Name: Anik Kumar Majumdar

Linear Regression Assignment

Weight (x)	Price (y)
2	35
4	60
5	20
3	50
6	50
5	55
7	60

Task-01:

Your objective is to manually compute the slope (m) & y-intercept (c) using Ordinary Least Squares Linear Regression. Once determined, apply these values to predict the price when the vegetable weight is 6.

Soln:

$$\text{Slope, } m = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sum (x-\bar{x})^2}$$

(i) $\bar{x} = \frac{\sum x}{N} = \frac{2+4+5+3+6+5+7}{7} = 4.6$

(ii) $\bar{y} = \frac{\sum y}{N} = \frac{35+60+20+50+50+55+60}{7} = 47.14$

(iii) Calculate $\sum (x-\bar{x})(y-\bar{y})$:

$(2-4.6)(35-47.14) = 31.56$

$(4-4.6)(60-47.14) = -7.72$

$(5-4.6)(20-47.14) = -10.86$

$(3-4.6)(50-47.14) = -4.58$

$(6-4.6)(50-47.14) = 4.00$

$(5-4.6)(50-47.14) = 1.14$

$(7-4.6)(60-47.14) = 30.86$

$\therefore \sum (x-\bar{x})(y-\bar{y}) = 31.56 - 7.72 - 10.86 - 4.58 + 4.00 + 1.14 + 30.86$

$= 44.4$

(iv) Calculate $\sum (x-\bar{x})^2$:

$(2-4.6)^2 = 6.76$

$(4-4.6)^2 = 0.36$

$(5-4.6)^2 = 0.16$

$(3-4.6)^2 = 2.56$

$(6-4.6)^2 = 1.96$

$(5-4.6)^2 = 0.16$

$(7-4.6)^2 = 5.76$

$\sum (x-\bar{x})^2 = 6.76+0.36+0.16+2.56+1.96+0.16+5.76$

$\therefore \sum (x-\bar{x})^2 = 17.72$

$$\text{Slope, } m = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sum (x-\bar{x})^2}$$

$$= \frac{44.4}{17.72}$$

$\therefore m = 2.51$

We get that,

$\bar{x} = 4.6$

$\bar{y} = 47.14$

$m = 2.51$

Formula fore intercept C:

If $\bar{y} = m\bar{x} + c$

then, $c = \bar{y} - m\bar{x}$

$\Rightarrow c = 47.14 - (2.51 * 4.6)$

$\therefore c = 35.594$

∴ Now, predict price when the vegetable weight is 6.

$y = mx + c$ | Here, $x = 6$

$\Rightarrow y = 2.51 * 6 + 35.594$

$\therefore y = 50.654$

Ans; The slope, $m = 2.51$

y-intercept, $c = 35.594$

& Apply these values to predict the price when the vegetable weight is 6,

$y = 50.654$

Weight	Price
2	35
4	60
5	20
3	50
6	50
5	55
7	60

Task-02:
Compute the residuals for each data point.

Soln: We get from Task-01:

$m = 2.51$

$C = 35.594$

Predicted Result:

If $x = 2$, then $\hat{y} = mx + C = 2.51 * 2 + 35.594 = 40.61$

If $x = 4$, then $\hat{y} = 2.51 * 4 + 35.594 = 45.63$

If $x = 5$, then $\hat{y} = 2.51 * 5 + 35.594 = 48.14$

If $x = 3$, then $\hat{y} = 2.51 * 3 + 35.594 = 43.12$

If $x = 6$, then $\hat{y} = 2.51 * 6 + 35.594 = 50.654$

If $x = 5$, then $\hat{y} = 2.51 * 5 + 35.594 = 48.14$

If $x = 7$, then $\hat{y} = 2.51 * 7 + 35.594 = 53.164$

We know that,

Residual = Observed value − Predicted Value

Residual Result:

Residual for $x=2$; $35 - 40.61 = -5.61$

Residual for $x=4$; $60 - 45.63 = 14.37$

Residual for $x=5$; $20 - 48.14 = -28.14$

Residual for $x=3$; $50 - 43.12 = 6.88$

Residual for $x=6$; $50 - 50.654 = -0.654$

Residual for $x=5$; $55 - 48.14 = 6.86$

Residual for $x=7$; $60 - 53.164 = 6.836$

Weight(x)	Price (y)	$\hat{y}$ (Predicted)	Residuals $(y-\hat{y})$
2	35	40.61	−5.61
4	60	45.63	14.37
5	20	48.14	−28.14
3	50	43.12	6.88
6	50	50.654	−0.654
5	55	48.14	6.86
7	60	53.164	6.836

Ans:

Weight(x)	Price(Y)
2	35
4	60
5	20
3	50
6	50
5	55
7	60

Task-03:

Calculate both the Mean Squared Error (MSE) & Mean Absolute Error (MAE).

Soln:

For,

$x = 2$; $Y = 35$ & $\hat{Y} = 40.61$

$x = 4$; $Y = 60$ & $\hat{Y} = 45.63$

$x = 5$; $Y = 20$ & $\hat{Y} = 48.14$

$x = 3$; $Y = 50$ & $\hat{Y} = 43.12$

$x = 6$; $Y = 50$ & $\hat{Y} = 50.654$

$x = 5$; $Y = 55$ & $\hat{Y} = 48.14$

$x = 7$; $Y = 60$ & $\hat{Y} = 53.164$

We know that,

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$

MSE calculation:

$$MSE = \frac{1}{7} ((35-40.61)^2 + (60-45.63)^2 + (20-48.14)^2 + (50-43.12)^2 + (50-50.654)^2 + (55-48.14)^2 + (60-53.164)^2)$$

$\therefore$ MSE = 167.34

Again,

we know that,

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |Y_i - \hat{Y}_i|$$

⊞ MAE Calculation:

$$MAE = \frac{1}{7}\Big(|35 - 40.61| + |60 - 45.63| + |20 - 48.14| + |50 - 43.12| + |50 - 50.654| + |55 - 48.14| + |60 - 53.164|\Big)$$

$\therefore$ MAE = 9.91

Ans: MSE = 167.34 & MAE = 9.91