GUJARAT TECHNOLOGICAL UNIVERSITY

Program Name: Master of Engineering

Level: PG

Branch: Civil Engineering (Transportation Engineering)

Course / Subject Code: ME01069041

Course / Subject Name: Statistical Techniques for Data Analysis

w. e. f. Academic Year:	2024-25
Semester:	1st Semester
Category of the Course:	PEC

Prerequisite :	Zeal to learn the subject
Rationale:	Statistics in engineering is used for designing experiments, analysing data, summarising and presenting information, and drawing reliable conclusions. It guides risk management, quality control, reliability analysis, and making informed decisions in design and operational processes. Statistics plays a crucial role in decision-making, planning, and interpreting results in Transportation engineering projects. The course typically begins with an introduction to basic concepts in probability theory, including sample spaces, events, and different approaches to defining probability. Emphasis is placed on developing a solid grasp of fundamental probability rules, such as the addition and multiplication rules, as well as understanding conditional probability, independence, and Bayes theorem. As the course progresses, students delve deeper into more advanced topics, such as random variables, probability distributions, and expectation. They explore common probability distributions, including binomial, Poisson, uniform, exponential, and normal distributions, and they learn how to calculate probabilities and expected values associated with these distributions. The course also equips students for measuring hypothesis testing. It also includes development of model and validation. It also discusses on curve fitting techniques and development of time series models.

PO

No	Program Outcomes
01	Engage in critical thinking and research to develop solutions to multifold real-world problems.
02	Communicate effectively with the engineering community at large level on complex design tasks & write and present technical reports.
03	Demonstrate a high level of professionalism in handling multidisciplinary and complex traffic engineering problems.
04	Plan, assess, create, integrate, carry out, and oversee complex transportation infrastructure

	projects in a sustainable local and global context.
05	Address societal issues pertaining to transportation by offering technologically advanced, reasonably priced solutions while upholding high standards of ethics and professionalism.

Course Outcome:

After Completion of the Course, Student will be able to:

No	Course Outcomes	RBT Level
01	Select a suitable method for processing and presentation of transportation data.	R
02	Apply probability distributions to analyze transportation data.	A
03	Choose appropriate hypothesis testing measures.	U
04	Analyze multivariate transportation data.	A
05	Differentiate various curve fitting techniques and develop Time Series models	A

**Revised Bloom's Taxonomy (RBT)*

Teaching and Examination Scheme:

Teaching Scheme (in Hours)			Total Credits L+T+ (PR/2)	Assessment Pattern and Marks				Total Marks
				Theory		Tutorial / Practical		
L	T	PR	C	ESE (E)	PA / CA (M)	PA/CA (I)	ESE (V)	
3	0	2	4	70	30	20	30	150

Course Content:

Unit No.	Content	No. of Hours	% of Weightage
1	Introduction: Statistics and Engineering, Basic concept – Population and Sample, Frequency distribution, Stem and Leaf displays, Descriptive measures, Quartiles and Percentiles	3	5
2	Probability: Sample space and Events, Counting, Probability, Axioms of probability, Elementary theorems of Probability, Conditional probability, Bayes' Theorem	3	10

3	Probability Distributions: Random variables, Binomial distribution, Hypergeometric distribution, Mean and variance of a probability distribution, Poisson distribution, Geometric and Negative Binomial distribution, Multinomial distribution	3	10
4	Probability Densities: Continuous random variables, Normal distribution, Other probability densities, Uniform distribution, Log-Normal distribution, Moment generating function, Sampling distribution of mean	3	10
5	Statistical Inference and Tests of Significance: Hypothesis testing, types of error in hypothesis, confidence interval, significance tests for comparing variances and means, tests with small and large samples, two-tail and one-tail student's t-test, analysis of variance (ANOVA), non-parametric tests (Chi-square test and Kolmogorov–Smirnov test), central limit theorem, practice with transportation data.	10	20
6	Sampling Techniques: Sample surveys, census, sampling bias, random sampling, stratified sampling, sequential sampling, cluster sampling, systematic sampling, sampling on successive occasions, non-sampling errors, applications in transportation engineering.	5	10
7	Regression and Correlation: Simple linear regression, residuals and variances, multiple linear regression, two-stage regression, forward, backward and step-wise regression, residual analysis, correlation analysis, type of correlations, coefficient of correlation, Karl-Pearson's coefficient, multivariate data analysis, factor analysis, applications in transportation engineering.	10	20
8	Parameter Estimation and Curve Fitting Techniques: Least square, generalized least squares, method of moments, maximum likelihood, algebraic and geometric curve fit, linear and non-linear curve fitting (polynomial, exponential, logarithmic, power, etc.), overfit, and under fit.	5	10
9	Time Series Models: Time series concept and components, utility, time series models, measurement of time series, graphical method, method of semi-average, moving average, least square, linear, parabolic and logarithmic trends, growth curves, ratio-to-trend and link relative method for seasonal variation, exercises with transportation data.	3	5
	Total	**45**	**100**

Suggested Specification Table with Marks (Theory):

Distribution of Theory Marks (in %)					
R Level	**U Level**	**A Level**	**N Level**	**E Level**	**C Level**
30	**20**	**30**	**10**	**10**	**--**

Where R: Remember; U: Understanding; A: Application, N: Analyze and E: Evaluate C: Create (as per Revised Bloom's Taxonomy)

(a) Books:

References/Suggested Learning Resources:

1. Probability and Statistics for Engineers-Miller, Freund-Hall, Prentice India Ltd.
2. Probability and Statistics for Engineers-Johnson Richard, Prentice India Ltd.
3. IntroductiontoProbability&StatisticsforEngineers&Scientists-Ross Sheldon, Elsevier Pub
4. Sampling techniques-Cochran, Wiley Series.
5. Statistics-Concepts and Controversies-David S. Moore-Freeman Company, New York.
6. Statistical and Econometric Methods for Transportation Data Analysis, Washington, S.P., Karlaftis, M.G., Mannering, F., Anastasopoulos, P., CRC Press, 2020, Third Edition.
7. Statistical Techniques for Transportation Engineering, Molugaram, K., Rao, G.S., Shah, A., Davergave, N., Butterworth-Heinemann, 2017, First Edition.
8. Multivariate Data Analysis, Joseph F.H., William C.B., Barry J.B., Anderson, R.E., Prentice Hall, 2018, Eighth Edition.
9. Probability and Statistical Inference, Robert V.H., Elliot, T., Zimmerman, D., Pearson, 2021, Tenth Edition.
10. Probability Concepts in Engineering Planning and Design, Alfredo H.S.A., Tang, W.H., Volume I & II, John Wiley & Sons, Singapore, 2007.
11. Quality Improvement through Statistical Methods, Bovas A., Springer Science & Business Media, 2012

(b) Open-source software and website:

List of Open-Source Software (May not be open source but useful for the subject):

Open Office, MATLAB, ANSYS, Excel and SPSS

Learning website:

http://courses.washington.edu/cee412/

https://ocw.mit.edu/courses/mathematics/18-05-introduction-to-probability

List of Experiments/Tutorials:

Minimum 20 problems from above topics, out of which half of the problems shall also be solved using self-developed computer programs in any language or software.

Assignment work on

- Probability distribution
- Sampling distribution
- Correlation
- Regression analysis
- Multivariate analysis
- Hypothetical testing

****