COEP Tech. University
Department of Mathematics
(MA- 21001) Probability and Statistics for Engineers
T.Y. B. Tech. Semester V (Autumn)
Academic Year 2023-24

Tutorial 5 on Unit 5 and 6 (Regression and Branch Specific Applications)

1. (a) What is simple linear regression?
   (b) What is multiple linear regression?
   (c) What is method of least squares?
   (d) What is meant by a polynomial regression?
   (e) Derive normal equations for the simple linear regression, multiple linear regression for two independent variables and for parabolic regression.

2. The raw material used in the production of a certain synthetic fiber is stored in a location without a humidity control. Measurements of relative humidity in the storage location and the moisture content of a sample of the raw material were taken over 15 days with the following data (in percentages) resulting.

relative humidity	46	53	29	61	36	39	47	49	52	38	55	32	57	54	44
Moisture content	12	15	7	17	10	11	11	12	14	9	16	8	18	14	12

Estimate regression line for above given data. Also fit a parabolic curve $Y = b_0 + b_1x + b_2x^2$ to above data. ANS: regression line is $Y = -2.51 + 0.32x$

3. The following data relate x, the moisture of a wet mix of a certain product to Y the density of the finished product.

x_i	5	6	7	10	12	15	18	20
Y_i	7.4	9.3	10.6	15.4	18.1	22.2	24.1	24.8

Fit a linear curve to this data. Plot a scatter diagram for this data. Predict density of finished product when the moisture of a wet mix of certain product is 17.

4. A study conducted at VPI and SU to determine if certain static arm-strength measures have an influence on the "dynamic lift" characteristic of an individual. Thirteen individuals were subjected to strength tests and then were asked to perform a weight lifting test in which weight was dynamically lifted overhead. The data are given here:

individual	1	2	3	4	5	6	7	8	9	10	11	12	13
x_i	17.3	19.3	19.5	19.7	22.9	23.1	26.4	26.8	27.6	28.1	28.2	28.7	29.0
y_i	87.3	99.3	79.5	59.7	25.9	55.1	63.4	100.8	78.6	89.1	100.2	56.7	57.0

Estimate α and β for the regression line. Find point estimate of $x = 30$.

5. The grades of a class of 9 students on a midterm report (x) and on the final examination (y) are as follows:

x_i	77	50	71	72	81	94	96	99	67
Y_i	82	66	78	34	47	85	99	99	68

   (a) Estimate the linear regression line.

   (b) Estimate the final grade of a student who received a grade of 85 on the midterm report.

6. A study was made by a retail merchant to determine the relation between weekly advertising expenditures and sales. The following data was recorded.

Advertising Costs (Rs.)	40	20	25	20	30	50	40	20	50	40	25	50
Sales(Rs.)	385	400	395	365	475	440	490	420	560	525	480	510

   (a) Plot a scatter diagram.

   (b) Find equation of regression line to predict weekly sales from advertising expenditures.

   (c) Estimate weekly sales when advertising costs are Rs.$35/-$

   (d) Plot residues versus advertising costs.

7. Student grades and classes missed for 12 students who took Statistics course is recorded as below.

Student	1	2	3	4	5	6	7	8	9	10	11	12
Final exam Grade (y)	85	74	76	90	85	87	94	98	81	91	76	74
Mid-Sem exam grade (x_1)	65	50	55	65	55	70	65	70	55	70	50	55
Classes missed x_2	1	7	5	2	6	3	2	5	4	3	1	4

   (a) Fit the multiple linear regression equation of the form $Y = b_0 + b_1x_1 + b_2x_2$

   (b) Estimate approximate final grade for a student who has mid-sem exam score 60 and missed 4 classes.

8. Fit a least square line to the data given below using x as the independent variable.

.

x	1	3	4	6	8	9	11	14
y	1	2	4	4	5	7	8	9

9. State whether the following statements are True or False and justify your answers.

.

   (a) The estimated simple linear regression equation maximizes the sum of the squared deviations between each value of y and the line.

(b) Slope of the linear function is equal to the change in independent variable divided by the corresponding change in the dependent variable.

(c) Method of least squares is used to fit a least square line to given finite number of x values and corresponding y values. Then $(\bar{x}, \bar{y})$ always lie on the least square line.

(d) Sum of squares of residues $SS_R = \sum_i (Y_i - A - Bx_i)^2$ is a chi-squared distribution with $(n-1)$ degrees of freedom.

10. Consider the following data of x and y values.

x	0	1	2	3	4	5	6
y	1	4	5	3	2	3	4

(a) Fit a polynomial to the above data $\mu_{Y|x} = \beta_0 + \beta_1 x + \beta_2 x^2$. [CO5, 3 Marks]

(b) Predict Y when $x = 2$.

11. In quality control concerned with any manufacturing process, What is meant by assignable cause and what is chance variation? What is control chart and upper control limit (UCL) and lower control limit (LCL)?

12. Assume that items produced are supposed to be normally distributed with mean 35 and standard deviation 3. To monitor this process, subgroups of size 5 are sampled. If the following represents the averages of the first 20 subgroups, does it appear that the process was in control?

Sample	1	2	3	4	5	6	7	8	9	10
$\bar{X}$	34	31.6	30.8	33.0	35	32.2	33	32.6	33.8	35.8
Sample	11	12	13	14	15	16	17	18	19	20
$\bar{X}$	35.8	35.8	34	35.0	33.8	31.6	33	33.2	31.8	33.6

13. Consider Markov chain with three states, $S = \{1, 2, 3\}$ that has the transition matrix

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}.$$

(a) Draw the state transition diagram for this chain.

(b) If we know $P(X_1 = 1) = P(X_1 = 2) = \frac{1}{4}$, find $P(X_1 = 3, X_2 = 2, X_3 = 1)$.

14. Consider the stochastic process so called 'random telegraph process'. This is a discrete-state continuous-time process $\{X(t)|; -\infty < t < \infty\}$ with the state space $\{-1, 1\}$. Assume that these two values are equally likely: $P[X(t) = -1] = \frac{1}{2} = P[X(t) = 1]$, $-\infty < t < \infty$.

(a) Find $\mu(t) = E(X(t))$. Is it independent of t?

(b) Find autocorrelation function $R(0) = E(X^2(t))$.

(c) Is this random telegraph process stationary in wide-sense? Justify.

Consider the given system which represents two-state Marcov Chain: Suppose that whether it rains tomorrow depends on previous weather conditions only through whether it is raining today. Suppose further that if it is raining today, then it will rain tomorrow with probability α and if it is not raining today, then it will rain tomorrow with probability β. If we say that system is in state 0 when it rains and state 1 when it does not, find transition probabilities of this Marcov chain and hence write its transition probability matrix. Draw the state transition diagram for this chain.

———THE END———-