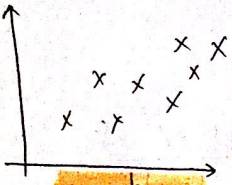
* Sec 12.2 : Least Squares Method.

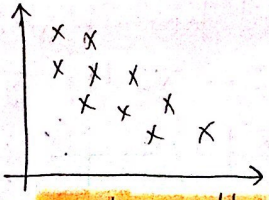
- Recall: Give 2 variables x and y :-

(x_1, y_1) (x_2, y_2) ---- (x_n, y_n).

if we plot a scatter diagram:-



+ve linear relationship



-ve linear relation

No relation.

→ To find the equation of the best trendline, we can use a method called the least square method.

→ The least squares method is a procedure for using sample data to find the estimated regression equation.

$\hat{y} = b_0 + b_1 x$ or $\hat{y} = A + Bx$.

; $\hat{y}$: the estimated value of y.

b_0 : y- intercept for the estimated regression line.

b_1 : the slope of the estimated regression line.

x_i : the independent variable ith value.

$$\rightarrow b_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \quad ; \quad b_0 = \bar{y} - b_1 \bar{x}$$

Ex: Consider the following data:-

(3, 50) (1, 20) (4, 30) (5, 60) (7, 50)

a) Find the estimated linear equation.

$\rightarrow$ To find b_1 :- $\bar{x} = 4$, $\bar{y} = 42$

x_i	y_i	$x - \bar{x}$	$y - \bar{y}$	$(x - \bar{x})(y - \bar{y})$	$(x - \bar{x})^2$
3	50	-1	8	-8	1
1	20	-3	-22	66	9
4	30	0	-12	0	0
5	60	1	18	18	1
7	50	3	8	24	9

$$\therefore b_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{-8 + 66 + 0 + 18 + 24}{1 + 9 + 0 + 1 + 9}$$

$$\therefore \boxed{b_1 = 5}$$

$\rightarrow$ To find b_0:

$$b_0 = \bar{y} - b_1 \bar{x}$$
$$= 42 - (5)(4)$$
$$\therefore \boxed{b_0 = 22}$$

So the estimated equation:

$$\hat{y}_i = 22 + 5x$$

b) Estimate y when $x = 10$.

$\longrightarrow \hat{y} = 22 + 5(10)$

$\therefore \boxed{\hat{y} = 72}$

* Using calculator:-

① mode → 3 → 1

② x_1 → , → y_1 → M+

x_2 → , → y_2 → M+

⋮

x_n → , → y_n → M+

بعد تخزين القيم نضغط ON

③ To find b_0 «A»

shift → 2 → ▷ → ▷ → 1 → =

④ To find b_1 : «B»

shift → 2 → ▷ → ▷ → 2 → =

نستطيع إيجاد mean و standard deviations و r_{xy} و S_{xy} بالخطوات التي تم شرحها في Sec 3.5 .

-Ex: Given the following sample of x and y:-

x	10	18	15	20	4
y	171	200	180	240	80

1) Find the estimated linear regression

Mode 3 1

10 , 171 M+

18 , 200 M+

⋮

4 , 80 M+

shift 2 ▷ ▷ 1 =

$\therefore b_0 = 57.3$

shift 2 ▷ ▷ 2 =

$\therefore b_1 = 8.72$

→ So the estimated equation is:-

$\hat{y} = 57.3 + 8.72x$

2) Estimated y when $x = 19$.

$$y(19) = 57.3 + 8.72(19) = 222.98$$

3) Find the correlation coefficient and interpret your answer.

shift 2 ▷ ▷ 3 =

$\therefore r_{xy} = 0.96$ strong +ve linear relationship

4) Find the covariance.

$$S_{xy} = r_{xy} S_x S_y.$$

To find S_x.

shift 2 3 =

$\therefore S_x = 6.47$

To find S_y:-

shift 2 ▷ 3 =

$\therefore S_y = 58.98.$

$\therefore S_{xy} = (0.96)(6.47)(58.98)$

$\boxed{S_{xy} = 366.34}$