(iii) The correlation coefficient between X and Y.

(iv) The mean values of X and Y.

(v) The standard deviation of Y given standard deviation of X = 3.

(vi) The angle between two lines of regression. (15)

7. (a) For multiple linear regression model, define residual vector e. Show that

(i) $E(e) = 0$

(ii) $Cov(e, \hat{Y}) = 0$

(iii) $e'\hat{Y} = 0$

(iv) $e'X = 0$

(b) Explain the concept of R^2 and Adjusted R^2. Discuss the importance of R^2 for model adequacy checking. (8,7)

8. (a) Suppose the postulated model is $E(Y) = \beta_0 + X$. But the model $E(Y) = \beta_0 + \beta_1 X + \beta_2 X^2$ is actually the true response function unknown to us. If we use observations of Y at X = –1, 0, 1, to estimate β_0 and β_1 from the postulated model, what biases will be induced?

(b) Find an unbiased estimate of error variance σ^2 in multiple linear regression model. (8,7)

(1000)

[This question paper contains 4 printed pages.]

Your Roll No..............

Sr. No. of Question Paper : 1260 **I**

Unique Paper Code : 2373010008

Name of the Paper : DSE: Regression Analysis

Name of the Course : **B.A./B.Sc. (P) NEP-UGCF**

Semester : V

Duration : 3 Hours Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.

2. Attempt any **SIX** Questions.

3. Use of simple calculator is allowed.

1. (a) Define Karl Pearson's coefficient of correlation. If X and Y are random variables and a, b, c, d are any numbers such that $a \neq 0$, $c \neq 0$, then

$$r(aX + b,\ cY + d) = \frac{ac}{|ac|} r(X, Y).$$

(b) Prove that for two independent variables, correlation coefficient r is 0. Is the converse true? Justify your answer with an example. Also find limits of r. (8,7)

2. (a) If X and Y are two variables with variances σ_X^2 and σ_Y^2, respectively and r is the coefficient of correlation between them. If $U = X + kY$ and $V = X + \left(\frac{\sigma_X}{\sigma_Y}\right)Y$. Find the value of k so that U and V are uncorrelated.

(b) The coefficient of rank correlation between the marks obtained by 10 students in Mathematics and Statistics was found to be 0.5. It was later discovered that the difference in ranks in two subjects obtained by one of the students was wrongly taken as 3 instead of 7. Find the correct coefficient of rank correlation. (8,7)

3. Consider the simple linear regression model

$$y = \beta_0 + \beta_1 x + \varepsilon$$

where $E(\varepsilon) = 0$, $V(\varepsilon) = \sigma^2$, ε's are uncorrelated.

(i) Show that $E(MSR) = \sigma^2 + \beta_1^2 S_{XX}$.

(ii) Show that $E(MSE) = \sigma^2$

Hence test for significance of regression, where symbols have their usual meanings. (15)

4. For a given model $Y_{nx1} = X_{nxp}\beta_{px1} + \varepsilon_{nx1}$ with $E(\varepsilon) = 0$, $V(\varepsilon) = \sigma^2 I$, $\rho(X) = p < n$. Find the least squares estimate $\hat{\beta}$ of β. Show that

(i) $\hat{\beta}$ is linear function of Y

(ii) $\hat{\beta}$ is unbiased for β

(iii) $Var(\hat{\beta}) = \sigma^2 \; (X'X^{-1})$ (15)

5. (a) Consider the model

$$E(y_1) = 2\beta_1 + \beta_2$$
$$E(y_2) = \beta_1 - \beta_2$$
$$E(y_3) = \beta_1 - \beta_3$$

Obtain least square estimate of $\beta_1 + 2\beta_2$. Also find its variance.

(b) Find 100 (1-α) % confidence interval for σ^2 in simple linear regression model. (9,6)

6. The equations of two lines of regression are

$$8X - 10Y + 66 = 0 \text{ and } 40X - 18Y = 214.$$

Find the following :

(i) The lines of regression of X on Y and Y on X.

(ii) Regression coefficients of Y on X and X on Y.