COMMON POOL OF GENERIC ELECTIVES (GE) COURSES OFFERED BY DEPARTMENT OF STATISTICS CATEGORY-VI

GENERIC ELECTIVE 4A: BASICS OF STATISTICAL INFERENCE

CREDIT DISTRIBUTION, ELIGIBILITY, AND PRE-REQUISITES OF THE COURSE

Course title & Code	Credits	Credit distribution of the Course			Eligibility Criteria	Pre-requisite of the Course (if any)
		Lecture	Tutorial	Practical/ Practice		
Basics of Statistical Inference	4	3	0	1	Class XII with Mathematics	Basic knowledge of probability, probability distributions and sampling distributions

Learning Objectives:

The learning objectives of this course are as follows:

- To introduce the concept of estimation theory and testing of hypothesis.
- To infer about the unknown population parameters based on random samples.
- To introduce the estimation/ inference about the population using hypothesis testing.

Learning Outcomes:

After successful completion of this course, students will be able to:

- Understanding of estimation theory, Point and interval estimations.
- Characteristics of a good estimator and different methods of estimation.
- Demonstrate the use of these techniques in data analysis.
- Develop the best/most powerful statistical tests to test the hypotheses regarding unknown population parameters by using the Neyman-Pearson theory.

SYLLABUS OF GE 4A

Theory

UNIT I: **(15 Hours)**

Estimation Theory

Estimation: Parameter space, sample space, point estimation, requirement of a good estimator, consistency, unbiasedness, efficiency, sufficiency, Minimum variance unbiased estimators, Factorization theorem, Fisher- Neyman Criterion: statement and applications, Cramer- Rao inequality: statement and application, MVB estimators and their applications, Statement of Rao-Blackwell theorem and Lehmann-Scheffe theorem..

UNIT II: (15 Hours)

Methods of Estimation

Methods of estimation: maximum likelihood, least squares and minimum variance, Properties of maximum likelihood estimators (illustration), Interval Estimation: confidence interval and confidence limits for the parameters of normal distribution, confidence intervals for large samples.

UNIT III: (15 Hours)

Test of Significance

Principles of test of significance: Null and alternative hypotheses, simple and composite, Type-I and Type-II errors, critical region, level of significance, power of the test, best critical region, most powerful test, uniformly most powerful test, uniformly most powerful unbiased critical region (UMPU), Neyman- Pearson Lemma: statement and its applications to construct most powerful test.

PRACTICAL/LAB WORK – 30 Hours

List of Practical / Lab Work:

1. Unbiased estimators and consistent estimators.
2. Efficient estimators and relative efficiency of estimators.
3. Sufficient estimators and factorization theorem.
4. Cramer- Rao inequality and MVB estimators.
5. Method of maximum likelihood estimation.
6. Method of least squares and minimum variance.
7. Confidence interval and confidence limits for the parameters of normal distribution.
8. Confidence intervals in case of large samples.
9. Type I and Type II errors, power of the test.
10. Most powerful critical region (NP Lemma).

Practical work to be conducted using electronic spreadsheet / EXCEL/ Statistical Software Package/ SPSS/ calculators.

ESSENTIAL READINGS:

- Miller, I. and Miller, M. (2013). John E. Freund's Mathematical Statistics, 8th Ed., Prentice Hall of India.
- S.C. Gupta and V.K. Kapoor (2020): Fundamentals of Mathematical Statistics, 12th Ed., Sultan Chand and Sons.
- R.V. Hogg, A.T. Craig and J.W. Mckean (2005): Introduction to Mathematical Statistics, 6th Edition, Pearson Education.

● A.M. Goon, M.K. Gupta and B. Das Gupta (2003): An Outline of Statistical Theory (Vol. II), 4th Ed.,World Press, Kolkata.

SUGGESTED READING:

● G. Casella and R.L. Berger (2002): Statistical Inference, 2nd Edition, Thomson Duxbury.
● E.J. Dudewicz and S.N. Mishra (1988): Modern Mathematical Statistics, John Wiley and Sons.
● V.K. Rohtagi and A.K. Md. E. Saleh (2009): An Introduction to Probability and Statistics, 2nd Edition, John Wiley and Sons.
● Mood A.M., Graybill F.A. and Boes D.C. (1974). Introduction to the Theory of Statistics, McGraw Hill.

Note: Examination scheme and mode shall be as prescribed by the Examination Branch University of Delhi, from time to time.

GENERIC ELECTIVE 4B: STATISTICAL COMPUTING USING R

CREDIT DISTRIBUTION, ELIGIBILITY, AND PRE-REQUISITES OF THE COURSE

Course title & Code	Credits	Credit distribution of the course			Eligibility criteria	Pre-requisite of the course (if any)
		Lecture	Tutorial	Practical/ Practice		
Statistical Computation using R	**4**	**2**	0	**2**	Class XII pass with Mathematics.	Basic knowledge of computers and basics of Statistics

Learning Objectives:

The learning objectives of this course are as follows:

- Review and expand upon core topics in probability and statistics.
- Practice of graphical interpretation, probability distribution and data analysis using `R'.

Learning Outcomes:

After completing this course, students would have developed a clear understanding of:

- Various Graphical representation and interpretation of data.
- Automated reports giving detailed descriptive statistics.
- Understanding data and fitting suitable distribution.
- Testing of hypothesis, p-value and confidence interval.

- Random number generation and sampling procedures.
- Importing data, Code editing in R and flow controls if (), for (), while ()

SYLLABUS OF GE 4B

Theory

UNIT I **(07 hours)**

Overview of the R language

Installing R and R studio; working on R studio, scripts and text editors, creating and saving R workspaces, installing packages and loading libraries.

Data types in R (Numeric, Integer, Character, Logical, and Complex) Data structures in R (Vector, Matrix, Data frames, List). Mathematical operators, Relational Operators, and Logical operators and use of functions: class(), names(), head(), tail(),rbind(), cbind(), rownames(), colnames() etc. Learn how to load data, importing a data file viz. .xlsx. handling missing data in R

UNIT II **(10 hours)**

Descriptive statistics and Graphs

Generate automated reports giving detailed descriptive statistics mean, median, mode, variance, skewness, five-point summary , frequency table. Statistical/mathematical functions, scan(), summary(),str(), table(), cut(),cumsum(), cumprod()etc.

Graphical representation of data: bar-plot, pie-chart, boxplot, frequency polygon, ogives , scatter plot, Fitting of curve lm(): linear, quadratic, exponential functions, correlation, and linear and multiple regression with the interpretation of results.

UNIT III **(10 hours)**

Decision-making and distributions

Introduction to flow control: if, if-else, while, and for loops, simple coding. Distribution functions(r,d,p,q) for Binomial, Poisson, Exponential, and Normal . Data distribution: qqplot(), qqnorm()

UNIT IV **(08 hours)**

Testing of Hypothesis and Time series

Basics of statistical inference in order to understand hypothesis testing, and compute p-values and confidence intervals. Applications on t-test, F-test, and Chi-square test with the interpretation of results. Time series analysis, components of a time series data, time series model, ts(), decomposition(), and smoothing with the interpretation of results.

PRACTICAL/LAB WORK – 30 Hours

List of Practical / Lab Work:

1. Graphical representation of data with bar-plot, pie-chart, and boxplot.
2. Histogram with equal and unequal class intervals, frequency polygon
3. Less than and more than Ogives.
4. Fitting of curve linear, quadratic, exponential functions,
5. Scatter plots, correlation
6. Linear and multiple regression

7. Drawing sample using SRSWR, SRSWOR
8. Drawing sample using stratified under proportion allocation and systematic sampling,
9. functions(r,d,p,q) for discrete distributions viz. Binomial, Poisson.
10. functions(r,d,p,q) for continuous distribution viz. Uniform,Exponential, and Normal .
11. Test the goodness of fit for Binomial, Poisson distribution.
12. Chi- Square test for independence of attributes.
13. Single, paired and independent samples t-test.
14. Components of a time series data.
15. decomposition(), and smoothing() under time series data

ESSENTIAL READINGS:

- Braun, W. J., and Murdoch, D. J. (2007). A First Course in Statistical Programming with R. Cambridge University Press. New York.
- Gardener, M. (2012). Beginning R: The Statistical Programming Language, Wiley Publications.

SUGGESTIVE READING:

- Crawley, M. J. (2012). The R Book. 2nd Ed., John Wiley & Sons.
- Dalgaard, P. (2008). Introductory Statistics with R. 2nd Ed., Springer.

Note: Examination scheme and mode shall be as prescribed by the Examination Branch University of Delhi, from time to time.