Course Code				**23STMAP233**			
Title of the Course				**Statistical Methods LAB**			
Offered to: (Programme/s)				**B.Sc.(Honors) - Statistics**			
L	**0**	**T**	**0**	**P**	**2**	**C**	**1**
Year of Introduction:		**2024-25**		**Semester:**		**3**	
Course Category:		**MAJOR**		**Course Relates to:**		**Local, Regional, National, Global**	
Year of Revision:		**NA**		**Percentage:**		**NA**	
Type of the Course:				**SKILL DEVELOMENT**			
Crosscutting Issues of the Course :				**NA**			
Pre-requisites, if any				**Basic Computers**			

Course Description:

This course focuses on the application of statistical techniques to analyze the relationship between two variables. Through hands-on exercises and data analysis, students will develop a strong foundation in bivariate analysis and regression techniques, preparing them for applications in various fields such as business, economics, and social sciences. It combines theoretical concepts with practical implementation using Microsoft Excel and SPSS.

Course Aims and Objectives:

S. No	COURSE OBJECTIVES
1	**Develop proficiency** in using MS Excel to generate bivariate random variable data.
2	**Apply the least squares method** to fit various curves (linear, quadratic, exponential, power) to data.
3	**Calculate and interpret** correlation coefficients (Pearson's and Spearman's) to measure the strength and direction of relationships between variables.
4	**Compute regression lines** for both ungrouped and grouped data to model the relationship between variables.
5	**Determine multiple and partial correlation coefficients** to assess the complex relationships among multiple variables.

.

Course Outcomes

At the end of the course, the student will be able to…

NO	COURSE OUTCOME	BTL	PO	PSO
CO1	create bivariate random variable tables using MS Excel and explore the relationship between the variables.	K6	2	2
CO2	apply curve fitting techniques to fit different types of curves (linear, quadratic, exponential, power) to data using the least squares method.	K3	2	2
CO3	apply Karl Pearson's and Spearman's rank correlation techniques to measure the strength and direction of relationships between variables.	K3	2	2
CO4	construct the regression lines for both grouped and ungrouped data and interpret the results.	K3	2	2
CO5	apply multiple and partial correlation methods to understand complex relationships among variables.	K3	2	2

For BTL: K1: Remember; K2: Understand; K3: Apply; K4: Analyze; K5: Evaluate; K6: Create

CO-PO-PSO MATRIX									
CO NO	**PO1**	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**	**PSO1**	**PSO2**
CO1		**3**							**3**
CO2		**3**							**3**
CO3		**3**							**3**
CO4		**3**							**3**
CO5		**3**							**3**

Use the codes 3, 2, 1 for High, Moderate and Low correlation Between CO-PO-PSO respectively

Course Structure

The lab list covers the key areas of a Statistical Methods, providing hands on practice with Excel & SPSS packages

Unit1: Bi-variate Random Variable table

Lab1: Create the Bi-variate random variable table with random number generation method in MS Excel.

1. **Random Number Generation:** Use software or hardware to generate random numbers that can be used to simulate bi-variate data. This will allow you to create controlled datasets for analysis.
2. **Joint Probability Distributions:** Determine joint probability mass functions (for discrete variables) or joint probability density functions (for continuous variables). This will provide a complete picture of the joint behavior of the variables.

Unit2: Curve Fitting - 1
Lab1: Fitting of the Straight line using principle of least squares method
Dataset: Real world data set
Experiment:

1. Determine the Line Equation
2. Calculate the constant values in model equation using principle of least squares method
3. Estimate the fitted values
4. Calculate the Residuals

Lab2: Fitting of the Second degree parabola using principle of least squares method
Dataset: Real world data set
Experiment:

1. Determine the Line Equation
2. Calculate the constant values in model equation using principle of least squares method
3. Estimate the fitted values
4. Calculate the Residuals

Unit3: Curve Fitting - 2
Lab1: Fitting of the Exponential curve using principle of least squares method
Dataset: Real world data set
Experiment:

1. Determine the Line Equation
2. Calculate the constant values in model equation using principle of least squares method
3. Estimate the fitted values
4. Calculate the Residuals

Lab2: Fitting of the Power curve using principle of least squares method
Dataset: Real world data set
Experiment:

1. Determine the Line Equation
2. Calculate the constant values in model equation using principle of least squares method
3. Estimate the fitted values
4. Calculate the Residuals

Unit4: Correlation Analysis

Lab1: Correlation between Height and Weight

Dataset: Real world data set
Experiment:

1. **Collect Data:** Measure the height and weight of a group of individuals. Record the data in a table.
2. **Create a Scatter Plot:** Plot the height values on the x-axis and the weight values on the y-axis.
3. **Calculate Correlation Coefficient:** Use the correlation coefficient formula (e.g., Pearson's correlation coefficient) to quantify the linear relationship between height and weight.

4. **Interpret Results:** Analyze the correlation coefficient to determine if there is a strong positive, strong negative, or weak correlation between height and weight.

Lab2: Correlation Between Movie Ratings and Box Office Success
Dataset: the ratings a movie receives from critics and its box office success.
Experiment:

1. **Rank Data:** Rank the movies based on their critic ratings and box office performance.
2. **Calculate Rank Differences:** For each movie, calculate the difference between its critic rating rank and its box office rank.
3. **Square the Differences:** Square each rank difference.
4. **Sum the Squared Differences:** Sum the squared rank differences.
5. **Calculate Spearman's Rank Correlation Coefficient:** Use the formula for Spearman's rank correlation coefficient to calculate the correlation between the two rankings.
6. **Interpret Results:** Analyze the Spearman's rank correlation coefficient to determine if there is a positive, negative, or no correlation between critic ratings and box office success.

Lab3: Correlation Coefficient between the Bivariate frequency data of two random variables.
Dataset: Real world data set
Experiment:

1. Record the data in a bi-variate frequency table.
2. **Calculate Correlation Coefficient:** Use the correlation coefficient formula (e.g., Pearson's correlation coefficient) to quantify the linear relationship
3. **Interpret Results:** Analyze the correlation coefficient to determine if there is a strong positive, strong negative, or weak correlation.

Unit5: Regression Analysis
Lab1: Predicting Sales Based on Advertising Expenditure
Dataset: To build a regression model to predict sales based on advertising expenditure.
Experiment:

1. **Plot the Data:** Create a scatter plot to visualize the relationship between advertising expenditure and sales.
2. **Calculate Regression Coefficients:** Use statistical software or formulas to calculate the slope (b1) and intercept (b0) of the regression line.
3. **Fit the Regression Line:** Write the equation of the regression line: $y = b0 + b1x$, where y is the predicted sales and x is the advertising expenditure.
4. **Evaluate the Model:** Assess the goodness of fit using metrics like R-squared, standard error of the estimate, and F-statistic.
5. **Make Predictions:** Use the regression equation to predict sales for different levels of advertising expenditure.

Lab2: Predicting Sales Based on Multiple Factors
Dataset To build a multiple regression model to predict sales based on multiple factors (e.g., advertising expenditure, price, and promotion).
Experiments:

1. **Collect Data:** Gather relevant data on sales and predictor variables.
2. **Build a Multiple Regression Model:** Use statistical software to fit a multiple regression model to predict sales based on the predictor variables.

3. **Evaluate the Model:** Assess the model's performance using metrics like R-squared, adjusted R-squared, F-statistic, and p-values for individual predictors.
4. **Calculate Partial Correlation Coefficients:** Calculate partial correlation coefficients to assess the unique contribution of each predictor variable to explaining the variation in sales, controlling for the effects of other predictors.

Text Book

1. Fundamentals of Mathematical Statistics, 12th Edition, 10th September 2020, S. C. Gupta and V. K. Kapoor, Sultan Chand & Sons, New Delhi.
2. Bernd Held., 2016, Microsoft Excel Functions & Formulas, Third Edition, Mercury Learning & Information.

Recommended References books:

1. Probability and Statistics,Volume I, D.Biswas, New central book Agency (P) Ltd, New Delhi.
2. An outline of Statistical theory, Volume Two,3rd Edition,2010(with corrections) A.M.Goon,M.K. Gupta, B.Dasgupta ,The World Press Pvt.Ltd., Kolakota.
3. Sanjay Arora and BansiLal:. New Mathematical Statistics, SatyaPrakashan , New Delhi.
4. https://www.tutorialspoint.com/excel_data_analysis/excel_data_analysis_tutorial.pdf

Semester End Lab Examination

23STMAP233: Statistical Methods

Offered to B.Sc. Hons STATISTICS

Max. Marks: 50 **Max. Time: 3Hrs**

Pass. Min: 20

(A)	**Evaluation Procedure**	**35 Marks**
I	Experiments (Exam & Execution)	**30 Marks**
II	Viva	**3 Marks**
III	Record	**2 Marks**
(B)	**CONTINUOUS ASSESMENT(Internal)**	**15 MARKS**

15 marks for the continuous assessment (Day to day work in the laboratory shall be evaluated for 15 marks by the concerned laboratory teacher based on the regularity/ record/viva). Laboratory teachers are mandated to ensure that every student completes 80%-90% of the lab assessments.

TOTAL: (A)+(B) = 50 MARKS