a comprehensive understanding of various multivariate data analysis techniques. 2. Develop proficiency in dimensionality reduction techniques such as Principal Components Analysis (PCA) and Factor Analysis. 3. Acquire advanced multivariate analysis skills, including conducting comparisons of several multivariate population means using one-way MANOVA, simultaneous confidence interval construction for treatment effects. 4. Master the analysis of multivariate relationships using techniques such as Canonical Correlation Analysis (CCA), cluster analysis, and distance methods.

TEACHING LEARNING STRATEGIES

- **Lecturing, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	Weightage
End Semester Evaluation(ESE)	**60**
Continuous Evaluation	

Tests	16
Assignment	08
Seminar	16
Total	**40**

Sample Questions to Test Outcomes:

1. Define a singular multivariate normal distribution. Give an example of a random vector following singular multivariate normal distribution.
2. If X ~Np(μ, Σ) and if Σ is a diagonal matrix then show that the components of X are independently normally distributed and conversely.
3. Give an example to show that the marginal distributions are normals does not imply that the joint distribution follows a multivariate normal distribution.
4. Define partial correlation. Explain how do you evaluate the partial correlation coefficients from a non-singular covariance matrix.
5. Define multiple correlation. Show that $0 \leq \rho 1.23...p \leq 1$.
6. Derive the distribution of sample mean of a sample of size n from Np(μ, Σ).
7. Find the MLE of the correlation matrix based on a random sample of size n from Np(μ, Σ),Σ > O, when μ is known.
8. Define Wishart distribution and derive its characteristic function.
9. Derive the characteristic function of a matrix-variate gamma distribution.
10. Derive the characteristic function of a Wishart matrix and show that Wishart distribution is a matrix variate generalization of $\chi 2$ distribution.

POOL F: DISCIPLINE SPECIFIC ELECTIVE COURSE

Course Code & Title	**MSBST03DSE09 - MACHINE LEARNING TECHNIQUES FOR BIOSTATISTICS**		
Programme	**M.Sc. Biostatistics**	**Semester**	III

Course Objectives	• To classify variables and articulate their roles within statistical modeling, distinguishing between categorical, numerical, and ordinal variables, and understanding their implications for analysis. • Demonstrate competency in applying least squares regression and nearest neighbors algorithms for data analysis. • Master supervised learning techniques and function approximation methods, utilizing them to model relationships between input and output variables effectively. • Demonstrate proficiency in utilizing roughness penalty methods and Bayesian approaches to estimate model parameters, understanding their roles in controlling model complexity and improving generalization. • Proficient in employing kernel methods and local regression techniques for modeling non-linear relationships in data.

Modules	Content	Module Outcome
Module I Introduction and overview of supervised learning **(15 Hours)**	Supervised Learning and Function Approximation, A Statistical Model for the Joint Distribution of input and output vectors, Function Approximation, Structured Regression Models, Linear Methods for Regression: Least squares, Subset selection, Shrinkage Methods, Methods using derived input directions, Multiple outcome shrinkage and selection, Lasso and related path algorithms.	• Explore structured regression models and classes of restricted estimators, including roughness penalty and Bayesian methods. • Understand the practical considerations in automatic selection of smoothing parameters and nonparametric logistic regression.
Module II: Linear methods for	Linear methods for classification using linear regression of an indicator matrix, linear discriminant analysis, logistic regression and separating hyperplanes. Basis expansions and	• Explore linear discriminant analysis (LDA) and its role in

classification **(15 Hours)**	regularizations: Piecewise polynomials and splines, Automatic Selection of the Smoothing Parameters, Nonparametric Logistic Regression, Multidimensional Splines.	multi-class classification problems. • Implement piecewise polynomials and splines for capturing complex functional forms in regression and classification tasks.
Module III: Kernal smoothing **(15 Hours)**	One-Dimensional Kernel Smoothers, Selecting the band width of the Kernel, Structured Local Regression Models in R^p, Local Likelihood and Other Models, Kernel Density Estimation and Classification: Kernel Density Estimation, Kernel Density classification and the Naïve Bayes classifier. Mixture Models for Density Estimation and Classification.	• Explore structured local regression models in multidimensional spaces and their advantages in capturing complex data patterns. • Implement kernel density classification methods for classifying data points based on their estimated densities. • Understand the mathematical properties of RBFs and kernels and their role in non-linear transformations of input data.
Module IV: Model assessment, inference and averaging **(15 Hours)**	Bias, Variance and Model Complexity, The Bias–Variance Decomposition, Optimism of the Training Error Rate, Estimates of In-Sample Prediction Error, The Bayesian Approach and BIC, Minimum Description Length, Cross-Validation, Bootstrap Methods, Conditional or Expected Test Error, introducing Model Inference and averaging: Local regression in IR, The EM Algorithm,	• Gain proficiency in techniques such as cross-validation and bootstrap methods for model evaluation and selection. • Understand the principles behind the Bayesian Information

	MCMC for Sampling from the Posterior, Bagging, Model Averaging and Stacking, Stochastic Search: Bumping.	Criterion (BIC) and minimum description length for model selection.
References	***Text Books*** 1. Hastie, T., Tshibirai, R. and Friedman, J. (2017) The Elements of Statistical Learning : Data Mining, Inference and Prediction, 2nd edition. Springer, New York. 2. James, G., Witten, D., Hastie, T. and Tibshirani, R.(2013) An Introduction to Statistical Learning with Applications in R. Springer, New York. ***Reference Books*** 1. James, G., Witten,D., Tibshirani, R. and Hastie, T. Neural Networks and Deep Learning: A Textbook. 2. Introduction to Machine Learning The Wikipedia Guide.	
Course Outcomes	After successful completion of this course, student will be able to: 1. Gain proficiency in applying least squares regression, nearest neighbors, and local methods for high-dimensional data analysis. 2. Master linear methods for regression and classification, including subset selection, shrinkage methods, and linear discriminant analysis. 3. Learn techniques for estimating in-sample prediction error, including cross-validation and bootstrap methods. 4. Explore advanced model inference methods, including maximum likelihood estimation, Bayesian inference, and the EM algorithm.	

TEACHING LEARNING STRATEGIES

- **Lecturing, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	Weightage
End Semester Evaluation(ESE)	**60**
Continuous Evaluation	

Tests	16
Assignment	08
Seminar	16
Total	**40**

Sample Questions to Test Outcomes:

1. How do least squares and nearest neighbors differ in their approach to supervised learning, and what are the advantages and limitations of each method in terms of handling high-dimensional data?
2. Can you explain the concept of the bias-variance tradeoff in the context of model selection? How do different methods such as Lasso, subset selection, and shrinkage methods address this tradeoff, and under what conditions would one method be preferred over another?
3. What are the key principles behind kernel methods and local regression in function approximation? How do these methods handle non-linear relationships between variables, and what are some practical considerations when choosing between different kernel functions or regression approaches?
4. How do linear methods for classification, such as logistic regression and linear discriminant analysis, differ in their approach to separating classes? Discuss the advantages and limitations of each method in terms of handling non-linearly separable data and the assumptions underlying their models.
5. What role does feature extraction play in filtering and smoothing techniques, particularly in the context of multidimensional splines and wavelet smoothing?
6. What are the key principles behind one-dimensional kernel smoothers, and how do they differ from other smoothing techniques such as splines or local regression?
7. How do structured local regression models in multidimensional spaces (Rp) extend the concepts of one-dimensional kernel smoothers?

POOL F: DISCIPLINE SPECIFIC ELECTIVE COURSE

Course Code & Title	**MSBST03DSE10-CATEGORICAL DATA ANALYSIS**		
Programme	**M.Sc. Biostatistics**	**Semester**	III
Course Objectives	• Demonstrate a comprehensive understanding of categorical data and their measures, including appropriate techniques for analysis. • Gain proficiency in conducting inference for contingency tables, applying appropriate statistical tests and interpreting results effectively. • Develop a deep understanding of generalized linear models, particularly focusing on binary and count data. • Gain proficiency in handling longitudinal data, understanding its characteristics. • Gain practical experience in fitting general linear mixed effect models, conducting inference for random effects, and interpreting results.		

Modules	Content	Module Outcome
Module I Categorical data and their measures **(15 Hours)**	Categorical data and their measures, Inference for contingency tables, Generalized linear models for binary and count data. Estimation, Inference and fitting of model.	• Demonstrate a thorough understanding of categorical data and their measures. • Proficient in conducting inference for contingency tables. • Experience in applying generalized linear models to binary and count data.
Module II Logit and logistic models **(15 Hours)**	Logistic, logit and log linear models with categorical predictors, Logit models with multi responses- Nominal and ordinal responses.	• Understanding Logistic, Logit, and Log-linear Models. • Proficient in specifying and estimating logistic, logit, and

		log-linear models with categorical predictors. • Apply logit models to datasets with multi-responses, encompassing both nominal and ordinal responses.
Module III: Longitudinal data and their characteristics **(15 Hours)**	Longitudinal data and their characteristics, The general linear model for longitudinal data-ML and REML estimation, EM algorithm, General linear mixed effect model. Inference for the random effects. BLUPs, Empirical Bayes, Shrinkage model building and diagnostics, Generalized additive mixed model.	• Develop a comprehensive understanding of longitudinal data and their unique characteristics. • Proficient in applying the general linear model to longitudinal data. • Fitting general linear mixed effect models to longitudinal data.
Module IV: Generalised linear model for longitudinal data **(15 Hours)**	Generalised linear model for longitudinal data, Random effect model, Transition models, Poisson and logistic regression models, Analysis and test. Classification of missing data mechanism- intermittent missing values and dropouts, weighted estimating equations, Modeling the drop out process.	• Explain the concepts and principles behind generalized linear models (GLMs) and their application to longitudinal data analysis. • Develop a thorough understanding of random effect models and their role in analyzing longitudinal data with correlated observations. • Understand the importance of model validation and diagnostic checks in longitudinal data analysis, and be able to apply these techniques effectively.

References	***Text Books*** 1. Agresti, A. (2012). *Categorical data analysis* (Vol. 792). John Wiley & Sons. 2. Diggle, P.J, Heagerty, P., Liang, K.,Y & Zeger,S., I (2003), Analysis of longitudinal data, Oxford university press. 3. Lindsey, J.,K.(1993) Models for repeated measurements, Oxford ***Reference Books*** 1. Weiss, R., E.(2005), Modelling longitudinal data, Springer, New York. 2. Little, R. J. A. & Rubin, D., B(2002), Statistical analysis with missing data, Wiley.
Course Outcomes	After successful completion of this course, student will be able to: 1. Conduct analysis of longitudinal data. 2. Apply statistical techniques to model longitudinal data and make predictions. 3. Understand analysis of longitudinal data with missing data. 4. Understand analysis of longitudinal data with time-dependent covariates.

TEACHING LEARNING STRATEGIES

- **Lecturing, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	**Weightage**
End Semester Evaluation(ESE)	**60**
Continuous Evaluation	
Tests	16
Assignment	08
Seminar	16
Total	**40**

Sample Questions to Test Outcomes:

1. What are categorical data, and what measures are commonly used to summarize and analyze them?
2. Discuss common statistical tests used for assessing the association or independence between categorical variables in contingency tables.
3. Provide examples of binary and count data scenarios where GLMs are appropriate and discuss the interpretation of model coefficients in these contexts.
4. Describe the process of model estimation and interpretation in logistic regression models.
5. Discuss strategies for handling issues such as missing data, multicollinearity, and model validation to ensure the reliability and generalizability of the model results.
6. Compare and contrast maximum likelihood (ML) and restricted maximum likelihood (REML) estimation techniques in the context of longitudinal data analysis. Provide examples to illustrate your points.
7. Explain the concept of a general linear mixed effect model (GLMM) and its relevance in analyzing longitudinal data.
8. Define the Generalized Additive Mixed Model (GAMM) and discuss its advantages in modeling longitudinal data with non-linear relationships.
9. Explain the concept of the Generalized Linear Model (GLM) for longitudinal data analysis.
10. Define transition models and discuss their significance in longitudinal data analysis.

POOL F: DISCIPLINE SPECIFIC ELECTIVE COURSE

Course Code & Title	**MSBST03DSE11-OPERATIONS RESEARCH**		
Programme	**M.Sc. Biostatistics**	**Semester**	III
Course Objectives	• Understand the fundamental concepts of linear programming problems (LPP) and their applications in real-world optimization scenarios. • Develop proficiency in graphical solution techniques for visualizing and analyzing LPPs, including identifying feasible regions and optimal solutions.		

	• Demonstrate the ability to determine feasible, basic feasible, and optimum basic feasible solutions to LPPs, both graphically and analytically. • Gain insight into the theoretical foundations of linear programming, including the analytical results and theoretical development of the simplex method. • Master advanced solution techniques such as the use of artificial variables, the Big-M method, and the two-phase simplex method to solve complex LPPs. • Explore practical optimization problems such as transportation and assignment problems, and learn to apply appropriate solution methods to address them effectively.

Modules	Content	Module Outcome
Module I Introduction to linear programming problem **(15 Hours)**	Introduction to linear programming problem (LPP),graphical solution, feasible, basic feasible, and optimum basic feasible solution to an LPP. Analytical results in general LPP, theoretical development of simplex method.	• Define linear programming problems and their significance in optimization. • Explain the concept of feasible solutions and their graphical representation. • Identify basic feasible solutions and optimal basic feasible solutions in LPPs.
Module II Simplex methods **(15 Hours)**	Artificial variables, Big-M method, two phase simplex method Duality, duality theorems, dual simplex methods.	• Derive analytical results for general linear programming problems. • Understand the theoretical development of the simplex method for solving LPPs. • Apply artificial variables, the Big-M method, and the two-phase simplex method to handle

		special cases in LPPs.
Module III Transportation and integer programming **(15 Hours)**	Transportation problem, assignment problem. Integer programming: Cutting plane methods, branch and bound technique. Network analysis, Critical path analysis.	• Explore the concept of duality in linear programming and understand the duality theorems. • Discuss the application of dual simplex methods in solving LPPs. • Analyze transportation and assignment problems and apply appropriate solution techniques.
Module IV Game theory and applications **(15 Hours)**	Game theory, pure and mixed strategies, conversion of two-person zero gain to a linear programming problem. Solution to game through algebraic, graphical and linear programming method.	• Examine integer programming and its solution methods. • Explore network analysis techniques such as Critical Path Method (CPM) and Program Evaluation and Review Technique (PERT). • Introduce game theory concepts, including pure and mixed strategies.
References	**Text Books** 1. K.V. Mital and Mohan, C (1996). Optimization Methods in Operations Research and SystemsAnalysis,3rd Edition, New Age International(Pvt.)Ltd. 2. Kanti Swarup, Gupta, P. K.and John, M. M. (1985): Operations Research., Sultan Chand & Sons. **Reference Books** 1. Hadley, G.(1964).Linear Programming, Oxford& IBH Publishing Co, New Delhi. 2. Taha.H.A.(1982): Operation Research, An Instruction, Macmillan. 3. HillerF.S.AndLieberman,G.J.(1995).IntroductiontoOperationsResearch,McGrawHill	
Course	After successful completion of this course, student will be able to:	

Outcomes	1. Identify and develop operational research models from the verbal description of the real system. 2. Understand the mathematical tools that are needed to solve optimization problems. 3. Understand various methods in Integer programming and Game theory.

TEACHING LEARNING STRATEGIES

- **Lecturing, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	Weightage
End Semester Evaluation(ESE)	**60**
Continuous Evaluation	
Tests	16
Assignment	08
Seminar/Viva	16
Total	**40**

Sample Questions to Test Outcomes:

1. Define Linear Programming Problem (LPP) and discuss its significance in optimization. What are the key components of an LPP, and how is it formulated?

2. Explain the graphical solution method for solving LPPs. How is the feasible region identified, and how are optimal solutions determined graphically?

3. Define feasible solutions, basic feasible solutions, and optimum basic feasible solutions to an LPP. What criteria are used to identify these solutions?

4. Discuss the theoretical development of the simplex method for solving LPPs analytically. How does the simplex method iteratively move from one basic feasible solution to another to find the optimal solution?

5. Explain the concept of artificial variables and the Big-M method in the context of solving LPPs. How are artificial variables used to convert an LPP into a form suitable for the simplex method?

6. Describe the two-phase simplex method for solving LPPs. What are the two phases, and how does this method handle special cases such as degeneracy?

7. Discuss duality in linear programming, including duality theorems and the dual simplex method. What is the relationship between the primal and dual problems, and how is duality used in optimization?

8. Define the transportation problem and the assignment problem. How are these types of problems formulated as linear programming problems, and what are their applications?

POOL F: DISCIPLINE SPECIFIC ELECTIVE COURSE

Course Code & Title	**MSBST03DSE12- DATA VISUALIZATION AND ANALYSIS USING PYTHON**		
Programme	**M.Sc. Biostatistics**	**Semester**	III
Course Objectives	• Gain proficiency in data acquisition and manipulation using Python libraries. • Understand and implement different data visualization techniques for comprehensive data analysis. • Learn the regression model building framework in Python, from problem definition to model validation. • Develop skills in multiple linear regression including handling categorical variables and diagnosing model issues with Python.		

Modules	**Content**	**Module Outcome**
Module I Introduction to PYTHON **(15 Hours)**	Introduction to PYTHON, Data acquisition processes, extraction, cleaning, annotation, integration, reduction, and transformation. Basic charts, multivariate visualization, pixel-oriented, geometric projection, icon-based, and hierarchical visualization.	• Ability to effectively extract, clean, annotate, integrate, reduce, and transform data using Python. • Understanding of data preprocessing techniques to prepare datasets for analysis, including data normalization,

		missing value imputation, and feature engineering.
Module II Data visualization tools **(15 Hours)**	Data visualization tools, rank analysis, trend analysis, multivariate analysis, distribution analysis, correlation analysis, and geographical analysis.	• Proficiency in utilizing advanced visualization methods such as pixel-oriented, geometric projection, icon-based, and hierarchical visualization. • Ability to communicate insights and findings derived from data visualization effectively.
Module III Regression model building framework **(15 Hours)**	Regression model building framework, covering problem definition, data pre-processing, model building, diagnostics, and validation. Simple linear regression, coefficients of determination, significance tests, residual analysis, and confidence/prediction intervals.	• Understanding of best practices in data visualization design. • Ability to leverage interactive visualization tools and dashboards to engage stakeholders and facilitate exploratory data analysis and decision-making processes.
Module IV Multiple linear regression **(15 Hours)**	Multiple linear regression, coefficients of multiple determination, interpretation of regression coefficients, categorical variables, diagnosing issues, heteroscedasticity and multicollinearity, outliers, autoregression, and variable transformation for robust regression model building.	• Understanding of the regression model building framework and evaluation using Python. • Capability to validate regression models using diagnostic techniques, including confidence and prediction intervals, to assess model accuracy and reliability for predictive analysis.

References	***Text Books*** 1. AndyKirk, Data Visualization a Handbook for Data Driven Design, Sage Publications, 2016 2. Philipp K.Janert, Gnuplot in Action, Understanding Data with Graphs, Manning Publications,2010. ***Reference Books*** 1. Alberto Cordoba, "Understanding the Predictive Analytics Lifecycle", Wiley,2014. 2. Eric Siegel, Thomas H. Davenport, "Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die", Wiley, 2013. 3. James R Evans, "Business Analytics-Methods, Models and Decisions", Pearson 2013. 4. R. N. Prasad, Seema Acharya, "Fundamentals of Business Analytics", Wiley,2015. 5. Perkovie, L. (2011). Introduction to computing using python: An Application development focus. Wiley Publishing. 6. McKinney, W. (2012). Python for data analysis: Data wrangling with Pandas, NumPy, and IPython. "O Reilly Media, Inc."
Course Outcomes	After successful completion of this course, student will be able to: 1. Apply Python libraries for data acquisition to extract, clean, integrate, and transform datasets efficiently. 2. Utilize Python libraries for various visualization methods to interpret and communicate complex data effectively. 3. Construct regression models using Python libraries, including appropriate pre-processing techniques and diagnostic tools. 4. Analyze simple linear regression results, including coefficients of determination and significance tests, using Python. 5. Implement multiple linear regression models in Python, addressing issues like multicollinearity and heteroscedasticity for robust predictions.

TEACHING LEARNING STRATEGIES

- **Lecturing, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	Weightage
End Semester Evaluation(ESE)	**60**
Continuous Evaluation	
Tests	16
Assignment	08
Seminar	16
Total	**40**

Sample Questions to Test Outcomes:

1. How would you describe the process of data acquisition and what are its key components?
2. Can you explain the difference between data cleaning and data transformation?
3. What are some common techniques used for reducing the dimensionality of datasets during data preprocessing?
4. Describe a situation where you might encounter missing data during the data acquisition process and how would you handle it?
5. How do you decide which visualization technique to use for a given dataset and analysis objective?
6. Explain the purpose of rank analysis tools in data visualization and provide an example of when you would use them.
7. What is the significance of residual analysis in regression model building and how is it performed?
8. Describe the steps involved in conducting a simple linear regression analysis and interpreting its results.
9. How do you diagnose multicollinearity in a multiple linear regression model and what are its potential consequences?

10. Can you provide an example of how you would use autoregression in a regression model building process and explain its importance?

POOL G: INTERDISCIPLINARY ELECTIVE COURSE (IDC)

Course Code & Title	**MSBST03IDC01-STATISTICAL DATA ANALYSIS USING SPSS**		
Programme Offered	**Department of Statistical Sciences**	**Semester**	III
Course Objectives	• Demonstrate proficiency in navigating SPSS interface components including the Data Editor window, SPSS Output window, and various drop-down menus. • Able to create, modify, import, and transform datasets using SPSS. • Gain an understanding of different types of variables and how to assign appropriate labels to them within SPSS. • Learn to conduct inferential statistical analysis techniques such as frequencies, cross tabs, independent sample t-tests, paired sample t-tests, ANOVA, correlation, regression, confidence intervals, and non-parametric tests using SPSS. • Develop the ability to interpret and communicate the results of statistical analyses conducted in SPSS.		

Modules	Content	Module Outcome
Module I Introduction to SPSS and its interface **(15 Hours)**	SPSS windows, Data editor window, Types of variables and labels, SPSS output window, Drop down menus of SPSS, Creating and modifying data files	• Proficiency in navigate various SPSS windows. • Able to differentiate between categorical and continuous variables and understand their significance in data analysis. • Creating and modifying data files in SPSS.
Module II	Import of data files, Transform drop	• Skills to transform variables

Data management in SPSS **(15 Hours)**	down menu, computing variables, Recode option, Data drop down menu, Split files, Weight cases, Select cases.	using the Transform drop-down menu in SPSS. • They will be able to perform operations such as computing new variables, recoding existing variables. • Able to utilize advanced data manipulation features of SPSS.
Module III Graphics in SPSS **(15 Hours)**	SPSS plots and graphs, Bar diagram, Pie diagram, Multiple bar diagram, Histogram, Box plot, P-P plot, Q-Q plot, Scatter diagram.	• Demonstrate proficiency in creating various plots and graphs in SPSS. • Develop the skills to interpret and analyze graphical representations generated in SPSS.
Module IV Basic statistical analysis **(15 Hours)**	Analyze drop down menu, Descriptive statistics, Frequencies, Cross tabs, Compare means-independent sample t test, paired sample t test, ANOVA, Correlation, Regression, Confidence intervals, Non-parametric test.	• Demonstrate proficiency in conducting various statistical analyses using the Analyze drop-down menu in SPSS. • Gain proficiency in conducting inferential statistics tests such as frequencies, cross tabs, independent sample t-tests, paired sample t-tests, ANOVA, correlation, regression.
References	***Text Book*** 1. Hinton P R, Brownlow C, McMurray, I. and Cozens, B.(2004): *SPSS Explained,* Routledge, Taylor and Francis group, New York. ***Reference Book*** 1. Sabine Landau, Brian S. Everitt (2003): A Handbook of Statistical Analyses Using SPSS, New York.	
Course Outcomes	After successful completion of this course, student will be able to: 1. Understand the usage of menus in SPSS window for drawing	

	various diagrams. 2. Computing descriptive statistics, the comparison of means, ANOVA, non-parametric tests, simple correlation and regression procedures and apply for real data sets. 3. Acquire the skills of plotting different graphs using SPSS.

TEACHING LEARNING STRATEGIES

- **Lecturing, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	**Weightage**
End Semester Evaluation(ESE)	**60**
Continuous Evaluation	
Tests	16
Assignment	08
Seminar	16
Total	**40**

Sample Questions to Test Outcomes:

1. Describe the use of Recode command in SPSS
2. Describe the use of Weight case facility in SPSS
3. What is the use of split file option in SPSS?
4. Distinguish between Value label and Values options in SPSS.
5. Explain the method of constructing and interpreting a Boxplot
6. Explain crosstab facility and its uses available in SPSS
7. How do you construct a frequency table and histogram using SPSS?
8. Explain Transform dropdown menu in SPSS.
9. Describe briefly the various options available in SPSS Analyze menu.

10. Explain different methods for constructing graphs in SPSS.

POOL G: INTERDISCIPLINARY ELECTIVE COURSE

Course Code & Title	**MSBST03IDC02-STATISTICAL DATA ANALYSIS USING R**		
Programme offered	**Department of Statistical Sciences**	**Semester**	III
Course Objectives	• Demonstrate a solid understanding of the fundamental concepts in R programming, including objects and their classes, operators, vectors, matrices, lists, and data frames. • Gain proficiency in indexing and accessing data within R, as well as importing and exporting data from various file formats. • Able to apply R programming to compute descriptive statistics. • Learn to create various graphical representations of data using R. • Develop the skills to plot cumulative distribution functions (CDFs) and probability density functions (PDFs) for various values of parameters in standard probability distributions using R. • Learn to generate random samples from standard probability distributions in R.		

Modules	Content	Module Outcome
Module I Introduction to R **(15 Hours)**	Introduction to R- Objects and their classes, operators, vectors and matrices, list and data frames, indexing and accessing data, importing and exporting data. Common built-in functions. Simple applications - Descriptive statistics.	• Demonstrate a solid understanding of the fundamental concepts of R programming. • Acquire proficiency in indexing and accessing data within R, enabling them to effectively manipulate datasets for analysis. • Able to apply basic descriptive statistics techniques in R to analyze datasets effectively.

Module II R-Graphics **(15 Hours)**	R-Graphics- Histogram, Box-plot, Stem and leaf plot, Scatter plot, Q-Q plot. Looping- for loop, repeat loop, while loop, if command, ifelse command.	• Demonstrate proficiency in creating various types of graphical representations using R, including histograms, box plots, etc. • Able to apply their knowledge of R graphics and looping structures to conduct exploratory data analysis and statistical inference tasks.
Module III Basic probability and distribution **(15 Hours)**	Basic concepts of probability and random variables, Probability distributions (Binomial, Poisson, Geometric, Uniform, Normal, Gamma, Beta), Plotting of cdf and pdf for different values of the parameters of standard distributions. Generations of random samples from standard distributions.	• Understanding of basic concepts of probability theory and random variables. • Understanding different probability distributions commonly used in statistical analysis. • Demonstrate the ability to plot cumulative distribution functions (CDFs) and probability density functions (PDFs) for different parameter values of standard distributions using R.
Module IV Descriptive statistics **(15 Hours)**	The Descriptive statistics, the comparison of means, ANOVA, non-parametric tests, correlation and regression procedures.	• Demonstrate mastery in descriptive statistics, including measures of central tendency, dispersion, and distributional shape. • Able to conduct and interpret various tests for comparing means, including independent samples t-tests, paired samples t-tests, etc.

References	***Text Books*** 1. Purohit, S. G, Ghore, S. D and Deshmukh, S. R. (2004): *Statistics Using R*. Narosa. ***Reference Books*** 1. Dalgaard, P. (2008): *Introductory Statistics with R, (Second Edition),* Springer.
Course Outcomes	After successful completion of this course, student will be able to: 1. Understand various built-in functions in R programming for statistical data analysis. 2. Understand different functions in R programming for writing computer programmes and develop computer programmes for different problems. 3. Understand different statistical test using R software

TEACHING LEARNING STRATEGIES

- **Lecturing, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	**Weightage**
End Semester Evaluation(ESE)	**60**
Continuous Evaluation	
Tests	16
Assignment	08
Seminar	16
Total	**40**

Sample Questions to Test Outcomes:

1. How will you install R in your computer?

2. How will you save, store and retrieve workspace in R?
3. Explain the seq() and rep() in R with examples.
4. How will you import data from excel to R?
5. What are the advantages of R over other statistical softwares?
6. Explain different types of arithmetic operators and assignment operators in R. Give examples.
7. Explain different ways of defining matrices in R.
8. Explain the different forms of sequence function in R. Give examples.
9. Expalin the built-in functions in R with examples.
10. Describe low level plotting functions in R.

	FOURTH SEMESTER								
Sl No	**Course Code**	**Title of Paper**	**Contact Hours/Week**			**Marks**			
			L	**T/S**	**P**	**ESE**	**CE**	**Total**	**Credits**
4.1	MSBST04DSC12	Project/Dissertation and Subject Viva				**60**	**40**	**100**	**12**
						60	**40**	**100**	**12**
DISCIPLINE SPECIFIC ELECTIVES (DSE)									
4.2	MSBST04DSExx	Elective-I (DSE) (Practical) (One course has to be chosen from **Pool H**)	3	2		60	40	100	3
4.3	MSBST04DSExx	Elective-II (DSE) (One course has to be chosen from **Pool I**)	3	2		60	40	100	3
Total Credits									**18**

L=Lecture, T/S=Tutorials/Seminar, P/I=Practical/Internship, CE =Continuous Evaluation, ESE = End Semester Evaluation

Sl No	**POOL H:- List of Courses for Elective (Practical)** **DISCIPLINE SPECIFIC ELECTIVES (DSE)**								
4.2.1	MSBST04DSE13	Biostatistical Computing Using SAS-III (**Practical**)		2	6	60	40	100	3
4.2.2	MSBST04DSE14	Biostatistical Computing Using R-III (**Practical**)		2	6	60	40	100	3
POOL I:- List of Courses for Elective **DISCIPLINE SPECIFIC ELECTIVES (DSE)**									
4.3.1	MSBST04DSE15	Categorical Data Analysis	4	1					3
4.3.2	MSBST04DSE16	Advanced Time Series Analysis	4	1					3
4.3.3	MSBST04DSE17	Actuarial Statistics	4	1					3
4.3.4	MSBST04DSE18	Statistical Quality Control	4	1					3

4.3.5	MSBST04DSE19	Advanced Bayesian Computing with R	4	1					3
4.3.6	MSBST04DSE20	Demographic Studies	4	1					3
4.3.7	MSBST04DSE21	Analysis of Longitudinal Data	4	1					3

POOL H: DISCIPLINE SPECIFIC ELECTIVE COURSE

Course Code & Title	**MSBST04DSE13-BIOSTATISTICAL COMPUTING USING SAS - III (PRACTICAL)**		
Programme	**M.Sc. Biostatistics**	**Semester**	IV
Course Objectives	• To introduce some advanced statistical computing techniques in biostatistics to extract information and visualization thereby enabling them to perform data analysis effectively and efficiently in SAS programming. • Illustrate different statistical techniques based on all the elective course in fourth semester.		

Modules	Content	Module Outcome
	Biostatistical Computing IV is a practical course. The practical is based on all the elective courses in the fourth semester.	• Describe different statistical technique to solve problems coming under all the elective courses in fourth semester.
Course Outcomes	After successful completion of this course, student will be able to: 1. Equipped with different theoretical methods in biostatistics to achieve the objectives. 2. Enhanced with the basic concepts of statistical theories besides developing their ability to handle real world problems with large scale data.	

TEACHING LEARNING STRATEGIES

- **Practical sessions through computers, statistical computations, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Hands on training**

ASSESSMENT RUBRICS

Components	Weightage
End Semester Evaluation	**60**
Continuous Evaluation	
Practical Tests	32
Record	8
Total	**40**

POOL H: DISCIPLINE SPECIFIC ELECTIVE COURSE

Course Code & Title	**MSBST04DSE14-BIOSTATISTICAL COMPUTING USING R - III (PRACTICAL)**		
Programme	**M.Sc. Biostatistics**	**Semester**	IV
Course Objectives	• To introduce some advanced biostatistical computing techniques in applied statistics to extract information and visualization thereby enabling them to perform data analysis effectively and efficiently in R programming. • Illustrate different biostatistical techniques based on all the elective course in fourth semester.		

Modules	Content	Module Outcome
	Biostatistical Computing IV is a practical course. The practical is based on all the elective courses in the fourth semester.	• Describe different statistical technique to solve problems coming under all the elective courses in fourth semester.
Course	After successful completion of this course, student will be able to:	

Outcomes	1. Equipped with different theoretical methods in biostatistics to achieve the objectives. 2. Enhanced with the basic concepts of biostatistical theories besides developing their ability to handle real world problems with large scale data.

TEACHING LEARNING STRATEGIES

- **Practical sessions through computers, statistical computations, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Hands on training**

ASSESSMENT RUBRICS

Components	**Weightage**
End Semester Evaluation	**60**
Continuous Evaluation	
Practical Tests	32
Record	08
Total	**40**

POOL I: DISCIPLINE SPECIFIC ELECTIVE COURSE

Course Code & Title	**MSBST04DSE15- CATEGORICAL DATA ANALYSIS**	
Programme	**M.Sc. Biostatistics**	**Semester IV**
Course Objectives	• Understand the fundamentals of statistical measurement scales and their implications in research design and analysis. • Study various sampling frameworks and techniques, including random sampling, stratified sampling, and cluster sampling, and apply them appropriately in research settings. • Learn about exact tests and their significance in situations where traditional statistical tests may not be applicable.	

Modules	Content	Module Outcome
Module I Overview of analysis strategies **(15 Hours)**	Scale of Measurements, sampling frameworks, overview of analysis strategies, Chi-square statistic, exact tests, difference in proportions, odds ratio and relative risk, sensitivity and specificity, McNemar's test.	• Understand the differences between nominal, ordinal, interval, and ratio scales of measurement. • Apply appropriate measurement scales in research design and data analysis. • Identify the implications of different measurement scales on statistical analysis techniques.
Module II Measure of association and contingency tables **(15 Hours)**	Mantel-Haenszel test, Measure of association, sets of 2 x r tables, sets of s x 2 tables, relationship between sets of tables.	• Demonstrate proficiency in selecting and implementing appropriate sampling frameworks for different research scenarios. • Understand the strengths and limitations of various sampling methods.
Module III Mantel-Haenszel methodology and application **(15 Hours)**	Association, exact tests for association, Measure of association, observer agreement, test for ordered differences, General Mantel-Haenszel methodology, Mantel- Haenszel applications.	• Gain familiarity with descriptive statistics, hypothesis testing, and regression analysis. • Identify appropriate analysis strategies based on research objectives and data characteristics. • Evaluate the assumptions and validity of different analysis techniques.
Module IV Advanced topics	Advanced topics: application to repeated measures, Wilcoxon-Mann-Whiteney test, Kruskal-Wallis test,	• Navigate and manipulate data tables efficiently.

(15 Hours)	Friedman's Chi-square test, Aligned rank test for randomised complete blocks, Durbin's test for balanced incomplete blocks, Rank analysis of covariance.	• Perform basic to advanced data transformations. • Generate informative tables and graphical displays to summarize data effectively.
References	***Text Books*** 1. Stokes, M. E., Davis, C. S., & Koch, G. G. (2012). *Categorical data analysis using SAS*. SAS institute. 2. Agresti, A. (2012). *Categorical data analysis* (Vol. 792). John Wiley & Sons. ***Reference Books*** 1. Powers, D., & Xie, Y. (2008). *Statistical methods for categorical data analysis*. Emerald Group Publishing. 2. Sloane, D., & Morgan, S. P. (1996). An introduction to categorical data analysis. *Annual review of sociology*, *22*(1), 351-375. 3. Lawal, B., & Lawal, H. B. (2003). *Categorical data analysis with SAS and SPSS applications*. Psychology Press.	
Course Outcomes	After successful completion of this course, student will be able to: 1. Demonstrate a thorough understanding of various statistical methods. 2. Proficiency in working with data tables, conducting data manipulation, and performing statistical analysis, enhancing their ability to manage and analyze large datasets in real-world research settings. 3. Cultivate critical thinking skills by critically evaluating research designs, selecting appropriate sampling frameworks. 4. Communicate statistical findings clearly and effectively through written reports, presentations, and graphical representations.	

TEACHING LEARNING STRATEGIES

- **Lecturing, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	Weightage
End Semester Evaluation	**60**
Continuous Evaluation	
Tests	16
Assignment	08
Seminar	16
Total	**40**

Sample Questions to test Outcomes:

1. Explain the different scales of measurements and how they impact statistical analysis. Provide examples for each scale.
2. Discuss the importance of sampling frameworks in research design. How do different sampling techniques affect the validity of study findings?
3. Give an overview of analysis strategies commonly used in research. Compare and contrast their strengths and limitations.
4. Demonstrate how to work with tables in the SAS system for data analysis and interpretation.
5. Explain the Chi-square statistic and its significance in hypothesis testing. Provide a real-world example illustrating its application.
6. Discuss the concept of odds ratio and relative risk in epidemiological studies. How are they calculated and interpreted?
7. Define sensitivity and specificity in diagnostic testing. How do these measures inform the accuracy of a diagnostic test?
8. Explain McNemar's test and its application in paired nominal data analysis. Provide a step-by-step example to illustrate its usage.
9. Describe the Mantel-Haenszel test and its role in analyzing categorical data. Provide examples of situations where this test would be appropriate.

10. Discuss advanced topics such as repeated measures analysis and non-parametric tests like Wilcoxon-Mann-Whitney, Kruskal-Wallis, and Friedman's Chi-square tests.

POOL I: DISCIPLINE SPECIFIC ELECTIVE COURSE

Course Code & Title	**MSBST04DSE16-ADVANCED TIME SERIES ANALYSIS**		
Programme	**M.Sc. Biostatistics**	**Semester**	IV
Course Objectives	• Understand the fundamental concepts of stochastic processes and their applications in modeling time series data, providing a solid foundation for further analysis. • Learn to analyze auto-covariance, auto-correlation, and spectral density properties of time series data, enabling students to characterize and interpret temporal patterns effectively. • Gain in-depth knowledge of autoregressive (AR), moving average (MA), autoregressive moving average (ARMA), and autoregressive integrated moving average (ARIMA) models, enabling students to select and apply appropriate models for different time series data sets. • Explore spectral analysis, periodgrams, correlograms, and diagnostic checks for model validation, equipping students with advanced analytical tools for interpreting and analyzing time series data effectively.		

Modules	Content	Module Outcome
Module I Revisit to foundations of time series **(15 Hours)**	Motivation, Time series as a discrete parameter stochastic process, Auto-Covariance, Auto-Correlation and spectral density and their properties. Exploratory time series analysis, Exponential and moving average smoothing, Holt-Winter smoothing, forecasting based on smoothing,	• Understanding Stochastic Processes. • Applying Stochastic Processes to Time Series Analysis. • Interpreting Stochastic Process Properties. • Applying Stochastic Processes in Forecasting:

	Adaptive smoothing.	
Module II Detailed study of the stationary ARMA models **(15 Hours)**	Detailed study of the stationary process: Autoregressive, Moving Average, Autoregressive Moving Average and Autoregressive Integrated Moving Average Models. Choice of AR/MA periods.	• Understanding exploratory Analysis Techniques. • Understanding Smoothing Methods. • Characterizing Stationary Processes. • Interpreting Time Series Properties.
Module III Estimation and forecasting of ARIMA models **(15 Hours)**	Estimation of ARMA models: Yule-Walker estimation for AR Processes, Maximum likelihood and least square estimation for ARMA Processes, Discussion (without proof) of estimation of mean, Auto- covariance and autocorrelation function under large samples theory, Residual analysis and diagnostic checking. Forecasting using ARIMA models.	• Understanding forecasting Techniques. • Identifying model Selection and Estimation. • Validation and Diagnostic Checks. • Utilize advanced analytical tools such as spectral analysis, periodograms, and correlograms to analyze and interpret time series data effectively, enhancing their ability to derive insights and make informed decisions.
Module IV Spectral and seasonal analysis **(15 Hours)**	Spectral analysis of weakly stationary process. Periodogram and correlogram analysis. Seasonal ARIMA models (Basic concepts only), ARCH and GARCH models (Basic concepts only)	• Understand seasonal and non-seasonal models such as seasonal ARIMA, ARCH, and GARCH models, enabling them to capture and forecast complex temporal patterns within data sets. • Apply advanced analytical techniques including spectral

		analysis, periodograms, and correlograms to analyze and interpret time series data effectively.
References	**Text Books** 1. Box G.E.P and Jenkins G.M. (1970).Time Series Analysis, Forecasting and Control. Holden -Day. 2. Brockwell P.J .and Davis R.A. (1987). Time Series: Theory and Methods, Springer Verlag. **Reference Books** 1. Abraham B and Ledolter J.C. (1983). Statistical Methods for Forecasting, Wiely 2. Anderson T.W. (1971). Statistical Analysis of Time Series, Wiely. 3. Fuller W.A. (1978). Introduction to Statistical Time Series, John Wiley.	
Course Outcomes	After successful completion of this course, student will be able to: 1. Understand exploratory time series analysis and its real data application. 2. Understand autoregressive models and their estimation methods. 3. Understand non-linear time series models and their estimation methods. 4. Apply statistical techniques to time series data and make predictions.	

TEACHING LEARNING STRATEGIES

- **Lecturing, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	Weightage
End Semester Evaluation(ESE)	**60**

Continuous Evaluation	
Tests	16
Assignment	08
Seminar/Viva	16
Total	**40**

Sample Questions to Test Outcomes:

1. What is the significance of time series analysis in understanding data trends and making predictions? Discuss the motivation behind studying time series data.
2. Define a discrete parameter stochastic process and explain how time series can be viewed as such a process. How does this perspective help in analyzing time series data?
3. Explain the concepts of auto-covariance, auto-correlation, and spectral density in the context of time series analysis. Discuss their properties and their roles in characterizing time series behavior.
4. Describe the techniques involved in exploratory time series analysis. How do tests for trend and seasonality contribute to understanding time series patterns?
5. Compare and contrast exponential smoothing, moving average smoothing, and Holt-Winter smoothing methods for time series forecasting. Provide examples illustrating their application.
6. Provide a detailed study of stationary processes, including autoregressive (AR), moving average (MA), autoregressive moving average (ARMA), and autoregressive integrated moving average (ARIMA) models. Discuss the choice of AR/MA periods in model selection.
7. Discuss different estimation methods for ARMA models, including Yule-Walker estimation for AR processes and maximum likelihood/least square estimation for ARMA processes. Explain the concept of residual analysis and diagnostic checking.
8. Explain the concept of spectral analysis for weakly stationary processes. Discuss the use of periodograms and correlograms in analyzing spectral density and autocorrelation functions.

9. Introduce the basic concepts of seasonal ARIMA models, ARCH models, and GARCH models in time series analysis. Discuss their relevance and applications in modeling time series data.

POOL I: DISCIPLINE SPECIFIC ELECTIVE COURSE

Course Code & Title	**MSBST04DSE17- ACTUARIAL STATISTICS**		
Programme	**M.Sc. Biostatistics**	**Semester**	IV
Course Objectives	• Develop a greater understanding of statistical principles and their application in actuarial statistics. • Describe the core areas of actuarial practice and relate to those areas actuarial principles, theories and models. • Describe estimation procedures for lifetime distributions. • Explain the concept of survival models. • Understand the application of knowledge of the life insurance environment. • Describe Net premiums and its various types. • Expand their applied knowledge in various specialized areas of actuarial studies and statistics.		

Modules	Content	Module Outcome
Module I Insurance and utility theory **(15 Hours)**	Insurance and utility theory, models for individual claims and their sums, survival function, curtate future lifetime, force of mortality. Life tables and its relation with survival function, examples, assumptions for fractional ages, some analytical laws of mortality, select and ultimate tables.	• Explains the utility theory and insurance. • Explain survival function and application. • Examine the properties of force of mortality. • Define Life tables and its relation with survival function, examples.
Module II	Multiple life functions, joint life and	• Explain Multiple life functions

Multiple life functions **(15 Hours)**	last survivor status, insurance and annuity benefits through multiple life functions evaluation for special mortality laws. Multiple decrement tables, central rates of multiples decrement, net single premiums and their numerical evaluations.	and its properties. • Articulate the insurance and annuity benefits through multiple life functions evaluation for special mortality laws. • Explains the Multiple decrement tables. • Describe net single premiums and their numerical evaluations.
Module III Compound Poisson distribution and its applications **(15 Hours)**	Compound Poisson distribution and its applications. Principles of compound interest: Nominal and effective rates of interest and discount, force of interest and discount, compound interest, accumulation factor, continuous compounding.	• Define Distribution of aggregate claims. • Derive the compound Poisson distribution and explain its applications. • Explain Principles of compound interest and its attributes.
Module IV Different types of insurance and amenities **(15 Hours)**	Insurance payable at the moment of death and at the end of the year of death-level benefit insurance, endowment insurance, differed insurance and varying benefit insurance, recursions, commutation functions. Life annuities: Single payment, continuous life annuities, discrete life annuities, life annuities with monthly payments, commutation functions, varying annuities, recursions, complete annuities immediate and apportion able annuities-due.	• Explain the Life insurance and its types. • Describe Insurance payable at the moment of death and at the end of the year of death-level benefit insurance • Explain the Life annuities and its types.

References	***Text Books*** 1. Beard, R.E., Penlikainen, T. and Pesonnen, E (1984): Risk Theory: The Stochastic Basis of Insurance, 3rd Edition, Chapman and Hall, Londan. ***Reference Books*** 1. Bowers, N.L., Gerber, H.U., Hickman, J.E., Jones, D.A. and Nesbitt, C.J. (1997): Actuarial Mathematics', Society of Actuarias, Ithaca, Illiois, U.S.A., second Edition. 2. Neill, A. (1977): Life Contingencies, Heineman.
Course Outcomes	After successful completion of this course, student will be able to: 1. Understand the principles of insurance and utility theory, and apply them to analyze individual claims and their aggregate sums. 2. Analyze the relationship between life tables and survival functions, and apply this knowledge to evaluate insurance and annuity benefits. 3. Demonstrate proficiency in using multiple decrement tables to assess insurance risks. 4. Apply the Compound Poisson distribution and principles of compound interest to model insurance-related phenomena. 5. Develop skills in evaluating and designing life annuities, including single payment, continuous, discrete, and monthly payment annuities.

TEACHING LEARNING STRATEGIES

- **Lecturing, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	Weightage
End Semester Evaluation(ESE)	**60**
Continuous Evaluation	

Tests	16
Assignment	08
Seminar/Viva	16
Total	**40**

Sample Questions to Test Outcomes:

1. Define insurance and utility theory. How does utility theory influence decision-making in insurance?
2. Discuss models for individual claims and their sums in the context of insurance. How are these models used to assess risk and determine premiums?
3. Explain the concepts of survival function, curtate future lifetime, and force of mortality in the context of life insurance. How are these concepts related?
4. Describe life tables and their relationship with the survival function. Provide examples to illustrate the use of life tables in actuarial calculations.
5. What assumptions are made for fractional ages in life tables? How do these assumptions impact the accuracy of actuarial calculations?
6. Discuss some analytical laws of mortality commonly used in actuarial science. What are select and ultimate tables, and how are they applied?
7. Explain multiple life functions and their significance in joint life and last survivor status insurance policies. How are insurance and annuity benefits evaluated using multiple life functions?
8. Define multiple decrement tables and central rates of multiple decrements. How are net single premiums calculated using multiple decrement tables?
9. Describe the compound Poisson distribution and its applications in insurance. How is it used to model the frequency and severity of insurance claims?
10. Discuss the principles of compound interest, including nominal and effective rates of interest, force of interest and discount, and continuous compounding. How are these principles applied in actuarial calculations for insurance products like endowment and annuity policies?

POOL I: DISCIPLINE SPECIFIC ELECTIVE COURSE

Course Code & Title	**MSBST04DSE18-STATISTICAL QUALITY CONTROL**		
Programme	**M.Sc. Biostatistics**	**Semester**	IV
Course Objectives	• Understand the principles and theory behind statistical process control (SPC) and the significance of control charts in quality management. • Develop proficiency in constructing and interpreting Shewhart control charts for variables (R, s charts) and attributes (p, np, c, u charts), as well as modified control charts. • Gain insight into the performance of control charts through operating characteristic (OC) and average run length (ARL) curves. • Learn about advanced control chart techniques, including moving average control charts, exponentially weighted moving average (EWMA) charts, and cumulative sum (CUSUM) charts. • Master various sampling plans such as single sampling, double sampling, multiple sampling, and sequential sampling plans, including rectifying inspection plans. • Analyze the performance of sampling plans using metrics like OC, average outgoing quality (AOQ), average sample number (ASN), and average total inspection (ATI) curves. • Understand and apply sampling plans for single specification limits with known and unknown variance, as well as sampling plans with double specification limits.		

Modules	Content	Module Outcome
Module I Introduction to quality and quality	Introduction to quality and quality assurance, total quality management, quality control, Statistical process control, theory of control charts, Shewhart control	• Understand the theory behind control charts and their role in SPC. • Demonstrate proficiency in constructing and interpreting

assurance **(15 Hours)**	charts for variables, R,s charts, p, np, c, u charts, modified control charts.	Shewhart control charts for variables (R, s charts) and attributes (p, np, c, u charts). • Explore modified control charts and their applications in various industries.
Module II Control charts and process capability indices **(15 Hours)**	O.C and ARL curves of control charts, moving average control charts, EWMA charts, CUSUM charts, process capability analysis, process capability indices.	• Analyze operating characteristic (OC) and average run length (ARL) curves of control charts. • Apply moving average control charts, exponentially weighted moving average (EWMA) charts, and cumulative sum (CUSUM) charts to monitor process variability. • Conduct process capability analysis and calculate process capability indices to assess the performance of a process.
Module III Sampling Plans **(15 Hours)**	Single sampling, double sampling, multiple sampling and sequential sampling plans, rectifying inspection plans, measuring performance of the sampling plans - OC,AOQ,ASN, ATI curves.	• Differentiate between single sampling, double sampling, multiple sampling, and sequential sampling plans. • Develop rectifying inspection plans to improve quality control processes. • Measure the performance of sampling plans using metrics such as OC, average outgoing quality (AOQ), average sample number (ASN), and average total inspection (ATI) curves.

Module IV Sampling plans with double specification limits **(15 Hours)**	Sampling plans for single specification limit with known and unknown and unknown variance. Sampling plans with double specification limits, comparison of sampling plans by variables and attributes, Continuous sampling plans I, II and III.	• Design sampling plans for single specification limits with known and unknown variance. • Implement sampling plans with double specification limits and compare them based on variables and attributes. • Explore continuous sampling plans (I, II, and III) and understand their applications in industries with continuous production processes.
References	**Text Books** 1. Montgomory,D.C.(2005),IntroductiontoStatisticalQualityControl.5thEdition.Wiley,New-York. 2. Gerant,E.L.andLeavenWorth,R.S.(1980).StatisticalQualityControl.McGrawHill **Reference Books** 1. Duncan, A.J.(1986).Quality Control and Industrial Statistics. 2. Mittage, H.J. and Rinne, H. (1993). Statistical Methods for Quality Assurance. Chapmanand Hall. 3. Oakland,J.S.and Follorwel, R.F.(1990).Statistical Process Control. East-West Press. 4. Schilling,E.G.(1982).AcceptanceSamplinginQualityControl.MarcelDekker.	
Course Outcomes	After successful completion of this course, student will be able to: 1. Understand the construction various control charts and their real data applications. 2. Understand various process capability indices and their applications. 3. Understand various different acceptance sampling plans for attributes and variables.	

TEACHING LEARNING STRATEGIES

- **Lecturing, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	Weightage
End Semester Evaluation(ESE)	**60**
Continuous Evaluation	
Tests	16
Assignment	08
Seminar/Viva	16
Total	**40**

Sample Questions to Test Outcomes:

1. Define Statistical Process Control (SPC) and discuss its importance in quality management. What are the key objectives of implementing SPC in manufacturing processes?
2. Explain the theory of control charts in SPC. What are the fundamental principles behind control charts, and how do they help in monitoring and controlling process variability?
3. Describe Shewhart control charts for variables, including R charts and s charts. How are these charts constructed, and what do they indicate about process stability and variation?
4. Discuss p, np, c, and u charts in SPC. What types of processes are these charts suitable for, and how are they interpreted in terms of process control?
5. Explain modified control charts and their applications in SPC. What modifications can be made to traditional control charts, and under what circumstances are these modifications necessary?
6. Define Operating Characteristic (OC) and Average Run Length (ARL) curves of control charts. How are these curves used to evaluate the performance of control charts?

7. Describe moving average control charts, exponentially weighted moving average (EWMA) charts, and cumulative sum (CUSUM) charts. What advantages do these charts offer over traditional control charts?
8. Explain process capability analysis and process capability indices. How are these measures used to assess the ability of a process to meet specified quality requirements?
9. Discuss single sampling, double sampling, multiple sampling, and sequential sampling plans. What are the differences between these sampling plans, and how are they applied in quality inspection?
10. Describe sampling plans for single specification limits and double specification limits. How do these sampling plans vary based on known and unknown variance, and how are they evaluated using OC, AOQ, ASN, and ATI curves?

POOL I: DISCIPLINE SPECIFIC ELECTIVE COURSE

Course Code & Title	**MSBST04DSE19- ADVANCED BAYESIAN COMPUTING WITH R**		
Programme	**M.Sc. Biostatistics**	**Semester**	**IV**
Course Objectives	• Understand the fundamental principles of statistical decision-making, including randomized decision rules and standard loss functions. • Explore the concept of prior information and its incorporation into decision-making. • Master the application of Bayes' theorem for inference, including the estimation of prior and posterior densities, and gain proficiency in analyzing parametric families and likelihoods, such as the exponential family. • Apply learned concepts and techniques to practical examples and real-world problems, using software packages such as Learn Bayes and Win-BUGS.		

Modules	**Content**	**Module Outcome**

Module I Fundamentals of Bayesian concepts **(15 Hours)**	Statistical decision problem, randomized decision rule, decision principle, standard loss functions, Prior information, subjective determination of prior density, non-informative priors, maximum entropy priors, conjugate priors, discrete prior. Parametric family and likelihood, exponential family, Bayes' theorem for inference, prior and posterior densities.	• Understand and explain the concept of statistical decision-making. • Apply decision principles effectively in various decision-making scenarios. • Evaluate and compare different decision strategies based on their performance in minimizing expected loss and achieving desired outcomes.
Module II Bayes models and Learn Bayes package **(15 Hours)**	single parameter models, normal distribution with known variance and unknown mean, normal with known mean and unknown variance, Poisson model, normal distribution with both parameters unknown, multinomial model, Dirichlet prior, Bioassay experiment, comparing two proportions, predictive distribution, beta-binomial distribution, multivariate normal distribution, Introduction to Learn Bayes package, Examples using Learn Bayes package.	• Demonstrate proficiency in formulating and specifying prior. • Apply Bayes' theorem for inference tasks and analyze the impact of prior specification on posterior inference. • Evaluate the suitability of different prior distributions for specific modeling scenarios.
Module III Introduction to Markov Chain Monte Carlo methods **(15 Hours)**	Computing integrals using Monte-Carlo simulation, approximation based on posterior mode, importance sampling, Markov Chain Monte Carlo methods, Metropolis-Hastings algorithm, random walk, Gibbs sampling.	• Implement computational methods such as Monte Carlo simulation, importance sampling, and Markov Chain Monte Carlo (MCMC). • Evaluate the performance and efficiency of different

		computational techniques in generating posterior samples and estimating posterior distributions.
Module IV Hierarchical models **(15 Hours)**	Hierarchical models, shrinkage estimators, posterior predictive model checking, comparison of hypotheses, Bayes factor, one sided test for normal mean, two-sided test for normal mean, normal linear regression model, prediction of future observations, examples and R codes, introduction to Win-BUGS package.	• Understand the concept of hierarchical models and their application in modeling complex data. • Perform posterior predictive model checking to assess the adequacy of hierarchical models and identify potential model misspecifications.
References	**Text Books** 1. Jim Albert (2007). Bayesian Computation with R, New York: Springer Verlag. 2. Berger,O.J.(1985).StatisticaldecisionTheoryandBayesianAnalysis,SecondEdition,SpringerVerlag. 3. Bensal, A. K.(2008).Bayesian Parametric Inference, New Age, Delhi. **Reference Books** 1. Ferguson,T.S.(1967).MathematicalStatistics:ADecisionTheoreticApproach,AcademicPress,New-York. 2. Bolstad, W.(2004). Introduction to Bayesian Statistics, Hoboken, NJ: John Wiley. 3. Gelman,A.,Carlin,J.,Stern,H.andRubin,D.(2003).BayesianDataAnalysis,NewYork:ChapmanandHall. 4. Gilks, W. R., Richardson, S and Spiegelhalter, D.J.(1996). Markov Chain Monte Carlo in Practice. Chapman & Hall/ CRC, New York.	
Course Outcomes	After successful completion of this course, student will be able to: 1. Understand the advantageous Bayes estimation over that based on frequentist approach. 2. Understand the LearnBayes package for various Bayesian	

	computations 3. Understand MCMC methods in various situations in which the exact computation is difficult. 4. Understand Gibbs sampling to generate random samples from a multivariate distribution.

TEACHING LEARNING STRATEGIES

- **Lecturing, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	Weightage
End Semester Evaluation(ESE)	**60**
Continuous Evaluation	
Tests	16
Assignment	08
Seminar/Viva	16
Total	**40**

Sample Questions to Test Outcomes:

1. Define a randomized decision rule and explain its significance in statistical decision-making.

2. Discuss the role of prior information in Bayesian inference and explain how it is incorporated into the decision-making process.

3. Describe the concept of maximum entropy priors and explain when they are useful in Bayesian analysis.

4. Compare and contrast Monte Carlo simulation and importance sampling methods for estimating posterior distributions.

5. Explain the Metropolis-Hastings algorithm and discuss its advantages and limitations in Markov Chain Monte Carlo (MCMC) sampling.

6. Provide an example of a hierarchical model and explain how it can be used to analyze data with nested levels of variability.

7. Discuss the concept of shrinkage estimators and explain how they address overfitting in hierarchical modeling.

8. Explain the process of posterior predictive model checking and discuss its importance in assessing model adequacy.

9. Compute the Bayes factor for two competing hypotheses and interpret the results in the context of model comparison.

10. Implement the Win-BUGS software package to perform Bayesian analysis on a given dataset, and interpret the results obtained.

POOL I: DISCIPLINE SPECIFIC ELECTIVE COURSE

Course Code & Title	**MSBST04DSE20- DEMOGRAPHIC STUDIES**		
Programme	**M.Sc. Biostatistics**	**Semester**	**IV**
Course Objectives	• To introduce students to key concepts and theories in demography. • To provide students with an understanding of demographic data sources and measurement techniques. • To familiarize students with demographic methods for analyzing population dynamics. • To demonstrate the applications of demographic analysis in various fields, including public health, economics, and social policy.		

Modules	Content	Module Outcome

Module I Scope of demography **(15 Hours)**	Definition and scope of demography, Historical development of demography, Importance and applications of demographic research, Population Size and Composition, Measures of population size: population counts, estimates, and projections, Population composition: age structure, sex ratio, and demographic characteristics, Interpretation of population pyramids	• Gain a comprehensive understanding of demographic concepts such as population size, composition, fertility, mortality, migration, and population aging. • Learn to interpret and analyze demographic measures including birth rates, death rates, life expectancy, migration flows, and population pyramids.
Module II Measures of mortality and fertility **(15 Hours)**	Fertility: Measures of fertility: birth rates, total fertility rates, age-specific fertility rates, Determinants of fertility: socioeconomic, cultural, and policy factors, Trends and patterns in fertility, Mortality: Measures of mortality: death rates, life expectancy, age-specific mortality rates, Causes of mortality: infectious diseases, chronic diseases, external causes, Epidemiological transition theory	• Acquire proficiency in identifying and utilizing various sources of demographic data, including censuses, surveys, and vital registration systems. • Learn data quality assessment techniques and sampling methods for demographic research
Module III Theories of migration **(15 Hours)**	Migration: Types of migration: internal, international, refugee movements, Measures of migration: net migration rates, migration flows, migration stocks, Theories of migration: push and pull factors, network theory, migration systems Population Aging, Concepts and measures of population aging,	• Develop skills in demographic analysis techniques, including standardization methods and demographic modeling approaches such as cohort-component projection and survival analysis. • Apply statistical methods to analyze demographic trends

	Causes and consequences of population aging, Challenges and opportunities of an aging population	and patterns, and interpret findings accurately.
Module IV Demographic Analysis Techniques **(15 Hours)**	Demographic Data Sources and Methods; Sources of demographic data: censuses, surveys, vital registration systems, Data quality issues and sampling techniques Demographic Analysis Techniques Standardization techniques, Demographic modeling: cohort-component projection method, population momentum, Survival analysis techniques for life tables and mortality data, Population Policy and Planning.	• Explore historical trends in population dynamics and understand the underlying causes and consequences of demographic changes. • Analyze contemporary population trends, including fertility, mortality, migration, and population aging, and assess their implications for society.
References	**Text Book** 1. Poston Jr, D. L., & Bouvier, L. F. (2010). *Population and society: An introduction to demography*. Cambridge University Press. 2. Cox PR (1957). *Demography*. Cambridge University Press **Reference Book** 1. Croxton F E and Crowder D J (1967) *Applied General statistics*, Prentice - Hall India. 2.Bogue, Donald J: Principles of Demography, John Wiley and Sons, New York,1969 3. Shrivastava O S: A Text Book of Demography with Economics of Man Power Supply and Manpower Demand, Vikas, New Delhi,1983	
Course Outcomes	After successful completion of this course, student will be able to: 1. Define and explain fundamental demographic concepts such as population size, composition, fertility, mortality, migration, and population aging. 2. Identify and utilize various sources of demographic data, including censuses, surveys, and vital registration systems, and assess data quality.	

	3. Apply demographic analysis techniques, including standardization methods and demographic modeling approaches, to analyze population dynamics and trends. 4. Interpret demographic measures and trends accurately, and assess their implications for social, economic, and political contexts.

TEACHING LEARNING STRATEGIES

- **Lecturing, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	Weightage
End Semester Evaluation(ESE)	**60**
Continuous Evaluation	
Tests	16
Assignment	08
Seminar/Viva	16
Total	**40**

Sample Questions to Test Outcomes:

1. Define demography and explain its significance in the field of statistics.
2. Describe the measures used to quantify population size and composition. How are these measures calculated?
3. Discuss the factors influencing fertility rates and their variations across different populations.
4. Compare and contrast crude birth rate, total fertility rate, and age-specific fertility rates. Provide examples illustrating their use.
5. Explain the concept of the epidemiological transition and its implications for population health.

6. Discuss the main determinants of mortality rates and how they have changed over time.
7. Describe the different types of migration and their impact on population dynamics. Provide examples of push and pull factors influencing migration.
8. Define population aging and discuss its causes and consequences for societies.
9. Explain the cohort-component projection method and its application in population forecasting.
10. Discuss the role of demography in informing public policy decisions related to healthcare, labor markets, and environmental planning.

POOL I: DISCIPLINE SPECIFIC ELECTIVE COURSE

Course Code & Title	**MSBST04DSE21- ANALYSIS OF LONGITUDINAL DATA**		
Programme	**M.Sc. Biostatistics**	**Semester**	**IV**
Course Objectives	• Master advanced statistical modeling techniques for longitudinal data. • Develop proficiency in estimating model parameters using maximum likelihood (ML), restricted maximum likelihood (REML). • Gain a comprehensive understanding of missing data mechanisms and strategies for handling missing values in longitudinal studies. • Learn to address challenges such as time-dependent covariates, intermittent missing values, and dropout processes in longitudinal data analysis.		

Modules	Content	Module Outcome
Module I General Linear Model for Longitudinal	General Linear Model for Longitudinal Data. ML and REML estimation, EM algorithm: General linear mixed-effects model, Inference for; the random	• Explore maximum likelihood (ML) and restricted maximum likelihood (REML) estimation methods. • Understand the expectation-

Data **(15 Hours)**	effects, BLUPs, Empirical Bayes, Bayes, Shrinkage Model building and diagnostic, relaxing parametric assumptions: generalized additive mixed model.	maximization (EM) algorithm. • Perform inference for random effects.
Module II Random effects models for binary and count data **(15 Hours)**	Generalized Linear Model for Longitudinal Data, Marginal models, for binary, ordinal, and count data: Random effects models for binary and count data: Transition models: Likelihood-based models for categorical data; GEE; Models for mixed discrete and continuous responses.	• Extend the framework to handle generalized linear models (GLMs) for longitudinal data. • Study marginal models for binary, ordinal, and count data, as well as random effects models. • Explore transition models and likelihood-based approaches for categorical data
Module III Modeling the dropout process **(15 Hours)**	Classification missing data mechanism; Intermittent missing values and dropouts; Weighted estimating equations; Modeling the dropout process (Selection and pattern mixture models).	• Investigate classification of missing data mechanisms and strategies for addressing intermittent missing values and dropouts. • Learn about weighted estimating equations and modeling the dropout process.
Module IV Multivariate longitudinal data **(15 Hours)**	Dangers of time dependent covariates, Lagged covariates; Marginal Structural models; Joint models for longitudinal and survival data; Multivariate longitudinal data; Design of randomized and observational longitudinal studies.	• Address challenges associated with time-dependent covariates and lagged covariates. • Explore marginal structural models and joint models for longitudinal data. • Discuss strategies for analyzing multivariate

		longitudinal data.
References	**Text books** 1. Diggle, P. J., Heagerty, P., Liang, K. Yand Zeger.S. L (2003). *Analysis of Longitudinal Data, 2nd Edn.* Oxford University press, New York. 2. Fitzmaurice, G.M., Laird, N.M and Ware, J.H. (2004).*Applied Longitudinal Analysis*, John Wiley & Sons, New York. **Reference Books** 1. Crowder, M. J. and Hand, D. J. (1990). *Analysis of Repeated Measures*. Chapman and Hall/CRC Press, London. 2. Davidian, M. and Giltinan, D. M. (1995). *Nonlinear Models for Repeated Measurement Data.* Chapmanand Hall/CRC Press, London. 3. Hand, D and Crowder, M. (1996). *Practical Longitudinal Data Analysis.* Chapman and Hall/CRC Press, New York. 4. Little, R. J. A and Rubin, O. B.(2002). *Statistical Analysis with Missing Data, 2nd Edition,* Wiley, New York. 5. McCullagh, P. and Nelder. J. A. (1989*). Generalized Linear Models. 2nd Edition,* Chapman and Hall/CRC Press, London. 6. Weiss, R. E. (2005*). Modeling Longitudinal Data. Springer*, New York	
Course Outcomes	After successful completion of this course, student will be able to: 1. Conduct analysis of longitudinal data. 2. Apply statistical techniques to model longitudinal data and make predictions. 3. Understand analysis of longitudinal data with missing data. 4. Understand analysis of longitudinal data with time-dependent covariates.	

TEACHING LEARNING STRATEGIES

- **Lecturing, Visualization, Team Learning**

MODE OF TRANSACTION

- **Lecture, Seminar, Discussion, Questioning and Answering**

ASSESSMENT RUBRICS

Components	Weightage
End Semester Evaluation(ESE)	**60**
Continuous Evaluation	
Tests	16
Assignment	08
Seminar/Viva	16
Total	**40**

Sample Questions to Test Outcomes:

1. What distinguishes maximum likelihood (ML) from restricted maximum likelihood (REML) estimation in the context of longitudinal data analysis?
2. Explain the concept of the expectation-maximization (EM) algorithm and how it is utilized in fitting general linear mixed-effects models for longitudinal data.
3. How would you handle intermittent missing values in a longitudinal dataset? Describe the strategies and techniques you would employ.
4. Discuss the advantages and limitations of using generalized estimating equations (GEE) for analyzing longitudinal data compared to mixed-effects models.
5. What are the key differences between marginal models and random effects models for handling binary data in longitudinal studies?
6. How do you assess the impact of time-dependent covariates on longitudinal outcomes? Describe the statistical methods used for this analysis.
7. Explain the concept of best linear unbiased predictions (BLUPs) and their relevance in estimating random effects in longitudinal models.
8. Describe the process of building a generalized additive mixed model (GAMM) for longitudinal data and discuss its advantages over parametric approaches.
9. What are selection models and pattern mixture models, and how are they used to address the issue of dropout in longitudinal studies?

10. Can you outline the steps involved in designing a longitudinal study, including considerations for handling missing data and analyzing multivariate longitudinal outcomes?