Assignement-I

Q-1 Create the relationship model for the given data as shown below and find the relationship b/w height & weight of TIET students. Also, unknown entry of CSED used having height 181cm, then Determine its weight

Dataset of CSED-TIET students

Entries	Height (X) cm	weight (Y) kg
1	151	63
2	174	81
3	138	56
4	186	91
5	128	47
6	136	57
7	179	76
8	163	72
9	152	62
10	131	48

Q-2 The rent of a property in a particular area is provided to you. Find the relationship between area & rent using the concept of LR. Also, predict when area is 790 ft^2 then what is estimated rent?

Entries	Area (ft^2)	Rent (₹)
1	340	500
2	1080	1700
3	640	1100
4	880	800
5	990	1400
6	510	500

Q-3 The marks obtained by a student in an examination according to spending study time in minutes. Total Marks or marks out of 2000 is given. Find the relationship b/w study time & marks by using the concept of linear regression. Also predict the marks for a student if he/she study for 790 minutes.

Entries	Study time (Minutes)	Marks obtained
1	350	520
2	1070	1600
3	630	1000
4	890	850
5	940	1350
6	500	490.

Q-4 Sales of a company in dollars for each years are shown below

x (Year) :	2005	2006	2007	2008	2009
y (sales) :	12	19	29	37	45

(a) Find the least sq. regression line, $y = ax + b$.

(b) Use the least sq. regression line as a model to estimate the sales of the company in 2012.

Assignement-II

By Dr. Karamjeet Singh.

Q-1 Given the probability distribution function of binomial distribution is

$$f(x, x_i) = \begin{bmatrix} n \\ x_i \end{bmatrix} p^{x_i} (1-p)^{n-x_i} \; \forall \; i=1,2,\ldots,n.$$

Find its maximum likelihood estimate for 'p'.

Q-2. For the Gamma distribution $G(x;\alpha,\lambda)$. Find maximum likelihood for 'α' when 'λ' is treated as constant.

Q-3 Obtain the ML estimator of α & β for the Uniform distribution having the following p.d.f

$$f(x;\alpha,\beta) = \begin{cases} \dfrac{1}{\beta-\alpha} & \text{for } \alpha \leq x \leq \beta \\ 0 \;; & \text{otherwise} \end{cases}$$

Q-4 Let $x_1, x_2, \ldots, x_n$ represents a random sample from each of the distribution having the following pdf.

(a) $f(x;\theta) = \theta x^{\theta-1}$, $0 < x < 1$, $0 < \theta < \infty$, 0, else where

(b) $f(x;\theta) = \frac{1}{2} e^{-(x-\theta)}$, $-\infty < x < \infty$, $-\infty < \theta < \infty$

In each case find MLE $\hat{\theta}$ for θ.

Numerical on Multiple Linear Regression (MLR)

Data set of following type

$\xrightarrow{}$ I.Vs

Y (D.V)	X_1	X_2
-2.7	4	7
4.5	5	6
2.5	5	7
10.5	5	2
6.7	3	2

A.T. Eqn. of MLR

$$Y = b_0 + b_1 X_1 + b_2 X_2 \quad \text{—} \quad (1)$$

Our objective is to estimate the approximate value of b_0, b_1 & b_2

$$\therefore \quad (A) \text{ — } b_0 = \bar{Y} - b_1 \bar{X}_1 - b_2 \bar{X}_2$$

$$(B) \text{ — } b_1 = \frac{(\sum x_2^2)\cdot(\sum x_1 y) - (\sum x_1 \cdot x_2)\cdot(\sum x_2 y)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 \cdot x_2)^2}$$

$$(C) \text{ — } b_2 = \frac{(\sum x_1^2)\cdot(\sum x_2 y) - (\sum x_1 \cdot x_2)(\sum x_1 y)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 \cdot x_2)^2}$$

Use this in (A) (B) & (C)

$$\text{Now, } \sum x_i^2 = \sum X_i^2 - \frac{(\sum X_i)^2}{N}$$

$$\text{Let } i = 1, \text{ then } \sum x_1^2 = \sum X_1^2 - \frac{(\sum X_1)^2}{N}$$

$$i = 2, \text{ then } \sum x_2^2 = \sum X_2^2 - (\sum X_2)^2 / N.$$

* Also, $\sum x_i y = \sum X_i Y - \frac{(\sum X_i)(\sum Y)}{N}$

Let, $i=1$, then $\sum x_1 y = \sum X_1 Y - \frac{(\sum X_1)(\sum Y)}{N}$

$i=2$, then $\sum x_2 y = \sum X_2 Y - \frac{(\sum X_2)(\sum Y)}{N}$

* Also, $\sum x_1 \cdot x_2 = \sum X_1 X_2 - (\sum X_1)\cdot(\sum X_2)/N$

↑ Use above in Eq^n. (A), (B) & (C).

Y	X_1	X_2	X_1^2	X_2^2	$X_1 Y$	$X_2 Y$	$X_1 X_2$
-2.7	4	7	16	49	-10.8	-18.9	28
4.5	5	6	25	36	22.5	27	30
2.5	5	7	25	49	12.5	17.5	35
10.5	5	2	25	4	52.5	21	10
6.7	3	2	9	4	20.1	20.1	6

$\sum Y = 21.5$ ↓ $\quad \sum X_2 = 24$ ↓ $\quad \sum X_2^2 = 142$ ↓ $\quad \sum X_2 Y = 60$ ↓

$\sum X_1 = 22$ $\quad \sum X_1^2 = 100$ $\quad \sum X_1 Y = 96.8$ $\quad \sum X_1 X_2 = 109$

Now, let's get started by putting the calculated values in formula's.

Consider

$$\sum x_1^2 = \sum X_1^2 - \frac{(\sum X_1)^2}{N}$$

$$= 100 - \frac{(22)^2}{5}$$

$$= \frac{500 - 484}{5} = \frac{16}{5} = 3.2$$

Consider

$$\sum x_2^2 = \sum X_2^2 - \frac{(\sum X_2)^2}{N}$$

$$= 142 - \frac{(24)^2}{5}$$

$$= \frac{710 - 576}{5} = \frac{134}{5}$$

$$= 26.8$$

consider

$$\sum x_1 y = \sum X_1 Y - \frac{(\sum X_1)\cdot(\sum Y)}{N}$$

$$= 96.8 - \frac{22 \times 21.5}{5}$$

$$= 96.8 - 94.6$$

$$= 2.2$$

consider

$$\sum x_2 y = \sum X_2 Y - \frac{\sum X_2 \cdot \sum Y}{N}$$

$$= 60 - \frac{24 \times 21.5}{5}$$

$$= -43.2$$

Now,

$$\sum x_1 \cdot x_2 = \sum X_1 \cdot X_2 - \frac{\sum X_1 \cdot \sum X_2}{N}$$

$$= 109 - \frac{22 \cdot 24}{5}$$

$$= 109 - 105.6$$

$$= 3.4$$

Now consider the Eq^n. B

$$b_1 = \frac{(\sum x_2)^2 \cdot (\sum x_1 y) - (\sum x_1 \cdot x_2) \cdot (\sum x_2 \cdot y)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 \cdot x_2)^2}$$

$$= \frac{26.8 \times (2.2) - (3.4) \times (-43.2)}{(3.2) \cdot (26.8) - (3.4)^2}$$

$$= \frac{58.96 + 146.88}{85.76 - 11.56} = \frac{205.84}{74.2} = 2.774$$

$$\therefore \boxed{b_1 = 2.774}$$

Similarly $b_2 = \dfrac{(\sum x_1^2)\cdot(\sum x_2 y) - (\sum x_1 \cdot x_2) * (\sum x_1 y)}{(\sum x_1^2)\cdot(\sum x_2^2) - (\sum x_1 \cdot x_2)^2}$

$$\therefore b_2 = \frac{(3.2)\cdot(-43.2) - (3.4)(2.2)}{(3.2)\cdot(26.8) - (3.4)^2}$$

$$= \frac{-138.24 - 7.48}{85.76 - 11.56} = \frac{-145.72}{74.2}$$

$$\boxed{b_2 = -1.964}$$

$\therefore$ Eq^n (A) becomes

$$\cancel{Y_2} \; b_0 = \overline{Y} - b_1\overline{X}_1 - b_2\overrightarrow{X_2}$$

$$b_0 = \frac{1}{5}\left[21.5 - (2.774)\cdot(22) + (1.964)\cdot 24\right]$$

$$b_0 = \frac{1}{5}\left[21.5 - 61.028 + 47.136\right]$$

$$b_0 = \frac{7.608}{5} = 1.522$$

Eq^n. (1) become

$$Y = b_0 + b_1 X_1 + b_2 X_2$$

$$Y = 1.522 + 2.744 X_1 + (-1.964) \cdot X_2$$

$$\Rightarrow Y = 1.522 + 2.744 X_1 - 1.964 X_2$$

is the required multiple Linear regression Eq^n.

Factor Analysis Model
Extracting Commonfactor
Determining number of factors
Transformation of factor analysis
factor Scores.