Semester	**SEVEN**
Course	**Major**
Paper Code	**C4ST230741P**
Paper Title	**Advanced Data Analysis**
No. of Credits	**6**
Theory / Practical /Composite	**Practical**
Classes per week	**6** **Unit 1: 2 periods/week** **Unit 2: 2 periods/week** **Unit 3: 2 periods/week**
Module	**One**

Course Outcomes

1. **Remember** key concepts, definitions, assumptions, and formulas related to advanced sampling techniques, experimental designs, categorical data models, asymptotic methods, time series models, and demographic measures.
2. **Understand** the theoretical principles underlying systematic, cluster, multi-stage and PPS sampling, generalized linear models, large-sample inference, time series modeling, and demographic processes.
3. **Apply** appropriate statistical methods to solve problems in advanced sampling, experimental design, categorical data analysis, asymptotic testing, time series forecasting, and population studies.
4. **Analysis** of data using regression-based estimators, logit/probit/Poisson models, chi-square tests, Box-Jenkins models, and demographic techniques to identify patterns and relationships.
5. **Evaluate** the efficiency, validity, and suitability of sampling designs, experimental setups, inferential methods, time series models, and demographic projections for real-world data.
6. **Create** statistical models, forecasts, and interpretations for survey data, categorical responses, large-sample problems, time series behavior, and population estimation and projection.

Syllabus

Content	No. of lectures	CO mapping	Cognitive levels
UNIT 1: ***Categorical Data Analysis & Asymptotic Methods*** 1. Measures of association for 3x3 contingency table. 2. Conditional and marginal odds ratio, Homogeneous association. 3. Fitting a logit model, Confusion matrix, ROC & AUC, Goodness-of-fit measures.	26	CO1 CO2 CO3 CO4 CO6	K1 K2 K3 K4 K6

4. Fitting a probit model, Confusion matrix, ROC & AUC, Goodness-of-fit measure. 5. Fitting a Poisson regression model, Goodness-of-fit measure. 6. Tests of significance and confidence intervals concerning sample standard deviation, coefficient of variation and correlation coefficient (both single sample and two sample cases). 7. Tests of significance and confidence intervals using variance stabilizing transformations. 8. Tests for goodness of fit, independence and homogeneity using Pearsonian chi-square statistic			
UNIT 2: ***Advanced Time Series Analysis & Demography*** 1. Test for randomness of a residual series. 2. Fitting Box Jenkins models. 3. Exponential smoothing, Holt Winters Method. 4. Measures of mortality. 5. Life Tables. 6. Measures of fertility and population growth. 7. Population Estimation, Projection and Forecasting. 8. Fitting of logistic equation by Rhode`s method.	26	CO1 CO2 CO3 CO4 CO5 CO6	K1 K2 K3 K4 K5 K6
UNIT 3: ***Advanced Sampling Techniques and Experimental Design*** 1. Systematic Sampling: drawing of sample and estimation. 2. Ratio and Regression estimation: Estimation of population mean/ total and mean squares. Compare the efficiencies of ratio and regression estimators relative to SRS. 3. Cluster sampling: estimation of mean or total, variance of the estimate, estimate of intra-class correlation coefficient, efficiency as compared to SRS. 4. Two stage sampling. 5. Double Sampling. 6. RRT and Snowball sampling. 7. Probability Proportional to size sampling: drawing of sample and estimation of population total. 8. Problems on IBD 9. Problems on randomized response surface	26	CO1 CO2 CO3 CO4 CO5 CO6	K1 K2 K3 K4 K5 K6

Reading/Reference list

Same as the relevant theory papers

Evaluation

Total Marks	100
Mode of Exam	Continuous Assessment

CO	CO Description	Cognitive levels
CO1	**Remember** key concepts, definitions, assumptions, and formulas related to advanced sampling techniques, experimental designs, categorical data models, asymptotic methods, time series models, and demographic measures.	K1
CO2	**Understand** the theoretical principles underlying systematic, cluster, multi-stage and PPS sampling, generalized linear models, large-sample inference, time series modeling, and demographic processes.	K2
CO3	**Apply** appropriate statistical methods to solve problems in advanced sampling, experimental design, categorical data analysis, asymptotic testing, time series forecasting, and population studies.	K3
CO4	**Analysis** of data using regression-based estimators, logit/probit/Poisson models, chi-square tests, Box-Jenkins models, and demographic techniques to identify patterns and relationships.	K4
CO5	**Evaluate** the efficiency, validity, and suitability of sampling designs, experimental setups, inferential methods, time series models, and demographic projections for real-world data.	K5
CO6	**Create** statistical models, forecasts, and interpretations for survey data, categorical responses, large-sample problems, time series behavior, and population estimation and projection.	K6