M. Sc. (Ag.) Semester: I Examination, 2023
Course: STAT 502
(Statistical Methods for Applied Sciences)

Time: **Two Hours**   Full marks: **30**

Questions are of equal value or as indicated in the margin

Answer any three questions

1. What do you mean by expectation of a random variable X? Obtain the mean and variance of the random variable X, the number of heads obtained in tossing of an unbiased coin twice.   **(2+8=10)**

2. (a) The p.d.f. of a random variable X is given as: $f(x) = ce^{-\frac{1}{2}\left(\frac{x-100}{5}\right)^2}$, $-\infty < x < \infty$. Write down the mean, mode the variance of the distribution. Also discuss about the skewness and kurtosis status of the distribution.

   (b) Write down four important properties of normal distribution.   **(6+4=10)**

3. (a) If the fitted multiple regression equation is $S = 1.833 + 1.076Y - 0.947N$

   (i) What are the values of b_0, b_1, b_2?

   (ii) Comment on the goodness of fit of the multiple regression equation when multiple correlation value $(R^2) = 0.896$

   Here

   S=Yearly Savings in '000' $Rs.$;

   Y=Monthly Income in '000' $Rs.$;

   N=Number of children

   (b) Draw all possible simple random samples of size two (i.e., $n = 2$), without replacement from a population with $N = 5$ with unit marked as A, B, C, D and E.   **(2+3+5=10)**

4. For given data set of values of y, x_1 and x_2; r_{y1}=0.91, r_{y2}=0.33 and r_{12}=0.81 Examine whether the computations are consistent. If the values are consistent, then calculate $R_{y.12}$.   **(10)**

5. (a) Prove that the maximum value of the variance of a binomial distribution having parameters n and p is $n/4$?

   (b) What do you mean by level of significance and power of a test?   **(5+5=10)**

Questions are of equal value or as indicated in the margin

Answer any five questions

Q.1   Explain correlation coefficient and its measures.   **10**

Q.2   The equations of two regression lines obtained in a correlation analysis study are as follows: **10**

   $8x - 10y + 66 = 0$ & $40x - 18y - 214 = 0$

   (a) Identify the regression lines of x on y and y on x.

   (b) Find the value of correlation coefficient also state clearly the sign of r.

   (c) Find mean value of x and y

   (d) Predicted value of y when x = 32

   (e) Predicted value of x when y = 50

Q.3   Identify the models and write normal equations for these given models   **10**

   (a) $Y = at^b$ (b) $Y = ab^t$

Q.4   **10**

Coefficients

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	-15.46	5.74	-2.69	0.04
X_1	2.23	0.92	2.43	0.05
X_2	0.34	0.10	3.41	0.01
X_3	1.45	0.66	2.20	0.07

R Square =0.98

Answer the questions from the given information

(a) What proportion of the total variation in Y is explained by the regression equation?

(b) Write down the prediction equation.

(c) Does X_1 contribute information in the prediction of Y?

(d) Does X_2 contribute information in the prediction of Y?

(e) Does X_3 contribute information in the prediction of Y?

Q.5   Write the formula with limits for following notation   **10**

   (a) r (b) $R^2_{1.23}$ (c) $r_{12.3}$ (d) partial regression coefficients $\beta_{12.3}$ and $\beta_{13.2}$

Q.6   What do you mean by Test of significance of a population correlation coefficient based on t-test. Suppose   **10**
   a random sample of 11 pairs of observations from a normal population gives a correlation coefficient (r) =
   0.239 and calculated value of t = 0.74. Is it significant at 5% level of significance?
   $t_9(5\%) = 2.262$

Q.7   What do you mean by multicollinearity? What are the different methods of detection of multicollinearity?   **10**
   Write down the remedies for handling multicollinearity.

Q.8   Write short notes on **any two** of the following: **10**

   (a) Regression

   (b) Scatter diagram

   (c) Forward selection of variables