JAWAHARLAL INSTITUTE OF POSTGRADUATE MEDICAL EDUCATION & RESEARCH (JIPMER) PUDUCHERRY

MSc BIOSTATISTICS

CURRICULUM

2021

TABLE OF CONTENTS

JIPMER

Jawaharlal Institute of Postgraduate Medical Education and Research, Puducherry (JIPMER) under Government of India since the year 1956, is one of the leading Medical Institutions of India. Spread over a sprawling 195-acre campus in an urban locale of Puducherry (formerly Pondicherry), JIPMER is 170 kms by road from Chennai.

JIPMER has been declared as an "Institution of National Importance" by an Act of Parliament, JIPMER, Puducherry, Act, 2008. A copy of the Act was Gazette notified on 14-7-2008 to enforce this Act. Prior to this, the Institute was functioning under the administrative control of Directorate General of Health Services, Ministry of Health and Family Welfare, New Delhi. The Institution is now empowered to award Medical Degrees under the clauses 23 & 24 of the said Act. Such Degrees shall be deemed to be included in the schedules to the respective Acts governing Medical Council of India/National Medical Commission, Indian Nursing Council, and Dental Council of India, entitling the holders to the same privileges as those attached to the equivalent awards from the recognized Universities of India.

JIPMER imparts Undergraduate (UG), Postgraduate (PG) and Super Specialty Medical Training through a large hospital complex (JIPMER Hospital) and a Nursing College. Some of the courses offered are MBBS., BSc., MSc., MD., MS, DM., and MCh. Courses. Full-time Ph.D. Programs are available in several disciplines.

DEPARTMENT OF BIOSTATISTICS

An independent full-fledged Department of Medical Biometrics and Informatics was established in the Institute in the year 2011. The department has been renamed as Department of Biostatistics in the year 2018. The department have been actively involved in teaching Research Methodology, Biostatistics and Demography to all Undergraduates, Postgraduate Medical & Paramedical students. Regular statistical consultations and guidance are provided to the postgraduate residents and students, faculty members, PhD scholars and staff in various departments of the institute.

The department is conducting MSc course in Biostatistics and Doctoral Programme in Biostatistics. The department is also conducting regular training programs in Research Methodology and Biostatistics for the faculty, residents, research scholars and students. The faculty members of the department are also regularly participating as resource person in the various workshops and CMEs on research methodology, Biostatistics, systematic review, and meta-analysis etc. conducted by the national institutes.

2. Course details:

2.1 Nomenclature of the course: MSc Biostatistics

2.2 Objectives:

- To train the students to pursue a career that encompasses Biostatistics as a decision science
- To train the students with inter-disciplinary expertise through strong advanced statistical methods
- To provide rigorous training in the modern areas of biostatistics related to the theory and application of statistical science to solve problems in epidemiology and public health, Clinical Trials, health services and policy, biomedical research, genetics, and other areas related to medicine
- To train in strong methodological foundations in Designing and analysing studies in biomedical research
- To train the students in advanced statistical analysis using different statistical packages
- To provide versatile training in statistical consultations

3. Regulations and Syllabus

3.1 Eligibility:

A candidate seeking admission should have qualified in B.Sc. degree in statistics (major) or B.Sc. degree in mathematical Science with Statistics as a subject from any recognized University or an examination of some other University/ Institute accepted by the institute as equivalent thereto shall be eligible for admission.

3.2 Method of Selection:

Candidates for the M.Sc. Biostatistics course will be selected on the basis of merit in the common All India Entrance Examination conducted by the Institute. The Entrance Examination will be a Computer Based Test (CBT) held online at centres across the country. The test is based on multiple choice questions and will be at the level of BSc.

3.3 Candidates intake per year:

Two candidates will be admitted to the course every year. There is no provision for sponsored / nominated candidates.

3.4 Duration of the course

2 years (Two years). Examinations will be held at the end of each academic year. There are no stipendiary provisions during the study period.

3.5 Medium of instruction: English

3.5 Subject details:

Year	Paper	Title of the paper
I year	I	Probability and Probability Distributions.
	II	Sampling & Sample Size Estimation in Biomedical Research
	III	Design of Experiments.
	IV	Statistical Inference in Biomedical Research
	V	Design and Analysis of Epidemiological Studies
	VI	Design and Analysis of Clinical Trials
	VII	Practical-I (Consist of problems related to papers I to III)
	VIII	Practical-II (Consists of problems related to papers IV to VI)
II year	I	Demography, Vital Statistics & Statistical Genetics
	II	Statistical Modeling of Biomedical Data
	III	Multivariate Analysis in Biomedical Research
	IV	Survival Analysis & Time Series Analysis
	V	Practical-III (Consists of problems related to papers I & II)
	VI	Practical-IV (Consists of problems related to papers III & IV)

SYLLABUS

First year Paper I: Probability and Probability distributions

Total teaching hours: 80 hours

Unit-I: Probability Theory (25%) (20 hours)

Random experiments, events, Sample space, probability, various definitions of probability, probability space, properties of probability, addition and multiplication theorems of probability, conditional probability, and Bayes theorem.

Random variable, probability mass function, probability density function, cumulative distribution function. Bivariate random variable, joint, marginal, and conditional distributions.

Expectation of a random variable and its properties, conditional expectation, addition and multiplication theorem of conditional expectation, moments and Moment Generating Function (MGF), Characteristics Function (CF), uniqueness theorem of MGF and CF and Probability Generating Function (PGF), properties of CF, MGF and PGF.

Unit- II: Probability distributions (50%) (40 hours)

Discrete univariate distributions

Binomial, Rectangular or Uniform, Poisson, Negative binomial, Hypergeometric, Logarithmic, Geometric and Multinomial distributions.

Definition, history, properties, mean, variance, moments, Moment Generating Function (MGF) and applications of all the above-mentioned discrete distributions.

Continuous univariate distributions

Gama, Beta, Normal distribution, Student t, F, Chi-square, Exponential, Weibull, Logistic and Lognormal distributions.

Definition, history, properties, mean, variance, moments and applications of all the above-mentioned continuous distributions.

Multivariate distributions

Hotelling's T^2, generalized T^2, Mahalanobis's D^2 and Wishart distributions, Cochran's theorems.

Unit III: Limit theorems (25%) (20 hours)

Inequalities: Chebyshev's, Markov's, Holder's, Jensen's and Minkowski's inequalities, Weak and strong convergence: (Convergence in probability, Convergence in distribution, almost sure convergence), Law of large numbers (weak and strong), Central limit theorem for independently and identically distributed random variables (statement only) and applications. Log, Inverse and square root transformations - concept and applications.

Textbooks (Latest edition):

1. Hogg RV, Tanis EA, Zimmerman DL. Probability and statistical inference. Pearson/Prentice Hall.
2. Rao CR. Linear statistical inference and its applications. New York: Wiley.
3. Rice JA. Mathematical statistics and data analysis. Cengage Learning.
4. Rosner, B. Fundamentals of Biostatistics. Cengage Learning.
5. Schay G. Introduction to probability with statistical applications. Birkhäuser.

Reference Books (Latest edition):

1. Casella G, Berger RL. Statistical inference. Pacific Grove, CA: Duxbury.
2. Rohatgi, VK, Saleh, KE. An Introduction to probability and statistics. Wiley.
3. Sullivan LM. Essentials of Biostatistics in Public Health. Jones & Bartlett Learning.
4. Daniel WW, Cross CL. Biostatistics: a foundation for analysis in the health sciences. Wiley.

First year Paper II: Sampling and Sample Size Estimation in Biomedical Research

Total teaching hours: 80 hours

Unit-I (10%): Introduction (8hours)

Data: Primary and secondary data, methods of collecting primary data: Census and sample surveys, their advantages and disadvantages. Sources of secondary data, its advantages and disadvantages. Population versus sample, sampling frame, selection and inclusion probabilities, sampling and non-sampling errors, sources of non-sampling errors, methods to reduce non-sampling errors, bias, mean squared error, advantages of sample studies, need and importance of sampling in Biomedical research.

Unit – II (30%): Random Sampling (24 hours)

Random Sampling - different types - Simple random sampling with replacement and without replacement (SRSWR & SRSWOR), different methods of selection of sample in SRS: lottery method and random number table method and their advantages and disadvantages. Estimation of mean, Proportion and standard error using SRS - application, advantages and disadvantages of SRS.

Stratified random sampling: Concept of Stratified Random sampling, stratification and selection of samples in stratified random sampling, allocation problem and construction of strata, determination of strata boundaries, number of strata, allocations for multiple characteristics, variance estimation, application, gain due to stratification, advantages and disadvantages.

Systematic sampling: Concept of systematic sampling, variance estimation, selection procedure for fractional interval, circular systematic sampling, advantages and disadvantages of systematic sampling, application, comparison with simple random sampling and stratified random sampling.

Cluster sampling: Concept of Cluster sampling, simple and multi-stage cluster sampling, equal and unequal cluster sizes, primary and secondary clusters, design effects, stratified-cluster sampling, application, comparison with simple random, stratified and systematic random sampling. Concept of multi-stage sampling.

Unit – III (10%) (8 hours)

Use of auxiliary information, ratio and regression methods of estimation under simple random sampling, ratio and regression estimators in stratified random sampling, estimation of variance, application, advantages and disadvantages.

Unit – IV (15%): Non-Random Sampling Methods (12 hours)

Concept of Non-Random Sampling, its types: convenient, purposive, quota, double, and snowball sampling methods. Advantages and disadvantages of non-random sampling methods.

UNIT – V (35%): Sample size determination (28 hours)

Need for and importance of sample size estimation in biomedical research, type-I and Type II errors, power of a test, level of significance, precision: absolute and relative, effect size. Sample size for estimating the mean & proportion, difference in means & proportions, correlation coefficients. Sample size for testing of hypothesis- One sample, two samples and more than two sample means, comparison of two proportions, sample size for testing the difference in paired samples.

Sample size for estimating and comparing the diagnostic validity of tool/test, estimating, and comparing the agreements, regression methods, survival analysis and other epidemiological and clinical trials.

Textbooks (Latest edition):

1. Cochran WG. Sampling techniques. John Wiley & Sons.
2. Som RK. Practical sampling techniques. CRC press.
3. Chow SC, Wang H, Shao J. Sample size calculations in clinical research. CRC press.
4. Lemeshow S, Hosmer DW, Klar J, Lwanga SK. Adequacy of sample size in health studies. Published on behalf of the World Health Organization by Wiley.

Reference Books (Latest edition):

1. Gupta AK, Kabe DG. Theory of sample surveys. World Scientific Publishing pvt. Limited.
2. Singh D, Chaudhary FS. Theory and Analysis of Sample Survey Designs. New Age International Publisher.
3. Schlesselman J J. Case-Control Studies: Design, conduct and analysis. Oxford university press.

First year Paper III: Design of Experiments

Total teaching hours: 80 hours

Unit- I: Introduction (15%) **(12 hours)**

Design of experiments, basic principles of experimental design: Randomization, Replication & blocking, cause of errors: assignable & chance causes, steps in planning experiments, accuracy of experiments.

Unit-II: Experimental designs (75%) **(60 hours)**

Completely Randomised design

Description of the design, model and its assumptions, components, hypothesis testing steps, test statistics, basic ANOVA table, interpretation, application.

Randomised complete block design

Description of the design, model and its assumptions, components, hypothesis testing steps, test statistics, basic ANOVA table, interpretation, application.

Balanced incomplete block design

Description of the design, model and its assumptions, components, hypothesis testing steps, test statistics, basic ANOVA table, interpretation, application.

Factorial design

2^n, 3^2 factorial experiments: Description of the design(illustrations, main effects and interactions), model and its assumptions, components, hypothesis testing steps, test statistics, basic ANOVA table, interpretation, Application. Response surface experiments; first order designs and orthogonal designs.

Split-plot design

Description of the design, model and its assumptions, components, hypothesis testing steps, test statistics, basic ANOVA table, interpretation, application.

Latin square design and incomplete Latin Square design

Description of the design, model and its assumptions, components, hypothesis testing steps, test statistics, basic ANOVA table, interpretation, application.

Repeated measures design, Cross over design, Analysis of Covariance and missing plot techniques.

Unit-III : Confounding (10%) **(8 hours)**

Principles of Confounding, Uses of confounded designs and statistical analysis. Confounding in factorial designs - Total and Partial confounding.

Textbooks (Latest edition):

1. Dean A, Voss D, Draguljić D. Design and analysis of experiments. New York: Springer.
2. Montgomery DC. Design and analysis of experiments. John wiley & sons.
3. Oehlert GW. A first course in Design and Analysis of experiments. Library of Congress Cataloging.

Reference Books (Latest edition):

1. Hinkelmann K, Kempthorne O. Design and analysis of experiments: Volume 2 Advanced Experimental Design. John Wiley and Sons.
2. Jobson JD. Applied multivariate data analysis: regression and experimental design. Springer Science & Business Media.

First year Paper IV: Statistical Inference in Biomedical Research

Total teaching hours: 80

UNIT: I - Theory of Estimation (20%)

Module-I: Point Estimation (10%) (8 hours)

Concept of statistical inference,Introduction to point estimation, parameter, statistic, estimator, properties of a good estimator: unbiased, sufficiency, efficiency and consistency, unbiased minimum variance estimator, Cramer-Rao- inequality, likelihood function, Maximum likelihood estimator and its properties, Factorization theorem, Rao-Blackwell theorem.

Module-II: Interval Estimation (10%) (8 hours)

Sampling distribution, distribution of sample mean, confidence interval & its interpretation, confidence interval for population mean: large and small samples, confidence interval for difference between two population means, population variance equal & unequal situations, confidence interval for population proportion & difference between two proportions, confidence interval for the variance and ratio of two variances.

UNIT: II - Testing of Hypothesis (80%)

Module-1: Parametric tests (30%) (24 hours)

Introduction to hypothesis testing and test of significance, statistical & clinical significance, , Neyman-Pearson lemma, likelihood ratio test, hypothesis testing – single population mean, difference between population means, Levene's test for equality of variance, ratio of two variances, paired comparison of mean, single population proportion, difference between two population proportions, paired comparison of proportions. Score statistic or Wald statistic, ANOVA model & post hoc tests, power calculation.

Module II: Non - Parametric Methods (30%) (24 hours)

Introduction & Analysis of contingency tables

Introduction to Non-parametric methods, advantages, and disadvantages of non-parametric methods. Analysis of 2 x 2 tables, assumptions and limitations, Chi-square test and Fisher-exact test, extension to r x c tables.

One sample methods & Test of goodness of fit:

Binomial test, Sign test, Chi-square test for goodness of fit, test of normality: Kolmogrov-Smirnov test.

Two independent samples comparisons & its extension:

Willcoxon rank sum/Mann Whitney U test, Kruskal-Wallis tests & multiple comparisons.

Two related samples comparisons & its extension:

Willcoxon-signed rank test, Mc-Nemar's test and paired comparison permutation test, Cochran Q test.

Module-III: Correlation and its test of significance (20%) (16 hours)

Introduction to correlation, types: Cramer coefficient, Phi coefficient, Lambda, Gamma, Kendall's Tau. Karl Pearson correlation coefficient and its properties. Spearman's rank correlation coefficient.

Textbooks (Latest edition):

1. Hogg RV, Tanis EA, Zimmerman DL. Probability and statistical inference. Pearson/Prentice Hall.
2. Rao CR. Linear statistical inference and its applications. New York: Wiley.
3. Rice JA. Mathematical statistics and data analysis. Cengage Learning.
4. Rosner, B. Fundamentals of Biostatistics. Cengage Learning; 2010.
5. Schay G. Introduction to probability with statistical applications. Birkhäuser.

Reference Books (Latest edition):

1. Casella G, Berger RL. Statistical inference. Duxbury advanced series; Wadsworth.
2. Daniel WW, Cross CL. Biostatistics: a foundation for analysis in the health sciences, Wiley.
3. Lehman EL, Romano JP. Testing Statistical Hypotheses, Springer.
4. Rohatgi VK, Saleh KE. An Introduction to probability and statistics. Wiley.

First year Paper V: Design and Analysis of Epidemiological Studies

Total teaching hours: 80 hours

Introduction to epidemiological studies (15%) (12 hours)

Introduction to epidemiology, need and importance of epidemiological research in public health, objectives & scope of research, classical, clinical, spatial and environmental epidemiology, research methods Vs methodology, types of research: Qualitative Vs Quantitative, Applied Vs Fundamental, Conceptual Vs Empirical, Field based Vs Laboratory based, Research process, good research criteria, formulation of research question and its key considerations, formulation of hypothesis, review of literature, steps in formulating a good research proposal; variables: exposure, outcome and extraneous variables.

Research designs (5%) (4 hours)

Concept and definition, need and importance, features of good research designs, different research designs: Qualitative and quantitative. Qualitative methods: FGDs, in-depth interviews, biographies, participatory methods, participant observation, etc. - Data collection, recording and Data analysis.

Quantitative study designs (25%) (20 hours)

Observational & Experimental study designs; Observational studies: Concept and objectives of observational studies, explorative, descriptive and analytical study designs and their importance, advantages and disadvantages; descriptive study designs, case study, case-series, cross sectional, analytical studies, case control, cohort-prospective, retrospective, nested case-control. Strengths and limitations of different study designs.

Measuring the occurrence of disease: (5%) (4 hours)

Rate, Ratio, proportion, Measures of morbidity - prevalence and incidence rate, association between prevalence and incidence, uses of prevalence and incidence, problems with incidence and prevalence measurements.

Measurement and scaling Techniques: (**5%**) **(4 hours)**

Measurement scales, sources of errors in measurements, validity, reliability and practicability, important scaling techniques: rating scales, arbitrary scales, differential scales, summated scales, cumulative scales, multidimensional scales, advantages and disadvantages of each scales.

Statistical measures of Risk: (10%) (8 hours)

Odds, Odds Ratio (OR) and log Odds Ratio, confidence interval for OR and statistical significance of OR; Relative Risk (RR) and confidence interval for RR, attributable risk.

Evaluation of Diagnostic/screening tool: (15%) (12 hours)

Introduction, need and importance of evaluation of diagnostic tools, validity and reliability

of diagnostic/screening tools, sensitivity, specificity, PPV, NPV and Likelihood Ratios, Receiver Operative characteristic curve, Area Under the curve, Statistical significance of the Area Under Curve.

Validity and reliability: (15%) (12 hours)

Concept and definition, importance of validity and reliability tool/instrument and measurements, different validity measures: internal and external validity, external validity : face validity, content validity, construct validity, criterion validity ; reliability and different methods and measures of reliability, test-retest method, split-half method, equivalence, inter-rater/observer agreement, intra-class correlation (ICC), Kappa statistics, internal consistency and Cronbach's alpha; Variation: measurement, biological and total variations & its effects.

Association & causation: (5%) (4 hours)

Association; causation; causal inference; errors and bias; Confounding; Controlling confounding; Measurement of interactions.

Textbooks (Latest edition):

1. Merrill RM. Fundamentals of Epidemiology and Biostatistics; combining the basics. Jones and Bartlett learning.
2. Lilienfeld DE, Stolley PD, Lilienfeld AM. Foundations of epidemiology. Oxford University Press, USA.
3. Merrill RM. Introduction to Epidemiology. Seventh edition. Jones & Bartlett Learning.
4. Rothman KJ, Greenland S, Lash TL, editors. Modern epidemiology. Lippincott Williams & Wilkins.

Reference Books (Latest edition):

1. Weiss NS. Epidemiology and Biostatistics: An Introduction to Clinical Research. Springer.
2. Woodward M. Epidemiology: Study design and data analysis. CRC press.
3. Schlesselman J J. Case-Control Studies: Design, conduct and analysis. Oxford university press.
4. Gordis L. Epidemiology. Sounders.
5. Epidemiology in Medicine. Hennekens HC, Buring JE. Lippincott Williams and Wilkins.
6. Park K. Park's Textbook of Preventive and social medicine, Bhanot Publishers.

First year Paper VI: Design and Analysis of Clinical Trial

Total teaching hours: 80 hours

Unit-I (20%): (16 hours)

Introduction to clinical trials: Basic concepts; Definitions; Historical perspectives of clinical trials, objectives of clinical trials, clinical trial and clinical study, The need and Ethics of clinical trials, bias and random error in clinical studies. Conduct of clinical trials, overview of Phase-I to IV trials, single and multi-center trials. Primary & secondary responses, concepts of blinding, randomization techniques.

Unit-II (20%): (16 hours)

Design of clinical trials: Parallel, cross over, factorial, cluster randomized trial designs and design effects, Single, double and triple blind trials. Type of control, design of studies with matched and unmatched controls, end points of clinical trials.

Unit-III (15%): (12 hours)

Design of Phase I trials, design of single-stage and multistage phase II trials, design and monitoring of Phase III trials with sequential stopping, design of bioequivalence trials.

Unit IV (15%): (12 hours)

Analysis & Reporting: Analysis of categorical and continuous outcomes from Phase-I to III trials, intention to treat and per protocol analysis, NNH and NNT.

Unit-V (20%): (16 hours)

Surrogate end Points: Selection and design of trials with surrogate end points, analysis of surrogate end point data, Meta-analysis of clinical trials.

Unit VI: (10%): (8 hours)

Planning and conduct of clinical trials: Protocol development; Multi-centric trials; Deviations from protocol; Stopping rules; Considerations of adverse effects and non-compliance.

Textbooks (Latest edition):

1. Evans S,Ting N. Fundamental concepts for new clinical trialists. CRC Press.
2. Friedman LM, Furberg C, DeMets DL, Reboussin DM, Granger CB. Fundamentals of clinical trials. New York: Springer.
3. Hackshaw A. A concise guide to clinical trials. John Wiley & Sons.
4. Shih WJ, Aisner J. Statistical design and analysis in clinical trials. CRC Press.

Reference Books (Latest edition):

1. Chen D, Peace KE, Zhang P. Clinical Trial Data Analysis Using R and SAS. Second Edition. CRC Press.
2. Dmitrienko A, Pulkstenis E. Clinical Trial Optimization Using R. CRC Press.
3. Hayes RJ, Moulton LH. Cluster randomized trials. CRC Press.
4. Korosteleva O. Clinical Statistics: introducing clinical trials, survival analysis and longitudinal data analysis. Jones and Bartlett Pulishers.
5. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to Meta-Analysis, Wiley.

Scheme of Assessment for MSc Biostatistics I year:

Theory examination:

	Theory papers	**Annual Examination**	**Internal Examination**	**Maximum marks**
Paper I	Probability and Probability Distributions.	70	30	100
Paper II	Sampling & Sample Size Estimation in Biomedical Research	70	30	100
Paper III	Design of Experiments	70	30	100
Paper IV	Statistical Inference in Biomedical Research	70	30	100
Paper V	Design and Analysis of Epidemiological Studies	70	30	100
Paper VI	Design and Analysis of Clinical Trials	70	30	100
Total Theory marks				**600**

.

Practical examinations and viva-voce examinations (Two days):

		Annual Examination	**Viva-voce Examination**	**Maximum marks**
Part I	Problems related to Probability and Probability Distributions, sampling & sample size estimation in biomedical research, and design of experiments.	70	30	100
Part II	Problems related to statistical inference in biomedical research, design, and analysis of epidemiological studies, design, and analysis of clinical trials.	70	30	100
Total Practical marks				**200**

Second year Paper I: Demography, Vital Statistics & Statistical Genetics

Total teaching hours: 80 hours

Unit-I: Demography & Vital Statistics (70%)

Module-I (10%): (8 hours)

Demography and Vital Statistics: Scope and Importance. Data Sources and their Limitations: Population Census, Vital Registration, Population Register, Demographic and Health Surveys, Other Sources, Vital Registration System in India- Civil Registration System (CRS), Sample Registration System (SRS), and Cause of Death Reporting system in India**.**

Module-II (10%): (8 hours)

Measure of Fertility (Crude and Specific rates of Fertility): CBR, GFR, GMFR, TFR, TMFR, Age Specific Fertility Rate, GRR, NRR. Uses of fertility indicators. Data sources, distribution of time to first birth, inter-live birth intervals and number of births (for both homogeneous and non-homogeneous groups of women), estimation of parameters, estimation of parity progression ratios from open birth interval data.

Module-III (10%): (8 hours)

Measure of Mortality (Crude and Specific rates of Mortality): Crude death rate, age specific death rate, Perinatal mortality rate, Neonatal mortality rate, Post-neonatal mortality, infant mortality rate. Mortality levels, trends, and differentials. Life tables: elements of life tables, construction of complete and abridged life tables and application of life tables. Expectation of life and its uses.

Module-IV (10%): (8 hours)

Migration: Migration Rates and Ratios, Indirect measures for estimating net internal migration. National growth rate method. Estimation of Measures of Mobility.

Module-V (10%): (8 hours)

Measurement of Population growth: Linear, Geometric, Exponential, Gompertz, Logistic Population growth models and their fitting to population data. Methods of population projection, uses of Leslie matrix. Stable and Quasi Stable populations, intrinsic growth rate. Stochastic models for population growth.

Module-VI (10%): (8 hours)

Demographic transition: concept and elements of demographic transition, different stages of demographic transition: pre-transition, transition, and post- transition stages. Demographic cycle: high stationary, early expanding, late expanding, low stationary and decline. Sex ratio. Dependency ratio. Divergence and convergence of population.

Module-VII (10%) **(8hours)**

International Classification of Diseases (ICD), Injuries and Causes of Death and International classification of functioning, disability, and health (ICF), Uses of ICD and ICF, difference between ICD and ICF, Structure of ICD and ICF, Coding in ICD and ICF.

Unit-II: Statistical Genetics (30%) **(24 hours)**

Physical basis of inheritance. Analysis of segregation, detection and estimation of linkage for qualitative characters, disturbed segregation. Gene and genotypic frequencies. Random mating and Hardy-Weinberg law. Application and extension of the equilibrium law. Fisher's fundamental theorem of natural selection. Disequilibrium due to linkage for two pairs of genes, sex-linked genes. Forces affecting gene frequency: selection, mutation and migration, equilibrium between forces in large populations. Polymorphism. Theory of path co-efficients. Regular systems of inbreeding. Polygenic system for quantitative characters, concepts of breeding value and dominance deviation. Genetic variance and its partitioning. Correlation between relatives, Heritability, Repeatability and Genetic correlation.

Textbooks (Latest edition):

1. Balding DJ. Bishop M. Cannings C. Handbook of Statistical Genetics. John Wiley.
2. Fisher RA. Statistical Methods for Research Workers. Oliver and Boyd.
3. Majumdar PK. Fundamentals of Demography. Surjeet Publications.
4. Cox PR. Demography. Fifth Edition, Cambridge University Press.
5. Campbell MJ, Machin D, Walters SJ. Medical statistics: A textbook for health sciences, Wiley.

Reference Books (Latest edition):

1. Keyfitz N, Caswell H. Applied mathematical demography. New York: Springer.
2. Lancaster HO. Expectations of life: a study in the demography, statistics, and history of world mortality. Springer Science & Business Media.
3. Narain P. Statistical Genetics. Wiley Eastern Limited.
4. Poston DL, Micklin M. Handbook of Population. Kluwer Academic/Plenum Publishers.

Second year Paper II: Statistical Modeling of Biomedical data

Total teaching hours: 80 hours

Unit-1- Simple Linear and Multiple Linear Regression (20%) (16 hours)

Simple Linear regression: Introduction, model, Assumptions, examples, and fitting. Methods of estimation: Ordinary Least Square (OLS) and Weighted Least Squares, Coefficient of determination, Residual Analysis, Inferences concerning intercept and slope, Confidence intervals for intercept and slope, Prediction intervals, Regression through the Origin, ANOVA approach to regression, F-distribution, Regression to mean, Outliers, Goodness of fit of the model.

Multiple Linear regression: Introduction with Visual presentations, model, examples, assumptions and Estimation. Fundamental Equation of Regression Analysis, ANOVA approach to Multiple regression, Regression diagnostics, Marginal effects of covariates, Pooled tests of significance Uncorrelated Predictors, Multicollinearity, Confounding, Goodness of fit of the model.

Unit-II - Logistic Regression Models (30%) (24 hours)

Introduction, model, assumptions, model fitting and Estimation. Conditional and unconditional models, Fundamental Equation of logistic regression Analysis, testing significance of model, interactions and confounding, Odds ratio, interpretations of the results, logistic regression for ordinal and Polytomous data

Multiple logistic regression model: Introduction, model, assumptions, model fitting, conditional and unconditional models, Estimation, Fundamental Equation of multiple logistic regression Analysis, testing significance of model, interactions and confounding, Odds ratio, interpretations of the results, forward, backward and step wise selection of independent variables, Goodness of fit of the model.

Multinomial logistic regression model: Introduction, assumptions, model fitting, conditional and unconditional models, Estimation, Fundamental Equation, testing significance of model, goodness of fit of the model.

Introduction to Count data and Poisson regression models, model fitting, testing the significance and goodness of fit.

Unit-III- Loglinear Models (15%) (12 hours)

Introduction, log Linear Models for two-way tables, Log linear models for Interaction and independence in three way tables. Interpreting Model parameters. Inference for log linear models, Model fitting and estimating parameters, iterative methods and their applications. Log linear models for related samples.

Unit-IV - Model for paired data (15%) (12 hours)

Measures for paired data - Stuart Maxwell Mc-Nemar's test for Marginal homogeneity for related square table, Cochran Mantel-Haenszel correlation statistic for testing marginal homogeneity in a R×C×K table. Measuring agreement between observers-Kappa measure of agreement.

Unit-V - Generalized Linear Modeling (20%) **(16 hours)**

The exponential family of distributions - Sufficient statistics of parameters of exponential family of distributions for Binomial, Multinomial, Poisson, Normal, Negative Binomial, Exponential and Weibull. Components of GLM: systematic, random and Link functions. Uses of log link, logit link, identity link and probit link.

Textbooks (Latest edition):

1. Rencher AC. Linear Models in Statistics, John Wiley & Sons, New York.
2. Hosmer DW., Lemeshow S. Applied logistic regression. Wiley.
3. Indrayan A, Malhotra RK. Medical biostatistics. CRC Press.

Reference Books (Latest edition):

1. Agresti A. Categorical data analysis. John Wiley & Sons.
2. Chen DG, Peace KE, Zhang P. Clinical trial data analysis using R and SAS. CRC Press.
3. Dobson AJ, Barnett AG. An introduction to generalized linear models. CRC press; 2018.
4. Fox J. Applied regression analysis and generalized linear models. Sage Publication.
5. Hilbe JM, Robinson AP. Methods of statistical model estimation. CRC Press.
6. McCullagh, P, Nelder JA, Generalized Linear models. CRC Press.
7. Peck EC, Montgomery DC, Vinning GG. Introduction to Linear Regression Analysis, 3/e, John Wiley & Sons.
8. Weisberg S. Applied Linear Regression. Third edition. Wiley-interscience.

Second year Paper III: Multivariate Analysis in Biomedical Research

Total teaching hours: 80 hours

Unit-I: Discriminant Analysis (20%) (16 hours)

The Discriminant Function for Two Groups, Relationship between Two-Group Discriminant Analysis and Multiple Regression, Discriminant Analysis for Several Groups: Discriminant Functions, Standardized Discriminant Functions. Tests of Significance: Tests for the Two- Group Case, Tests for the Several-Group Case. Interpretation of Discriminant functions: Standardized Coefficients, Partial *F*-Values, Correlations between Variables and Discriminant Functions. Stepwise selection of Variables. Classification analysis – classification into two and several groups, linear classification function, quadratic classification functions, estimating misclassification rate, improved estimates of error rates – partitioning the sample and Holdout method

Unit-II: Canonical Correlation & Spatial Correlations (20%) (16 hours)

Canonical Correlations and Canonical Variates, Properties of Canonical Correlations, Tests of Significance - Tests of No Relationship between the *y*'s and the *x*'s. Interpretation - Standardized Coefficients, Correlations between Variables and Canonical Variates. Relationships of Canonical Correlation Analysis to Other Multivariate Techniques - Regression, MANOVA and Discriminant Analysis. Spatial Correlations

Unit-III: Principal Component Analysis (20%) (16 hours)

Geometric and Algebraic Bases of Principal Components, Principal Components and Perpendicular Regression, Plotting of Principal Components, Principal Components from the Correlation Matrix, Deciding How Many Components to Retain, Information in the Last Few Principal Components, Interpretation of Principal Components - Special Patterns in sample covariance matrix or correlation matrix, Rotation, Correlations between Variables and Principal Components, Selection of Variables.

Unit-IV: Factor Analysis (20%) (16 hours)

Confirmatory and Exploratory factor Analysis, Orthogonal Factor Model: Model Definition and Assumptions, Nonuniqueness of Factor Loadings. Estimation of Loadings and Communalities: Principal Component Method, Principal Factor Method & other methods (introduction). Choosing the Number of Factors *m*. Rotation- Orthogonal Rotation, Oblique Rotation, Interpretation. Factor Scores, Validity of the Factor Analysis Model, The Relationship of Factor Analysis to Principal Component Analysis.

Unit-V: Cluster Analysis (20%) (16 hours)

Measures of Similarity or Dissimilarity, Hierarchical Clustering - Single Linkage (Nearest Neighbor), Complete Linkage (Farthest Neighbor), Average Linkage, Centroid, Median, Ward's Method, Divisive Methods. Non-hierarchical Methods – Partitioning - K means, Choosing the Number of Clusters, Cluster Validity, Clustering Variables.

Textbooks (Latest edition):

1. Anderson TW. An Introduction to Multivariate Analysis. Wiley.
2. Johnson RA, Wichern DW. Applied Multivariate Statistical Analysis. Prentice Hall.
3. Rencher AC. Methods of Multivariate Analysis. 2nd Edition. John Wiley & Sons.

Reference Books (Latest edition):

1. Brian S. Everitt, Graham Dunn. Applied Multivariate Data Analysis. John Wiley & Sons.
2. Gnanadesikan R. Methods for statistical data analysis of multivariate observations. John Wiley & Sons.
3. Harlow LL. The Essence of Multivariate Thinking: Basic Themes and Methods. Routledge.
4. Ho R. Handbook of Univariate and Multivariate Data Analysis with IBM SPSS Taylor & Francis Group.
5. Jobson JD. Applied multivariate data analysis: regression and experimental design. Springer Science & Business Media.
6. Rao CR. Linear statistical inference and its applications. New York: Wiley.

Second year Paper IV: Survival Analysis and Time Series Analysis

Total teaching Hours: 80 hours

Unit-I: Survival Analysis (75%) **(60 hours)**

Module-I (10%): **(8 hours)**

Introduction to survival analysis: concept, definition and importance in biomedical research. Censoring: concept of censoring and type of censoring. Survival function, probability density function, hazard function or rate, cumulative hazard function, relationship between the three types of function, survival curve, estimating medium survival time, estimation of these function in the absence and presence of censoring, application of these functions in survival analysis.

Module-II (20%): **(16 hours)**

Survival distributions- Exponential, Weibull, Pareto, Lognormal and Gamma distribution. Parametric estimation with complete and censored samples. Large sample tests under censored data. The E–M algorithm.

Module-III (20%): **(16 hours)**

Nonparametric methods of estimating survival function - introduction, Kaplan- Meier estimates, life table estimates, clinical life tables, life table vs. Kaplan-Meier estimates, The Mantel-Haenszel test, Comparing survival curves: Log rank test, Tarone-Ware tests, Weighted log rank estimators.

Module-IV (25%): **(20 hours)**

Regression methods for survival analysis - Cox-proportional hazard models: Introduction, proportionality assumption. Statement of asymptotic properties of the estimator. Estimation of the baseline hazard function, test of proportionality, interpretation of coefficients, Residuals and model checking. Graphical methods: Hazard plots and Survival plots. Application of Cox-proportional hazard models in Epidemiology and Public Health, Concepts of extended Cox model and MLE of Cox PH model. Introduction to accelerated models, risk models, discrete-time survival models and frailty models

Unit-II: Time series Analysis (25%) **(20 hours)**

Module-1(10%): **(8 hours)**

Introduction to time series analysis. Autocovariance and autocorrelation functions and their properties. Partial autocorrelation function. Exploratory Time Series Analysis, Tests for trend and seasonality. Deseasonalising and detrending an observed time series. Smoothing: Exponential, Moving Average, Holt and Winters, Adaptive smoothing.

Module-II (15%): **(12 hours)**

Stationary processes: Moving average (MA), Autoregressive (AR), ARMA and Invertibility. Box-Jenkins models. ACF and PACF plots of these processes. Sample ACF and PACF. Gaussian time series. Model identification and Model checking. Non-stationary

time Series models: Unit root test non-stationarity, unit root test, ARIMA and Seasonal ARIMA Models.

Textbooks (Latest edition):

1. Hosmer DW, Lemeshow S, May S. Applied survival analysis: Regression modelling of time to event data. Wiley.
2. Chatfield C. The analysis of time series an introduction. Chapman and Hall CRC Press.
2. Box GEP, Jenkins GM, Reinsel GC. Time Series Analysis Forecasting and Control.Wiley.
3. James HD. Time Series Analysis. Princeton University Press.
4. Indrayan A, Malhotra RK. Medical biostatistics. CRC Press.
5. Klein JP, Van Houwelingen HC, Ibrahim JG, Scheike TH, editors. Handbook of survival analysis. CRC Press.
6. Lee ET, Wang J. Statistical methods for survival data analysis. John Wiley & Sons.

Reference Books (Latest edition):

1. Belle GV, Fisher LD, Heagerty PJ, Lumley T. Biostatistics a methodology for the health sciences. Second edition. John Wiley.
2. Kirchgässner G, Wolters J. Introduction to Modern Time Series Analysis. Springer.
3. Hamilton JD. Time Series Analysis. Princeton University Press.
4. Klein JP, Moeschberger, ML. Survival Analysis: Techniques for Truncated and Censored Data. Second Edition. Springer-Verlag.
5. Vittingghoff E, Glidden DV, Shiboski SC, McCulloch CE. Regression methods in Biostatistics Linear, Logistic, Survival and Repeated Measures models. Springer.

Scheme of Assessment for MSc Biostatistics II year:

Theory examination:

	Theory papers	Annual Examination	Internal Examination	Maximum marks
Paper I	Demography, Vital Statistics & Statistical Genetics	70	30	100
Paper II	Statistical Modeling of Biomedical Data	70	30	100
Paper III	Multivariate Analysis in Biomedical Research	70	30	100
Paper IV	Survival Analysis & Time Series Analysis	70	30	100
	Total marks			**400**

Practical examinations and viva-voce examinations (Two days):

		Annual Examination	Viva-voce Examination	Maximum marks
Part I	Problems related to Demography, Vital Statistics & Statistical Genetics, Statistical Modeling of Biomedical Data	70	30	100
Part II	Problems related to Multivariate Analysis in Biomedical Research and Survival Analysis & Time Series Analysis	70	30	100
	Project	150	50	200
Practical and viva-voce examinations and Project work				**400**

Project:

1. Submission of project work is mandatory for the appearing in the final examination. The project work shall be carried out by the candidate under the guidance of a postgraduate teacher in the department. The topics for the project shall be selected within the first year of the course and the final project shall be submitted one month before the final examination. The project work is mainly to orient the candidate towards research methodology. Collaboration works with other departments are encouraged.
2. The project will be sent for review by an external faculty and will be evaluated by the External/Internal Examiners at the time of viva voce examination of the candidate during the final summative exam.
3. The Dean’s office will send the project to the library for record.

Regulations of MSc Biostatistics course:

Attendance:

- Students are required to attend 75% or more of all theory classes held, and 75% or more of practical in each subject to be eligible to appear in the annual examination. Under no condition will a student with less than the prescribed attendance in any subject be allowed to appear in the annual examination of that subject.
- Students with less than 75% attendance in theory and/or practical at the end of any year must start afresh in those subjects by joining with the junior batch of students. No extra classes will be arranged to make such students eligible for the annual examinations. The attendance accrued in the previous academic year in those subject(s) will not be transferred. Students will be required to secure 75% attendance afresh in theory and practical/clinical of subjects detained, after joining the junior batch to become eligible to appear in the annual examination.
- The 25% leverage in attendance includes all types of leaves (including leave on medical grounds). For absence because of illness or any medical condition, a duly approved medical leave from Dean (Academic) with medical and fitness certificate issued/verified by authorized JIPMER clinical faculty member is mandatory. Certificate must be submitted before or within 10 days after availing medical leave.
- Attendance cannot be improved upon by attending classes during the gap between the annual regular examination and supplementary examination held within 6 weeks of the former.
- Students who are detained in all the subjects of a year due to lack of attendance should join the classes with junior batch within 7 days of declaration of the eligibility/detention list or when classes commence, whichever is earlier.
- Students who are detained in one or more subject(s) because of lack of attendance but are eligible to appear for annual examination in at least one subject of the year should join classes with junior batch within 7 days of completion of the last final theory/practical examination or when classes start, whichever is earlier. Attendance accrued in the previous academic year will lapse and attendance will be calculated afresh from the date of joining the junior batch.
- A show-cause notice will be issued to students on continuous unauthorized absence without prior permission for two weeks or more. If such absence extends to a period more than one month for any reason, the student is liable for termination from the course. The decision of the competent authority is final.
- There is **no condonation permissible** for shortage of attendance.

Examination rules and regulations:

1. A student needs to pass in all theory paper(s) and practical examination to progress to the next year.
2. In the first and second years, a student who fails in any one or more theory subject(s) and/or practical examination in the will reappear in a supplementary examination (to be held within 6 weeks) in the theory and the corresponding practical examination (2nd attempt). If he/she passes these subjects at this supplementary examination, he/she will continue with the regular batch (in case of first year).
3. In case of fail result in any theory subject and/or practical in the 2nd attempt (supplementary examination), the student will study with the next junior year's batch and will have to reappear for examination (theory and practical) again the next year (3rd attempt). A student failing in any one or more theory subject(s) and/or practical examination in this attempt will reappear in a supplementary examination in the theory and the practical examination (to be held within 6 weeks) in the subjects (4th attempt).
4. If a student fails even in the 4th attempt, no further chances will be given, and his/her name will be struck off the rolls of JIPMER.
5. A student needs to complete the entire 2-year course within 4 years from the date of enrollment (twice the duration of the course), beyond which he/she will not be allowed to appear in any examination. Any exception for extenuating reasons (e.g. prolonged illness of the candidate, family problems, natural calamities, etc.) will be made only after approval of the Academic Advisory Committee and Head of the institution.
6. No grace marks will be awarded for either theory or practical examinations, under any circumstances.

Pass criteria:

1. A minimum 40% in each of the theory paper and overall aggregate, in theory (all papers put together) should obtain a minimum of 50% marks.
2. A minimum of 50% in practical separately.
3. A minimum of 50% of the grand total (theory, practical, project and internal assessments).

MODEL QUESTION PAPER
MSc Biostatistics
First year Paper I- Probability and Probability Distributions

Duration: Three Hours **Maximum Marks: 70**

Answer ALL Questions. Each question carries TEN marks

1. a) Define Random experiment, Probability and Random variable.

   b) State and prove Bayes theorem. (4+6)

2. State and prove addition and multiplication theorems on probability. (10)

3. a) Explain the axioms of probability.

   b) Explain the properties of characteristic function. (5+5)

4. Derive the Moment generating function of binomial and Poisson distributions. (10)

5. a) Prove that Geometric distribution possess lack of memory property.

   b) Describe the relation between F and χ^2 distributions. (5+5)

6. a) State Central limit theorem and explain its importance.

   b) Explain the characteristics of Normal distribution. (5+5)

7. State and prove Tchebychev's inequality**.** (10)

****

MODEL QUESTION PAPER

MSc Biostatistics

First year Paper II- Sampling & Sample Size Estimation in Biomedical Research

Duration: Three Hours **Maximum Marks: 70**

Answer ALL Questions. Each question carries TEN marks

1. a) Explain various data collection methods.
   b) What are the sources of secondary data? (5+5)

2. a) Explain the procedure of simple random sampling with and without replacement with examples.
   (5+5)
   b) Define (i) Population (ii) sample (ii) Sampling Frame (iv) Bias (v) Sampling error.

3. a) Compare the efficiency of stratified random sampling with Systematic sampling.
   b) Discuss the advantages and disadvantages of non-random sampling. (5+5)

4. Explain various non-random sampling methods. (10)

5. a) Define SRSWR and SRSWOR?
   b) Define Design effect and briefly explain its importance. (5+5)

6. a) Define Stratification. Distinguish between cluster and strata.
   b) Describe sampling and non-sampling error along with their sources. (5+5)

7. Illustrate the method of estimating the mean and variance in simple random sampling, stratified sampling, and systematic sampling techniques. (10)

****

MODEL QUESTION PAPER
MSc Biostatistics
First year Paper III- Design of Experiments

Duration: Three Hours **Maximum Marks: 70**

Answer ALL Questions. Each question carries TEN marks

1. a) Define Design of Experiment. Describe the qualities of a good experimental design.
   b) Explain the principles of Design of experiments with examples. (5+5)

2. Describe the completely randomized design. Discuss the analysis plan of data arising from Completely Randomized Design? (10)

3. a) How is Randomized Block Design efficient over completely randomized design?
   b) Discuss the advantages and disadvantages of RBD. (5+5)

4. Explain the concept of Balanced Incomplete block design and its analysis. (10)

5. Explain the model and analysis of 2^3 factorial design. (10)

6. Describe the test procedure and analysis of Latin Square Design. (10)

7. a) What is Post-hoc tests, explain about at least any two of them. (5+5)
   b) Explain the fixed effect and random effect models.

****

MODEL QUESTION PAPER
MSc Biostatistics
First year Paper IV- Statistical Inference in Biomedical Research

Duration: Three Hours **Maximum Marks: 70**

Answer ALL Questions. Each question carries TEN marks

1. a) Distinguish between (i) Point and interval estimate. (ii) estimator and estimate.
   b) Explain maximum likelihood estimation procedure. (5+5)

2. a) Given a random sample $\{X_1,X_2,\ldots X_n\}$ from a normal population $N(\mu,\sigma^2)$, where σ is known. What is the confidence level for the interval $\left[\bar{x} \pm 1.75\frac{\sigma}{\sqrt{n}}\right]$.
   b) Define type I, type II error, power of the test.
   c) Construct the confidence interval for ratio of Population variances when sampling from Independent Normal Distribution. (2+3+5)

3. a) What is the consistent property of an estimator?
   b) Let $X_1, X_2, ..., X_n$ be a random sample from a population with probability density function $f(x,\theta,\sigma^2) = \frac{1}{\sqrt{2\pi}} \exp\left[\frac{-(x-\theta)^2}{2\sigma^2}\right]; \mu, x \in R, \sigma > 0.$
   Prove that $\bar{X}$ is a consistent estimator of θ. (3+7)

4. State and Prove Neymann-Pearson's Lemma. (10)

5. a) Describe the context, assumptions and procedure involved in Mann-Whitney U test.
   b) Describe the non parametric methods for comparing more than two related samples. (5+5)

6. a) Explain Pearson's correlation coefficient and its properties along with its test of significance. (5+5)
   b) Cramer's coefficient.

7. a) Explain the test procedure for testing the difference between two independent population means.
   b) Mention the advantages and disadvantages of non-parametric tests. (5+5)

****

MODEL QUESTION PAPER

MSc Biostatistics
First year Paper V- Design and Analysis of Epidemiological Studies

Duration: Three Hours **Maximum Marks: 70**

Answer ALL Questions. Each question carries TEN marks

1. a) Define Epidemiology and enumerate its scope and uses in public health.
   b) Distinguish between Research methodology and Research methods. (5+5)
2. a) Explain different stages in the Research Process.
   b) Distinguish between Qualitative and Quantitative study designs. (6+4)
3. Explain different Analytical study designs with suitable examples. (10)
4. Explain different Measurement and Scaling methods in Epidemiology. (10)
5. Explain different measures to estimate the Diagnostic/ Prognostic accuracy of a tool. (10)
6. Illustrate Cross sectional, Case-Control and Cohort Designs. (10)
7. a) Define Error and Bias.
   b) Explain different types of Bias in case-control and cohort study designs. (4+6)

****

MODEL QUESTION PAPER
MSc Biostatistics
First year Paper VI- Design and Analysis of Clinical Trials

Duration: Three Hours **Maximum Marks: 70**

Answer ALL Questions. Each question carries TEN marks

1. a) Define Dose Limiting Toxicity (DLT) and Maximum tolerated dose (MTD).
   b) Explain with flow charts the designs used in Phase I trials. (4+6)

2. Describe in detail the statistical analysis and the outcome measures in phase I trial. (10)

3. Elaborate in detail the single and multi stage designs used in Phase II trials. (10)

4. a) What is the difference in sample size calculation between Phase II and III trials?
   b) Explain in detail the information required to calculate sample size. (3+7)

5. a) Explain bioequivalence trials.
   b) Explain in detail reporting of Phase II trials. (4+6)

6. Explain Analysis of Categorical outcomes in Phase III trials. (10)

7. Explain different phases in clinical trials. (10)

****

MODEL QUESTION PAPER
MSc Biostatistics
Second year Paper I- Demography, Vital Statistics & Statistical Genetics

Duration: Three Hours **Maximum Marks: 70**

Answer ALL Questions. Each question carries TEN marks

1. Explain different sources of Vital statistics in India. (10)

2. a) Discuss in detail the estimation of parity progression ratios from open birth intervals.
   b) Explain in detail the elements and applications of life tables. (4+6)

3. Explain the different mortality and morbidity indicators. (10)

4. a) Explain in detail about the components of Demographic Transition.
   b) Elaborate in detail the different stages of Demographic Transition. (3+7)

5. a) Explain the different methods of Population projection.
   b) Discuss in detail the structure and coding in ICD. (5+5)

6. a) Distinguish segregation and disturbed segregation.
   b) Explain in detail Hardy-Weinberg Law. (5+5)

7. Explain in detail the forces affecting gene frequency. (10)

****

MODEL QUESTION PAPER
MSc Biostatistics
Second year Paper II- Statistical Modeling for Biomedical Data

Duration: Three Hours **Maximum Marks: 70**

Answer ALL Questions. Each question carries TEN marks

1. a) What do you mean by Statistical modeling?.
   b) What are the general criteria for evaluating goodness of fit of a statistical model? (5+5)

2. Elaborate about variable selection procedure in multiple linear regression models. (10)

3. a) Derive the Maximum likelihood estimator of regression coefficient in matrix notation.
   b) What is the need of weighted least square estimation technique? (5+5)

4. Describe the test of significance for regression coefficients in Multiple Linear Regression Model. Interpret the regression coefficients with a case scenario. (10)

5. a) Create a case-scenario for binary Logistic Regression model by considering predictors of continuous, categorical, and ordinal nature.
   b) Brief about Wald statistics with example. (5+5)

6. a) Explain a log-linear model for two way tables.
   b) Write the analysis of the Stuart-Maxwell test for marginal Homogeneity. (5+5)

7. a) Elaborate on Generalized Linear model and its components.
   b) Derive GLM for binary data. (5+5)

****

MODEL QUESTION PAPER
MSc Biostatistics
Second year Paper III- Multivariate Analysis in Biomedical Research

Duration: Three Hours **Maximum Marks: 70**

Answer ALL Questions. Each question carries TEN marks

1. a) Define Eigenvalues and Eigenvectors.

   b) Explain Positive Semi definite and Orthogonal Matrix with an example. (4+6)

2. a) Explain in detail the Hotelling T^2 test to compare two mean vectors.
   b) Create a case scenario where Hotelling T^2 can be used to compare two groups. (6+4)

3. Elaborate in detail the Discriminant Function for two groups and explain the Stepwise selection of variables. (10)

4. Explain in detail the classification and estimation of misclassification rate. (10)

5. a) State the objectives of Principal Component Analysis (PCA).
   b) Explain the procedure of PCA. How to interpret the Principal Components? (3+7)

6. Elaborate in detail the estimation of Factor Loadings and Communalities using Principal Component and Factor method. (10)

7. a) Explain the cross-validation approach to check the validity of a clustering result.
   b) Explain in detail the k-means partitioning method. (5+5)

****

MODEL QUESTION PAPER
MSc Biostatistics
Second year Paper IV- Survival Analysis & Time Series Analysis

Duration: Three Hours **Maximum Marks: 70**

Answer ALL Questions. Each question carries TEN marks

1. a) Define Survival Analysis. Explain the descriptive measures of survival experience of a cohort.
   b) Explain the concept of censoring and its types with examples. (5+5)

2. a) Define survival and hazard function and establish a relationship between them.
   b) What is the Kaplan-Meier Estimator? Explain the construction of confidence interval for K-M Estimate.
   (5+5)

3. a) Explain the Log Rank test for comparing two different treatments.
   b) Describe parametric survival models and how it is different from cox regression hazard model.
   (5+5)

4. Explain in detail about the Cox-Proportional Hazard model in Survival Analysis. (10)

5. a) Explain in detail the various methods of measurement of Trend. (5+5)

   b) What is the need for smoothing? Explain in detail the different types of smoothing.

6. a) Explain in detail autocorrelation and partial Autocorrelation Function with its properties.
   b) Discuss in detail about the non-stationary time series models. (5+5)

7. Elaborate in detail the Box Jenkins Model. (10)

****