Undergraduate Program (with Statistics Major) Examination 2025
Semester-II
Statistics
Course: MJST03T (Descriptive Statistics II (Theory))
Full Marks: 60 Time: 3 Hours

Group A

(Answer any three questions. Each question carries 5 marks.)

1. How does intra-class correlation coefficient differ from ordinary correlation? 5

2. Let x and y be two variables having the same variances $s^2(> 0)$ and the correlation coefficient between them be r. Find all possible conditions under which the variables $u = ax+by$ and $v = cx-dy$, $(a, b, c, d > 0)$ are uncorrelated. 5

3. Explain the procedure of splicing two index number series into a single one. 5

4. Write a short note on 'homogeneity error in the measurement of price and quantity index and its control'. 5

5. Given n individual income values $y_1, y_2, \ldots y_n$ such that $y_1 < y_2 < \ldots < y_n$, show that the Gini coefficient can be expressed as $G = 1 + \frac{1}{n} - \frac{2}{n^2\overline{y}}\sum_{i=1}^{n}(n + 1 - i)y_i$, where $\overline{y}$ denotes the sample mean. 5

Group B

(Answer any three questions. Each question carries 15 marks.)

6. (a) What do you mean by cost of living index number? Can it be interpreted as a measure of standard of living?
   (b) Explain how we can compare the cost of living index number in two different situations.
   (c) Mention some uses of this index. (2+2)+7+4

7. (a) What do you mean by Lorenz curve? Prove that it is necessarily concave upwards.
   (b) Why is Pareto distribution suitable for graduating income distribution for the upper income group? Can we apply the log-normal distribution for the same?
   (c) What is Pareto's law? Given a frequency distribution of income, how can we justify whether the income variable actually follows a Pareto distribution? How can we estimate the Pareto coefficient from the given frequency distribution? (3+3)+(2+1)+(2+2+2)

8. (a) Give an example in each case where
      i. two variables have no relationship at all

(P.T.O)

ii. for two related variables, one of them depends on the other but the converse is not true.

(b) Find the point of intersection and angle between the two regression lines of y on x and x on y for a set of correlated bivariate data $(x_i, y_i), i = 1, 2, \dots n$. Interpret the special cases of (i) null correlation, (ii) perfect positive correlation and (iii) perfect negative correlation.

(c) Let there be k groups of paired data of sizes $n_j (j = 1, 2, 3, \dots k)$ on two variables (x, y), with means $(\overline{x_j}, \overline{y_j})$, variances (s^2_{xj}, s^2_{yj}) and correlation coefficient r_j, $j - 1, 2, \dots k$. Find the correlation coefficient r of the combined data of size $\sum_{j=1}^{k} n_j$ in terms of these values. Hence explain the phenomenon (if required, graphically) that r_j may be zero for each $j = 1, 2, \dots k$ and yet r may be non-zero. 2+6+7

9. (a) Let $Y = a + bx$ be the least square regression of y on x based on n pairs of observations $(x_i, y_i), i = 1, 2, \dots n$. If e stands for the least-square residuals, obtain the correlation coefficient between (i) Y and e, (ii) y and Y.

(b) Use the above result to bring out the significance and limitations of the correlation coefficient.

(c) Suggest a suitable measure of goodness of least square fit using a polynomial of degree $p \geq 1$, and examine the behavior of the measure as p increases. 6+3+6

10. (a) Let x and y be subject to observational errors, so that what one observes really are $x^* = x + \epsilon_x$ and $y^* = y + \epsilon_y$ instead of x and y, where $var(x) = s^2_x(> 0), var(y) = s^2_y(> 0)$ and $var(\epsilon_x) = var(\epsilon_y) = s^2(> 0)$. If ϵ_x and ϵ_y are uncorrelated both with x and y and they are also mutually uncorrelated, find the expression of the correlation coefficient between x^* and y^* in terms of the correlation coefficient between x and y. What can be said about their relative magnitudes? Comment on your finding.

(b) Consider a variable y which is related to x as

$$y = 1 - \frac{1}{1 + \exp(a + bx)}$$

where a and b are unknown constants. Apply a suitable transformation and derive the least square estimates of a and b on the basis of n pairs of observations on (x, y) with y as the dependent variable.

(c) Define Kendall's rank correlation coefficient. Find the lower and upper limits of the measure and explain the situations when the limits are attained. 5+5+5

Undergraduate Program with Statistics Major Examination, 2025
Semester II
Statistics (Practical)
Course: MJST01P (Descriptive Statistics II (Practical))
Full Marks: 20 Time: 2 Hours

(1) The following table shows the heights X (in cm) and weights Y (in kg) of 15 students:

Height X	152.5	157.5	160.0	170.0	175.0
Weight Y	56.0	56.5	55.0	55.0	58.0
	61.5	59.0	62.5	65.0	63.0
	57.5		50.0		61.0
	58.5		56.5		

(a) Calculate the correlation coefficient of weight and height.

(b) Find the regression equation of weight on height. Also Estimate the weight of a student of height 172 cm.

(c) Calculate the correlation ratio of weight on height. 3+3+3

(2) The heights in inches of 3 brothers belonging to each of six families are recorded below. Compute intra-class correlation coefficient. 4

Family	1	2	3	4	5	6
Height of brothers	69.5	71.2	65.6	62.2	68.0	64.4
	70.6	70.8	67.2	63.6	70.5	64.3
	72.3	72.0	66.3	63.5	70.5	64.6

(3) On a certain date the Ministry of Labour price index was 204.6. Percentage increases in price over some base period were: *House Rent* 65, *Clothing* 220, *Fuel and Light* 110, *Miscellaneous* 125. The weights of the different items in the group were as follows: *Food* 60, *House Rent* 16, *Clothing* 12, *Fuel and Light* 8, *Miscellaneous* 4. Find the percentage increase in price in the food group. 4

(4) Practical Note Book and Viva-voce. 3

Four Year Undergraduate Programme
Semester-II Examination 2025

Subject: Statistics Major
Paper: MJST04T

Probability and Probability Distributions II (Theory)

Full Marks: 60 **Time: 3 hours**

*Answer any **four** out of the following six questions of equal marks.*
Notations are of usual meanings.

1. (a) Derive the moment generating function (MGF) of the binomial distribution. Using it or otherwise, obtain the mean and variance of the distribution. (5)
   (b) Let $X_1 \sim P(4)$, $X_2 \sim P(5)$, and they are independent. Find the probability distribution of $X_1 + X_2$. (5)
   (c) If a fair die is rolled 8 times, find the probability distribution of the number of times a '6' appears. Define the distribution clearly. (5)

2. (a) Define Weibull distribution and find its mean and variance. (6)
   (b) Derive the mode of the Poisson distribution. (6)
   (c) Explain the memoryless property of the exponential distribution. Name the discrete distribution which satisfies this property. (3)

3. (a) Define the bivariate normal distribution and find the marginal distributions of the random variables involved in the bivariate random vector. (2+6)
   (b) Given the joint PDF:

$$f(x,y) = \begin{cases} 6(1-x-y), & 0 < x < 1, 0 < y < 1, x+y < 1 \\ 0, & \text{otherwise} \end{cases}$$

   Find:
   i. Marginal distributions of X and Y.
   ii. Conditional distribution of Y given $X = x$. (7)

4. (a) Derive the mean deviation about the mean of the normal distribution (8)
   (b) Define log-normal distribution and find its mean and variance. (7)

5. (a) Define convergence in probability and convergence in distribution of a sequence of random variables, and state their relationship. (5)
   (b) Suppose $X_n = \frac{1}{n}Z$, where $Z \sim N(0,1)$. Show that $X_n \to 0$ in probability. Does X_n converge in distribution? Justify. (5)
   (c) Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables, where each $X_i \sim \text{Uniform}(0,1)$. Then show that

$$\lim_{n\to\infty} P\left(\left|\bar{X}_n - \frac{1}{2}\right| > \varepsilon\right) = 0$$

   . (5)

6. (a) State Lindberg-Levy and Liapounov's Central Limit Theorems (CLTs). (4)

(b) Let $X_1, X_2, ..., X_n$ be i.i.d. with $\mu = 100$, $\sigma^2 = 64$. Use CLT to estimate $P(98 < \bar{X} < 102)$ for $n = 64$. Given $P(0 < z < 2) = 0.4772$ where $Z \sim N(0, 1)$ distribution. (5)

(c) Let $X_1, X_2, \ldots, X_{120}$ be i.i.d Poisson random variables with parametr $\lambda = 3$. Find the approximate value of $P(\sum_{i=1}^{120} > 390)$ given that $\Phi(1.6) \approx 0.9460$. (6)

Four Year Undergraduate Programme
Sem-II Examination 2025

Subject: Statistics Major
Paper: MJST04P
Probability and Probability Distributions II (Practical)

Answer the following six questions of equal marks. Nonations are of usual meanings. Necessary tables will be provided in the examination hall

Full Marks: 20 Time: 2 hours

1. The probability of winning a game is 0.4. If a person plays the game 12 times, find the probability that he wins more than 8 times.

2. In a book, printing errors occur at an average rate of 2 per 100 pages. Find the probability of exactly 3 errors in 150 pages.

3. A normal population has mean 100 and variance 25. Find the value of x such that 80% of the population lies below x.

4. A machine produces sheets of length uniformly distributed between 49 cm and 51 cm. Find the probability that a sheet is longer than 50.5 cm.

5. For a bivariate normal distribution with $\mu_X = 20$, $\mu_Y = 30$, $\sigma_X = 5$, $\sigma_Y = 6$, and $\rho = 0.4$, find $P(Y > 35 \mid X = 22)$.

6. Viva-voce and practical note-book.

4 Year UG Programme (with Statistics Major) Examination 2025
Semester II
Subject: Statistics
Paper: [SECST02] Introduction to R

Full Marks: 60 Time: 3 Hrs.

Group – A (Answer any ten questions) $10 \times 3 = 30$

1. Answer the following questions with proper justification. Provide code /pseudo-code wherever required.

   (a) Create two 3×3 matrices and multiply them using R.

   (b) Write $\sin(\log \pi)$ and show five significant digits.

   (c) Write a function which takes the vector [1,2,5,8,4,9] as input and their mean, sd and median, as output.

   (d) Using R, create a data frame of five students with their name, age, gender and location.

   (e) Write down the usage of the functions: Apply, Sapply, Tapply.

   (f) Suppose you have an array of numbers starting from 10 to 1200. Using R, how to find the numbers lying between 250 to 500?

   (g) Using loop find the median, quartiles and interquartile range for the data: 7 22 13 37 14 18 18 13 17 24 18.

   (h) Using loop find greatest n such that: $1^5 + 2^5 + 3^5 + \cdots + n^5 \leq 10^8$.

   (i) Generate 1000 random samples from each of $\chi^2_{(4)}$ and $U[2, 3]$ distributions.

   (j) Suppose there are 1200 students in Zoology. Among them 800 are girls and 400 are boys. You want to select 60 among them for a conference with same proportion as in the population. Use Stratified Sampling to select these lucky students.

   (k) Create a code that can identify a three digit number as an Armstrong number.

   (l) Create a function that can show the final position of a one dimension random walk (left/right) after n steps.

   (m) Using R, draw the plot of $\sin^2(x)$ and $\sin(x^2)$ in a single graph for $x \in [-3, 3]$.

Group – B (Answer any five questions) $5 \times 6 = 30$

2. Plot the multiple graphs of normal distribution $N(0.4, 1.2)$, $N(0.4, 0.6)$ and $N(0.4, 1.8)$.

3. Suppose you have 3 sets of data (A): 1.78 2.88 0.90 1.16 -0.16 4.19 2.26 1.28 0.86 1.76 0.74 2.32 1.84 2.23 1.34 1.70 4.62 3.20 2.40 0.60; (B): 0.88 3.02 2.96 1.10 4.33 1.25 2.90 2.82 2.31 1.44 2.44 1.49 1.95 3.91 0.03 1.09 8.78 2.42 0.27 1.64; (C): 0.99 0.88 0.99 0.84 0.84 0.86 0.98 0.87 0.89 0.91 0.98 0.92 0.68 0.72 0.86 0.92 0.96 0.89 0.90 0.91. Compute various measures of skewness to check whether these datasets are skewed or not. Compare the boxplots of A, B, C.

4. Suppose we have a bivariate data of 20 students: Marks in School (x): 148 134 131 146 135 159 136 161 166 165 160 136 176 145 144 153 158 145 126 170 and Marks in College (y): 423 405 362 369 333 417 301 425 372 438 415 393 349 306 380 338 450 326 381 359. Plot the scatter diagram. Also write down the correlation coefficient between x, y and the correlation matrix.

5. Suppose the marks of 100 students in a class is as follows:
37 52 07 52 35 52 68 42 66 82 14 59 70 75 74 38 51 38 30 69 52 64 67 53 59 61 65 68 45 41 70 25 89 42 59 52 84 45 25 64 65 71 34 77 86 81 70 74 68 72 10 15 68 30 30 57 62 34 74 51 60 27 51 77 52 54 65 27 37 68 80 94 50 36 43 44 51 54 71 35 44 55 58 47 75 28 45 46 44 45 81 12 81 92 56 60 36 65 49 78.
Draw a histogram on the data. Choose the class width of your own. Also draw the normal approximation curve.

6. The following data represent number of telephone calls received in a particular hour

No. of Calls	0	1	2	3	4	5	6	7	8	9 & more
Frequency	7	33	54	38	35	15	7	4	1	1

Using R, fit a poisson distribution when λ is unknown.

7. Using loop find all primes in the interval [100,1500].

Undergraduate Program (with Statistics Major) Examination 2024
Semester-II
Statistics
Course: MJST03T (Descriptive Statistics II (Theory))
Full Marks: 60 Time: 3 Hours

Group A

(Answer any three questions. Each question carries 5 marks.)

1. What is scatter diagram? Explain its role to derive the working formula of Pearson's correlation coefficient. 2+3

2. If the regression lines of Y on X and X on Y are respectively $a_1x + b_1y + c_1 = 0$ and $a_2x + b_2y + c_2 = 0$, then show that $a_1b_2 \leq a_2b_1$.

   cm 5

3. Prove that the correlation index is a non-decreasing function of its order. 5

4. Prove that Edgeworth-Marshall's price index number lies between Laspeyres' and Paasche's index. 5

5. Let X and Y have zero means and unit variances and the correlation coefficient between $aX + bY$ and $bX + aY$ is $\dfrac{(1 + 2ab)}{(a^2 + b^2)}$. Find the correlation coefficient between X and Y. 5

Group B

(Answer any three questions. Each question carries 15 marks.)

6. (a) What do you mean by time reversal test and factor reversal test?
   (b) Prove that none of these tests are satisfied by Laspeyres' and Paasche's price index numbers. Further show that both of them are satisfied by Fisher's ideal index number.
   (c) If L_p, P_p and L_q denote, respectively Laspeyres' price index, Paasche's price index and Laspeyres' quantity index, show that $L_q(P_p - L_p)$ may be looked upon as the weighted covariance between price and quantity relatives, the weights being the base year values. Hence justify the statement that Laspeyres' price index usually exceeds Paasche's.

   4+4+7

7. (a) Define Gini coefficient of concentration. Write down its mathematical expression.
   (b) Prove that for a grouped data with k income groups arranged from the bottom of the distribution, it can be expressed as $G = 1 + \sum_{i=1}^{k} p_i \left(z_i + z_{i-1}\right)$, where p_i and z_i respectively represent the population share of the ith income group and the cumulative share of income upto the ith income group $(i = 1, 2, \ldots k; z_0 = 0)$

(c) Show that if the income variable follows a Pareto distribuion with Pareto coefficient ν, then the Gini coefficient is equal to $\dfrac{1}{2\nu - 1}$ (2+2)+5+6

8. (a) Prove that $0 \leq r^2 \leq {e^2}_{yx} \leq 1$, where the symbols have their usual meanings. Discuss in detail the cases of equality.

   (b) Geometrically interpret the situations:

      i. $r = 0, e_{yx} = 1$

      ii. $r = 0, e_{yx} = e_{xy} = 1$

   (c) Define Kendall's rank correlation coefficient. Explain how it can be interpreted as a product-moment correlation coefficient. (6+3)+2+(2+2)

9. (a) What do you mean by intra-class correlation? How does it differ from ordinary correlation?

   (b) Suppose you have p families each containing k members. Denote the height of the jth member of the ith family by $y_{ij}(i = 1, 2, \ldots p, j = 1, 2, \ldots k)$. Derive the working formula of intra-class correlation in this case.

   (c) Find limits of the measure. Interpret the marginal cases. Why is it called a skew coefficient? (2+2)+5+(2+2+2)

10. Write short notes on any three of the following: 5×3

   (a) Lorenz curve and its applications

   (b) Merits and demerits of chain-base index numbers

   (c) Limitations of correlation coefficient

   (d) Deflation of index numbers

   (e) Least square principle.

Undergraduate Program with Statistics Major Examination, 2024
Semester II
Statistics (Practical)
Course: MJST01P (Descriptive Statistics II (Practical))
Full Marks: 20   Time: 2 Hours

(1) Calculate Spearman’s rank correlation coefficient for the following data on marks of eight students in Statistics and Accountancy. 5

Student No.	1	2	3	4	5	6	7	8
Marks in Statistics	52	65	45	38	72	65	45	25
Marks in Accountancy	61	52	52	25	79	43	61	33

(2) For a set of 10 pairs of values of X and Y, the two regression lines are $x + 2.112y = 20.16$ and $x + 2.357y = 21.83$, and the standard deviation of Y is 14. Later, two pairs $(X = 3, Y = 8)$ and $(X = 5, Y = 4)$ are replaced by $(X = 8, Y = 3)$ and $(X = 4, Y = 5)$. Find the two regression lines for the new set of 10 pairs. Also determine the proportion of variability explained by the regression lines. 5+2

(3) Owing to the change in prices, the consumer price index of the working class in a certain area rose in a month by one quarter of what it was before and became 225. The index of food became 252 from 198, that of clothing from 185 to 205, that of fuel and lighting from 175 to 195, and that of miscellaneous from 138 to 212. The index of house rent, however, remained the same at 150. It was known that the weights of clothing, house rent and fuel and lighting were the same. Find out the exact weights of all the groups. 5

(4) Practical Note Book and Viva-voce. 3

Four Year Undergraduate Programme
Sem-II Examination 2024

Subject: Statistics Major
Paper: MJST04T
Probability and Probability Distributions II (Theory)

Answer any **four** out of the following six questions of equal marks.
Nonations are of usual meanings.

Full Marks: 60 Time: 3 hours

1. (a) Show that Poisson distribution is a limiting case of the binomial distribution.
   (b) Define geometric distribution and find its mean and variance.
   (c) 15 unbiased coins are tossed and numbers of 'Head' appeared are noted. If X be the random variable giving the number of 'Head' appeared. Derive the probability distribution of X.

   5+5+5

2. (a) Suppose the lifetime of a certain type of electronic component follows an exponential distribution with a mean lifetime of 5 years. Calculate the probability that a component will last more than 7 years. Additionally, given that a component has already lasted 4 years, determine the probability that it will last at least 3 more years.
   (b) Define gamma distribution and find its mean and variance.
   (c) Describe the relationship between the exponential distribution and the gamma distribution.

   6+6+3

3. (a) Define bivariate normal distribution. Show that the conditional expectation of X given $Y = y$ is a linear equation of y.
   (b) Consider a bivariate distribution with the joint PDF:

$$f(x, y) = \begin{cases} (x + y) & \text{if } 0 \leq x \leq 1 \text{ and } 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

   Find the conditional distribution of Y given $X = x$.

(2+8)+5

4. (a) Derive the r^{th} order central moment of a $N(\mu, \sigma^2)$ variable. Show the final expressions for both cases when r is an even and odd number.

   (b) Define *Beta* distribution of Type-I and find its mean and variance.

8+7

5. (a) Define convergence in probability and convergence in distribution of sequence of random variables.

   (b) Consider a sequence of random variables X_n where $X_n = \frac{Y_n}{n}$ and Y_n is a sequence of independently and identically distributed random variables with mean 0 and variance σ^2. Prove that X_n converges to 0 in probability but does not necessarily converge to 0 in distribution.

   (c) Give two well-known examples of convergence in distributions.

4+8+3

6. (a) State the Central Limit Theorem (CLT) for a sequence of independent and identically distributed (i.i.d.) random variables. Explain the conditions under which the CLT holds.

   (b) Suppose $X_1, X_2, \ldots, X_n$ are independently and identically distributed random variables with mean $\mu = 50$ and variance $\sigma^2 = 25$. Let $\overline{X_n}$ be the sample mean. Using CLT approximate the probability that the sample mean $\overline{X_n}$ falls between 48 and 52 for $n = 100$.

7+8

Four Year Undergraduate Programme
Sem-II Examination 2024

Subject: Statistics Major
Paper: MJST04P
Probability and Probability Distributions II (Practical)

Answer the following five questions of equal marks. Nonations are of usual meanings.

Full Marks: 20 Time: 2 hours

1. A certain drug is effective in 70% of the cases. A doctor administers the drug to 8 patients.
   (a) What is the probability that exactly 5 patients will respond to the drug?
   (b) Find the probability that fewer than 3 patients will respond to the drug.
2. A bookstore receives an average of 3 online orders per hour.
   (a) What is the probability that exactly 1 order is received in an hour?
   (b) Calculate the probability that no more than 2 orders are received in an hour.
3. The length of time (in hours) that a battery lasts is normally distributed with a mean of 500 hours and a standard deviation of 50 hours.
   (a) Calculate the probability that a battery lasts less than 450 hours.
   (b) Determine the value below which 10% of the batteries are expected to last.

4. Let X and Y are jointly normally distributed with means $\mu_X = 100, \mu_Y = 80$, variances $\sigma_X^2 = 36, \sigma_Y^2 = 49$ and correlation coefficient $\rho = 0.4$

   (a) Find the regression line of X on Y.

   (b) If $Y = 85$ what is the expected value of X?

5. Viva-voce and practical note-book.

4-Year Undergraduate Program
with Statistics major
Examination 2024
Semester II
Subject: Statistics
Paper: [SECST 02]

Full Marks: 60 **Time: 3 Hrs.**

Group A (Answer any ten questions) **$10 \times 3 = 30$**

1. Answer the following questions with proper justification. Provide code /pseudo-code wherever required.

(a) Name some popular R packages for data analysis.

(b) Generate a box plot of a numeric variable grouped by a categorical variable in a dataset.

(c) What are vectors in R, and how are they different from lists?

(d) How do you perform matrix multiplication in R?

(e) Create a bar plot of the frequency of a categorical variable in a dataset (OR) Generate a violin plot of a numeric variable grouped by a categorical variable in a dataset.

(f) Create a user-defined function that inputs x, y from the user and generates the value $\sin(x + y)$.

(g) Write a function in R to calculate the factorial of a number.

(h) How do you input data from an Excel file into R? How do you input data from a text file into R?

(i) Create a vector `x` with elements $0, 0.4, 0.8, 1.2, \ldots, 8$.

(j) Identify and remove rows with missing values in a dataset in R. Check the structure and summary of a dataset in R.

(k) Write a function in R to calculate the greatest common divisor (GCD) of two numbers

(l) Find the correlation coefficient between two numeric vectors in R.

(m) Check the first few rows of a dataset in R. Identify duplicate rows in a dataset in R.

Group B (Answer any five questions) **$6 \times 5 = 30$**

2. Write an R function to calculate a numeric vector's mean, median, mode, variance, skewness, and kurtosis from the first principle without using a built-in function in R. Illustrate with the vector `c(4, 2, 1, 5, 8, 15, 16, 23, 42)`.

3. Use an `if-else` statement to check if a number is positive, negative, or zero. Use a `while` loop to calculate the factorial of a given number.

4. Write a function that checks whether a given number is prime.

5. Write an R code to generate the first 20 numbers in the Fibonacci sequence. Write the function to plot them.

6. Create a vector of 25 random numbers from a Poisson distribution and 25 from a normal distribution. Plot a histogram of these numbers.

7. Generate a sequence of numbers from 0 to 2π and write a function to plot the sine function over this range.

B.Sc. (Honours) Sem-II, Examination 2024

Subject:Statistics

Paper: [CC-3]

Probability and Probability Distribution (Theory)

Time: 3 hours Full Marks: 40

Answer any 4 questions out of following 6 questionsof equal marks:

Notations carry usual meanings.

1. (a) A discrete random variable X has the following probability mass function (pmf):

$$P(X = x) = \begin{cases} \frac{1}{6} & \text{for } x = 1, 2, 3, 4, 5, 6 \\ 0 & \text{otherwise} \end{cases}$$

Find the expected value, $E(X)$, of the random variable X. Also calculate the variance $Var(X)$ of the random variable X.

(b) Prove the memory less property of the exponential distribution.

2. (a) Let X be a continuous random variable with the probability density function (pdf):

$$f(x) = \begin{cases} 2x & \text{for } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Compute the cumulative distribution function (CDF) $F(x)$ of X and find the probability $P(0.5 \leq X \leq 0.7)$.

(b) Write down the pdf of the gamma distribution and find its mean and variance.

3. A normally distributed random variable Y has a mean of μ and a standard deviation σ. Find the moment generating function of Y. Hence or otherwise find its skewness and kurtosis.

4. (a) If a binomial random variable Z has parameters $n = 10$ and $p = 0.5$, find the probability of exactly 4 successes.

(b) Find the mean and variance of the Poisson distribution.

5. Consider two discrete random variables X and Y with the joint probability mass function given by:

$$P(X = x, Y = y) = \begin{cases} \frac{1}{12} & \text{for } (x, y) = (1, 1), (1, 2), (2, 1), (2, 2) \\ 0 & \text{otherwise} \end{cases}$$

Find the marginal probability mass functions $P(X = x)$ and $P(Y = y)$. Also find $E(X/Y = 2)$.

6. (a) A box contains 5 red balls and 7 blue balls. Two balls are drawn sequentially without replacement.

   i. What is the probability that both balls drawn are red?

   ii. Given that the first ball drawn was red, what is the probability that the second ball drawn is also red?

   (b) Consider two events A and B with probabilities $P(A) = 0.4$, $P(B) = 0.5$ and $P(A \bigcap B) = 0.2$.

   i. Are the events A and B independent?

   ii. Find the value of $P(A \cup B)$.

B.Sc. (Honours) Sem-II, Examination 2024

Subject:Statistics

Paper: [CC-3B]

Probability and Probability Distribution (Practical)

Time: 2 hours — Full Marks: 20

Answer the following questions:

Notations carry usual meanings.

1. In a certain game, a player has a 60% chance of winning each round. The player plays 10 rounds.
   (a) Calculate the probability that the player wins exactly 7 rounds.
   (b) Determine the probability that the player wins at least 8 rounds.
2. The heights of adult men in a certain population are normally distributed with a mean of 175 cm and a standard deviation of 10 cm.
   (a) Find what proportion of men are taller than 185 cm?
   (b) Calculate the probability that a randomly selected man is between 170 cm and 180 cm tall.
3. Consider the joint probability distribution of two discrete random variables X and Y:

$X \backslash Y$	1	2	3
1	0.1	0.2	0.1
2	0.2	0.1	0.3

   (a) Find the marginal probability distribution of X.
   (b) Calculate the covariance between $X$and Y.
4. A call center receives an average of 5 calls per minute.
   (a) What is the probability that exactly 3 calls will be received in a given minute?
   (b) Calculate the probability that more than 6 calls will be received in a given minute.
5. Viva-voce and practical note-book.

B.Sc. Semester II Examination 2023
Subject: Statistics
Paper: CC-3
(Probability and Probability Distributions)

Full Marks: 40 Time: 3 hours

Answer any four of the following six questions of equal marks.
(Notations carry usual meanings)

1. (a) Suppose that each day the weather can be uniquely classified as 'fine' or 'bad'. Suppose further that the probability of having fine weather on the last day of a certain year is P_0 and we have the probability p that the weather on an arbitrary day will be of the same kind as on the preceding day. Let the probability of having fine weather on the n^{th} day of the following year be P_n. Show that

$$P_n = (2p-1)P_{n-1} + (1-p)$$

Deduce that

$$P_3 = (2p-1)^3(P_0 - \frac{1}{2}) + \frac{1}{2}$$

(b) Let a function be

$$p(x) = \begin{cases} \frac{x}{15}, & \text{if } x = 0,1,2,3,4,5. \\ 0 & \text{otherwise} \end{cases}$$

i. Do you think $p_X(x)$ is a p.m.f of X?
ii. Find $P(X = 3 \text{ or } 4)$
iii. Find $P[\frac{1}{2} < X < \frac{5}{2}/X > 1]$

5+5

2. (a) In a factory 2 machines M_1 and M_2 are used for manufacturing screws which may be uniquely classified as good or bad. M_1 produces per day n_1 boxes of screws, of which on the average, p_1% are bad while the corresponding numbers for M_2 are n_2 and p_2. From the total production of both M_1 and M_2 for a certain day, a box is chosen at random, a screw is taken out of it. It is found to be bad. Find the probability that the selected screw is manufactured by the machine M_2.

(b) Let the joint pdf of X and Y be

$$f_(x,y) = \begin{cases} 2(x+y-3xy^2), & \text{for } 0<x<1; 0<y<1 \\ 0, & \text{otherwise} \end{cases}$$

Find $Cov(X,Y)$ and $Var(X)$.

4+6

3. (a) Define 'mutual independence' and 'pair-wise independence' of events. Prove that mutual independence implies paire-wise independence but the converse may not be true.

   (b) Prove that if X and Y are independent random variables then $E(XY) = E(X)E(Y)$.

6+4

4. (a) Define Geometric distribution and deduce its moment generating function. Hence find the mean and variance of the distribution.

   (b) Suppose for a random variable X , $M_X(t) = \frac{pe^t}{1-qe^t}$ where p and q are constants. Find the m.g.f. of $(3X + 2)$.

6+4

5. (a) Deduce the expression for r^{th} order central moment of the normal distribution.

   (b) Find the moment generating function of the random variable whose moments are $\mu'_r = (r + 1)!2^r$.

5+5

6. (a) Let X follow exponential distribution with mean λ. Find

$$\frac{P[X > s + t]}{P[X > t]}$$

   Interpret the result.

   (b) The distribution function of a random variable X is given by

$$F_X(x) = \begin{cases} 1 - (1 + x)e^{-x}, & \text{for } x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

   Find the probability density function of X. Hence find $E(2X - 3)$.

5+5

B.Sc. Semester II Examination 2023
Subject: Statistics
Paper: CC 3B
(Probability and Prob. Distributions-Practical)

Full Marks: 20 Time: 2 hours

Answer <u>all</u> the following questions.
(Notations carry usual meanings)

1. Let a two dimensional variable (X, Y) have the joint probability distribution as follows.

Y \ X	1	2	3
-1	$\frac{1}{9}$	$\frac{1}{18}$	$\frac{1}{18}$
0	$\frac{1}{18}$	$\frac{2}{9}$	$\frac{3}{9}$
1	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$

Answer the following questions.

(a) Find $P(X > 1\, and\, Y = 0)$

(b) Are X and Y independent?

(c) Find $Cov(X, Y)$.

(d) Find $V(X - Y)$.

2+1+2+2

2. Data was collected from a pathological laboratory which shows the frequencies of number of red blood corpuscles$x$ per cell.

x	0	1	2	3	4	5
$f(x)$	142	156	69	27	5	1

Fit an appropriate distribution to the above data citing reasons and comment on the fitting.

6

3. The monthly incomes of a group of 10,000 persons were found to be normally distributed with mean Rs.750 and s.d. Rs.50. Show that of this group, about 95% had income exceeding Rs.668 and 5% had income exceeding Rs.832.

4

4. Viva-voce and Practical Notebook

3

B.Sc. Examination, 2023
Semester-II
Statistics
Course: CC-4
(Algebra)

Time: 3 Hours **Full Marks: 60**

Questions are of value as indicated in the margin.
Notations have their usual meanings

Group-A (Answer any five questions)

1. Answer any five questions. [5 × 2]

(a) Explain whether $det(A^5) = (det(A))^5$, where $A_{n\times n}$ is a square matrix.

(b) Define the rank of a matrix. When do we say a set of vectors to be linearly independent?

(c) Let $P(x)$ be a polynomial of degree n. If $P(a) = 0$ then show that $P(x) = (x - a) \cdot Q(x)$ for some polynomial $Q(x)$.

(d) Illustrate a suitable use of row-reduced echelon form with an example.

(e) A is a 3×3 non-null real matrix and $A^2 - A - I_3$ is a null matrix. Show that A^{-1} exists and $A^{-1} = A - I_3$.

(f) What will be the degree of the polynomial obtained by the product of two different polynomials of degree m and n?

(g) Define the determinant of a matrix.

(h) State Vieta's Theorem.

Group-B (Answer any five questions) [5 × 10]

2. (a) Let A be an $n \times n$ matrix. Show that the following are equivalent (TFAE):

(i) A is invertible

(ii) $det(A) \neq 0$

(iii) The number 0 is not an eigenvalue of A.

(b) For which $n \in \mathcal{N}$ is $(x^2 + x + 1) \mid (x^{2n} + x^n + 1)$ (where $\mathcal{N}$ is set of natural numbers)?

(c) For which $n \in \mathcal{N}$ is $37 \mid 1\underbrace{0\cdots0}_{\text{n times}}1\underbrace{0\cdots0}_{\text{n times}}1$? (you may use (2b), or otherwise).

[4 + 4 + 2]

3. (a) Show that the rank of an orthogonal matrix of order n is n.

(b) Show that basis is not unique for a vector space.

(c) If $(I + A)^{-1}(I - A)$ is skew-symmetric matrix, then show that A is an orthogonal matrix.

[2 + 4 + 4]

4. (a) Define the dimension of a vector space with a suitable example.

(b) Find reduced row-echelon form of $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}$.

(c) Describe The Gram-Schmidt Orthogonalization procedure.

.

[2 + 4 + 4]

5. (a) Let $A_{m \times n}$ be a real matrix. Show that rank of A is the same as the rank of $A^T A$.

(b) If A and B are two $n \times n$ matrices and A has an inverse, then show that

$$(A + B)A^{-1}(A - B) = (A - B)A^{-1}(A + B)$$

(c) Investigate for what values of a, m the following system of equations

$$\begin{aligned} x + y + z &= 1 \\ x + 2y - z &= m \\ 5x + 7y + az &= m^2 \end{aligned}$$

has only one solution, no solution, or infinitely many solutions.

[2 + 4 + 4]

6. (a) Define Similar matrix and Idempotent matrix.

(b) Prove that, if A and B are two matrices such that $AB = A$ and $BA = B$ then A, B are idempotent.

(c) Let A be an idempotent matrix. Show that all eigenvalues of A are either 0 or 1.

[2 + 4 + 4]

7. (a) Define the eigenvalue and eigenvector of a square matrix.

(b) Find eigenvalues of the following matrix. Also, find the algebraic multiplicity of each of the eigenvalues.

$$\begin{bmatrix} 2 & 0 & 0 & 0 \\ 5 & 3 & 0 & 0 \\ 9 & 1 & 3 & 0 \\ 1 & 2 & 5 & -1 \end{bmatrix}$$

(c) Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Find the characteristic polynomial. Hence, or otherwise, Show that $A^2 - 5A - 2I = 0$.

.

[2 + 4 + 4]

8. (a) Show that, "similarity" of matrices is an equivalence relation. In other words, show that, "similarity" is a reflexive, symmetric, and transitive relation.

(b) If $n \times n$ matrices A and B are similar matrices, then prove that they have the same characteristic polynomial and hence the same eigenvalues

(c) Calculate the rank of the following matrix:

$$\begin{bmatrix} 4 & -6 & 0 \\ -6 & 0 & 1 \\ 0 & 9 & -1 \\ 0 & 1 & 4 \end{bmatrix}$$

.

[2 + 4 + 4]