Worksheet 4

Fred Azizi

2023-10-03

$$S_y = \sqrt{\frac{1}{n-1}\left(\sum (y_i-\bar{y})^2\right)}$$

1. The following data has mean income and housing for 10 cities in Florida. Values are in dollars ($) and rounded to the nearest thousand.

$$r = \frac{S_{xy}}{S_x \; S_y}$$

$$S_{xy} = \frac{\sum (x_i-\bar{x})(y_i-\bar{y})}{n-1}$$

$$S_x = \sqrt{\frac{1}{n-1}\left(\sum (x_i-\bar{x})^2\right)}$$

City	Income ($\boldsymbol{x}$)	Housing (y)	$x_i-\bar{x}$	$(x_i-\bar{x})^2$	$y_i-\bar{y}$	$(y_i-\bar{y})^2$	$(x_i-\bar{x})(y_i-\bar{y})$
A	26	109	-3.6	12.96	-3	9	10.8
B	29	97	-0.6	0.36	-15	225	9
C	25	115	-4.6	21.16	3	9	-13.8
D	28	99	-1.6	2.56	-13	169	20.8
E	38	122	8.4	70.56	10	100	84
F	32	145	2.4	5.76	33	1089	79.4
G	25	100	-4.6	21.16	-12	144	55.2
H	22	76	-7.6	57.76	-36	1296	772.2
I	29	113	-0.6	0.36	1	1	-0.6
J	42	144	12.4	153.76	32	1024	396.8

S_x S_y

$\sum x_i = 296$ $\sum y_i = 1120$ 396.4

$\bar{x} = 29.6$ $\boxed{\bar{y} = 112}$ $\sum (x_i-\bar{x})^2$

$\sum (y_i-\bar{y})^2 = 4066$ 915

a. Calculate the correlation coefficient between x and y. What can you conclude about the relationship between the 2 variables?

b. Calculate the least square line.

c. Calculate the coefficient of variation

(d) Cal the ... determination

$$S_{xy} = \frac{915}{10-1} = 101.66$$

$$S_x = \sqrt{\frac{1}{9}\,396.4} \approx 6.20$$

$$S_y = \sqrt{\frac{1}{9}\cdot 4066} \approx 21.26$$

$$r = \frac{S_{xy}}{S_x \, S_y} = \frac{101.66}{6.2 \cdot 21.26} \approx 0.77$$

(Extra space for question 1)

linear

strong Positive relationship between x, Y.

(b) LS Line ⟶

$$\begin{cases} b_1 = \dfrac{S_{xy}}{S_x^2} = \dfrac{101.66}{(6.2)^2} \approx 2.64 \\ b_0 = \bar{Y} - b_1 \bar{X} = 112 - 2.64 \cdot 29.6 = 33.86 \end{cases}$$

(c) $$CV_x = \frac{S_x}{\bar{X}} = \frac{6.2}{29.6} \approx 0.21$$

$$CV_y = \frac{S_y}{\bar{y}} = \frac{21.26}{112} \approx 0.19$$

(d) Coef of determination

$$= r^2 = (0.77)^2 = 0.5929$$

2. A sample of 30 observations has a standard deviation of 4. Find the sum of squared deviations from the sample mean.

$$n = 30$$

$$\boxed{S_x} = 4 \quad \rightarrow \quad \overset{S_x}{\overset{\shortparallel}{\sqrt{\frac{1}{\underset{30}{n}-1} \cdot \sum_{i=1}^{30} (x_i - \bar{x})^2}}} = 4$$

$$\sum_{i=1}^{30} (x_i - \bar{x})^2 = ?$$

$$\left(\sqrt{\frac{1}{29} * \sum_{i=1}^{30} (x_i - \bar{x})^2} \right)^2 = (4)^2$$

$$\cancel{29} * \quad \rightarrow \quad \frac{1}{\cancel{29}} * \sum_{i=1}^{30} (x_i - \bar{x})^2 = 16 * 29$$

$$\rightarrow \quad \sum_{i=1}^{30} (x_i - \bar{x})^2 = 16 * 29 = 464$$

3. Following observations are given for two variables.

y	x
5	2
8	12
18	3
20	6
22	11
30	19
10	18
7	9

y and x

a. Compute and interpret P_{86}.

b. Compute and interpret the correlation coefficient.

c. Draw the relevant diagram for the data above.

$L_1 \quad L_2 \quad L_3 \quad L_4 \quad L_5 \quad L_6 \quad L_7 \quad L_8$

(a) $x: \quad 2 \quad 3 \quad 6 \quad 9 \quad 11 \quad 12 \quad 18 \quad 19$

$$L_p = (\overset{8}{n}+1)\cdot\frac{P}{100} = (9)\cdot\frac{86}{100} = \boxed{7}.74$$

obs obs obs

$$L_7 + (L_8 - L_7)\cdot 0.74$$

$$18 + (\underbrace{19-18}_{1}) \times 0.74 = 18.74 = P_{86}^{x}$$

$\rightarrow Y$

4. 4 candidates are running for mayor; Adams, Brown, Collins and Dalton (We assume one of the candidates is going to win, there is no run off). The following probabilities are assigned:

$$P[\text{Adams wins}] = 0.42 \quad P[\text{Brown wins}] = 0.09$$

$$P[\text{Collins wins}] = 0.27 \quad P[\text{Dalton wins}] = 0.22$$

Determine the probabilities for the following events (use 2 decimal places):

a. Adams loses.

b. Either Brown or Dalton wins.

c. Adams, Brown, or Collins wins.