4/51 2111N8054A

M.Sc. (Integrated)
Artificial Intelligence and Machine Learning
M.Sc. Sem-V Examination
CC-302
Regression Theory

Time: 2 and 30 hours] **November 2024** **[Max. Marks:70**

Instructions: All questions are compulsory. Use of non-programmable scientific calculator is allowed.

Q.1 **(a)** A study compared the effects of four 1-month point-of-purchase promotions on sales. The unit sales for five stores using all four promotions in different months follow. **(07)**

Free sample	78	87	81	89	85
One-pack gift	94	91	87	90	88
Cents off	73	78	69	83	76
Refund by mail	79	83	78	69	81

Do the promotions produce different effects on sales? Use $\alpha = 0.01$ (Use $F_{tab} = 5.29$)

(b) In the following calculations, use $\alpha = 0.05$. Use Fisher's LSD procedure to determine which means are different. ($\bar{\bar{x}} = 62$) (Use $F_{tab} = 4.26$) **(07)**

Treatment 1	63	47	54	40	$s_j^2 = 96.67$
Treatment 2	82	72	88	66	$s_j^2 = 97.34$
Treatment 3	69	54	61	48	$s_j^2 = 81.99$

OR

(a) The calculations necessary to set up the analysis of variance table. Use $\alpha = 0.05$ to test for any significant differences. where $SST = 354.93$, $SSBL = 312.32$ & $\bar{\bar{x}} = 11.73$. Set up the ANOVA table. (Use $F_{tab} = 4.46$) **(07)**

	A	**B**	**C**
1	10	9	8
2	12	6	5
3	18	15	14
4	20	18	18
5	8	7	8

(b) Develop the analysis of variance computations for the following data. At $\alpha = 0.05$, is there a significant difference between the treatment means? Set up the ANOVA table. **(07)**
(Use $F_{tab} = 3.39$)

A	136	120	113	107	131	114	129	102			$s_j^2 = 146.86$
B	107	114	125	104	107	109	97	114	104	89	$s_j^2 = 96.44$
C	92	82	85	101	89	117	110	120	98	106	$s_j^2 = 173.78$

Q.2 **(a)** Cost accountants often estimate overhead based on the level of production. At the Standard Knitting Co., they have collected information on overhead expenses and units produced at different plants, and want to estimate a regression equation to predict future overhead. **(07)**

Overhead	191	170	272	155	280	173	234	116	153	178
Units	40	42	53	35	56	39	48	30	37	40

a. Develop a scatter diagram for these data and indicate about the relationship between the two variables?
b. Develop the regression equation for the cost accountants.
c. Predict overhead when 50 units are produced.

(b) Use the following data. **(07)**

PTO

N 805 A-2

Use $s = 6.5141$ & $\bar{\bar{x}} = 10$ & $\sum(x_i - \bar{\bar{x}})^2 = 190$ & $\hat{y} = 7.6 - 0.9x$ & $t_{\frac{\alpha}{2}} = 3.182$

x_i	2	6	9	13	20
y_i	7	18	9	26	23

a. Develop a 95% confidence interval for the expected value of y when $x = 12$.
b. Estimate the standard deviation of an individual value of y when $x = 12$.
c. Develop a 95% prediction interval for y when $x = 12$.

OR

(a) The following data are the monthly salaries y and the grade point averages x for students who obtained a bachelor's degree in business administration with a major in information systems. The estimated regression equation for these data is $\hat{y} = 1790.5 + 581.1x$. **(07)**

GPA	2.6	3.4	3.6	3.2	3.5	2.9
Monthly Salary (Rs.)	3300	3600	4000	3500	3900	3600

a. Compute SST, SSR and SSE.
b. Compute the coefficient of determination r^2.
c. What is the value of the sample correlation coefficient?

(b) Given are data for two variables, x and y. The estimated regression equation for these data is $\hat{y} = -7.02 + 1.59x$. **(07)**

x_i	6	11	15	18	20
y_i	6	8	12	20	30

a. Compute the residuals.
b. Develop a plot of the residuals against the independent variable x.

Q.3 (a) Consider the following data for a dependent variable y and two independent variables, x_1 and x_2. **(07)**

x_1	30	47	25	51	40	51	74	36	59	76
x_2	12	10	17	16	5	19	7	12	13	16
Y	94	108	112	178	94	175	170	117	142	211

Develop an estimated regression equation relating y to x_1. Estimate y if $x_1 = 45$.

(b) The following estimated regression equation based on 10 observations was presented. **(07)**

$$\hat{y} = 29.1270 + 0.5906\,x_1 + 0.4980\,x_2$$

Here SST = 6724.125, SSR = 6216.375, $S_{b1} = .0813$, and $S_{b2} = .0567$.

a. Compute MSR and MSE.
b. Compute F and perform the appropriate F test. Use $\alpha = 0.05$. (Use $F_{tab} = 4.74$)
c. Perform a t - test for the significance of β_1. Use $\alpha = 0.05$. (Use $t_{tab} = 1.895$)

OR

(a) **I.** The following estimated regression equation relating sales to inventory investment and advertising expenditures was given. **(07)**

$$\hat{y} = 25 + 10\,x_1 + 8\,x_2$$

The data used to develop the model came from a survey of 10 stores; for those data, SST = 16,000 and SSR = 12,000. Compute and interpret R^2 and R_a^2.

II. The following estimated regression equation based on 30 observations was presented.

$$\hat{y} = 17.6 + 3.8x_1 - 2.3x_2 + 7.6\,x_3 + 2.7\,x_4$$

The values of SST and SSR are 1805 and 1760, respectively. Compute R^2 and R_a^2.

(b) The following estimated regression equation based on 10 observations was presented. **(07)**

$$\hat{y} = -18.37 + 2.01\,x_1 + 4.74\,x_2$$

Here SST = 15182.9, SSR = 14052.2, $S_{b1} = .2471$, and $S_{b2} = .9484$.

a. Compute MSR and MSE.
b. Compute F and perform the appropriate F test. Use $\alpha = 0.05$. (Use $F_{tab} = 4.74$)

c. Perform a t - test for the significance of β_1. Use $\alpha = 0.05$. (Use $t_{tab} = 1.895$)

Q.4 **(a)** Consider the following time series data. **(07)**

Week	1	2	3	4	5	6	7
Value	24	13	20	12	19	23	15

a. Construct a time series plot. What type of pattern exists in the data?
b. Develop the three-week moving average forecasts for this time series. Compute MSE and a forecast for week 8.

(b) The following time series shows the sales of a particular product over the past 12 months. **(07)**

Month	1	2	3	4	5	6	7	8	9	10	11	12
Sales	105	135	120	105	90	120	145	140	100	80	100	110

a. Construct a time series plot. What type of pattern exists in the data?
b. Use $\alpha = 0.3$ to compute the exponential smoothing forecasts for the time series.

OR

(a) Consider the following time series. **(07)**

t	1	2	3	4	5	6	7
Y_t	82	60	44	35	30	29	35

a. Construct a time series plot. What type of pattern exists in the data?
b. Develop the linear trend equation for this time series.
c. What is the forecast for $t = 8$?

(b) The quarterly sales data (number of copies sold) for a college textbook over the past three years follow. **(07)**

Quarter	**1**	**2**	**3**	**4**
Year1	4	2	3	5
Year2	6	3	5	7
Year3	7	6	6	8

a. Construct a time series plot. What type of pattern exists in the data?
b. Show the four-quarter and centred moving average values for this time series.

Q.5 Attempt any **SEVEN** out of **TWELVE:** **(14)**

(1) An experiment has been conducted for four treatments with eight blocks. Complete the following analysis of variance table.

Source of Variation	**Sum of Squares**	**Degrees of Freedom**	**Mean Square**	**F**
Treatments	900			
Blocks	400			
Error				
Total	1800			

(2) Write down the formula of LSD and rejection rule for Fisher's LSD.
(3) The required condition for using an ANOVA procedure on data from several populations is that the _____.
(4) Larger values of r^2 imply that the observations are more closely grouped about the_____.
(5) Write down the formula of Holt's Exponential moving average.
(6) A variable that takes on the values of 0 or 1 and is used to incorporate the effect of qualitative variables in a regression model is called _____.
(7) In a questionnaire, respondents are asked to mark their marital status. Marital status is an example of which type of scale of measurement?
(8) In multiple regression analysis, the correlation among the independent variables is termed_____.

(9) If the estimate of the trend component is 158.2, the estimate of the seasonal component is 94%, the estimate of the cyclical component is 105%, and the estimate of the irregular component is 98%, then find the forecast of production which will be done by multiplicative model.

(10) The following linear trend expression was estimated using a time series with 17 time periods.

$$T_t = 129.2 + 3.8t$$

Find the trend projection for time period 18.

(11) The time series component that reflects variability due to natural disasters is called_____.

(12) If $r = 1$, the angle between two regression lines is _____.

**** ****