Hoofdstuk 4

Bijlage 4.1:

Mean

Population mean: $$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$$

Sample mean: $$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Bijlage 4.2:

Variance

Population variance: $$\sigma^2 = \frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N}$$

Sample variance:* $$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n - 1}$$

The population variance is represented by σ^2(Greek letter *sigma* squared).

Bijlage 4.3:

x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
8	$(8 - 7) = 1$	$(1)^2 = 1$
4	$(4 - 7) = -3$	$(-3)^2 = 9$
9	$(9 - 7) = 2$	$(2)^2 = 4$
11	$(11 - 7) = 4$	$(4)^2 = 16$
3	$(3 - 7) = -4$	$(-4)^2 = 16$
	$\sum_{i=1}^{5}(x_i - \bar{x}) = 0$	$\sum_{i=1}^{5}(x_i - \bar{x})^2 = 46$

Bijlage 4.4:

Shortcut for Sample Variance

$$s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}\right]$$

Bijlage 4.5:

The following are the number of summer jobs a sample of six students applied for. Find the mean and variance of these data.

17 15 23 7 9 13

SOLUTION

The mean of the six observations is

$$\bar{x} = \frac{17 + 15 + 23 + 7 + 9 + 13}{6} = \frac{84}{6} = 14 \text{ jobs}$$

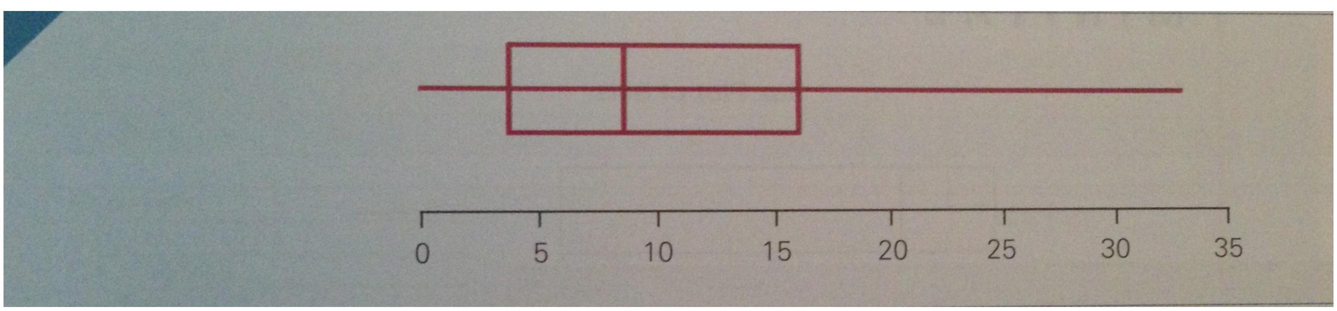
The sample variance is

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n - 1}$$

$$= \frac{(17 - 14)^2 + (15 - 14)^2 + (23 - 14)^2 + (7 - 14)^2 + (9 - 14)^2 + (13 - 14)^2}{6 - 1}$$

$$= \frac{9 + 1 + 81 + 49 + 25 + 1}{5} = \frac{166}{5} = 33.2 \text{ jobs}^2$$

Bijlage 4.6:



Bijlage 4.7:

Coefficient of Correlation

Population coefficient of correlation: $\rho = \dfrac{\sigma_{xy}}{\sigma_x \sigma_y}$

Sample coefficient of correlation: $r = \dfrac{s_{xy}}{s_x s_y}$

Bijlage 4.8:

SOLUTION

Because we've already calculated the covariances we need to compute only the standard deviations of X and Y.

$$\bar{x} = \frac{2 + 6 + 7}{3} = 5.0$$

$$\bar{y} = \frac{13 + 20 + 27}{3} = 20.0$$

$$s_x^2 = \frac{(2 - 5)^2 + (6 - 5)^2 + (7 - 5)^2}{3 - 1} = \frac{9 + 1 + 4}{2} = 7.0$$

$$s_y^2 = \frac{(13 - 20)^2 + (20 - 20)^2 + (27 - 20)^2}{3 - 1} = \frac{49 + 0 + 49}{2} = 49.0$$

The standard deviations are

$$s_x = \sqrt{7.0} = 2.65$$

$$s_y = \sqrt{49.0} = 7.00$$

The coefficients of correlation are:

Set 1: $r = \dfrac{s_{xy}}{s_x s_y} = \dfrac{17.5}{(2.65)(7.0)} = .943$

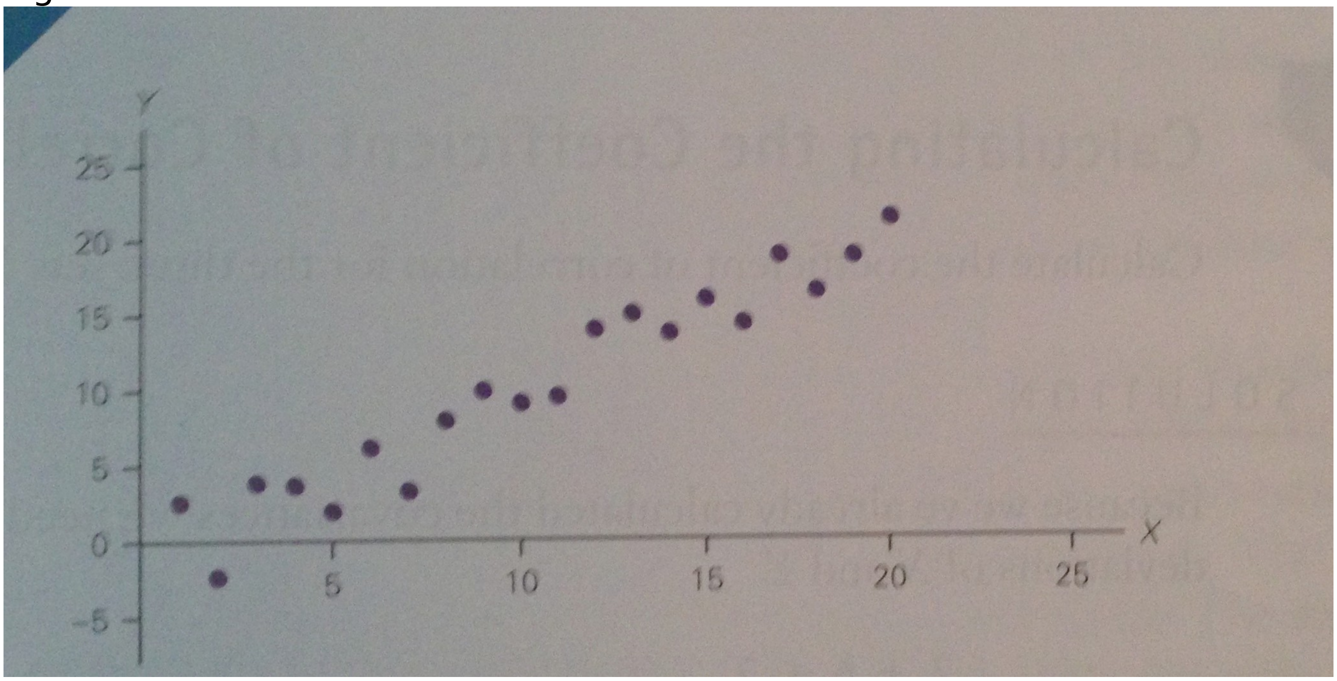
Set 2: $r = \dfrac{s_{xy}}{s_x s_y} = \dfrac{-17.5}{(2.65)(7.0)} = -.943$

Set 3: $r = \dfrac{s_{xy}}{s_x s_y} = \dfrac{-3.5}{(2.65)(7.0)} = -.189$

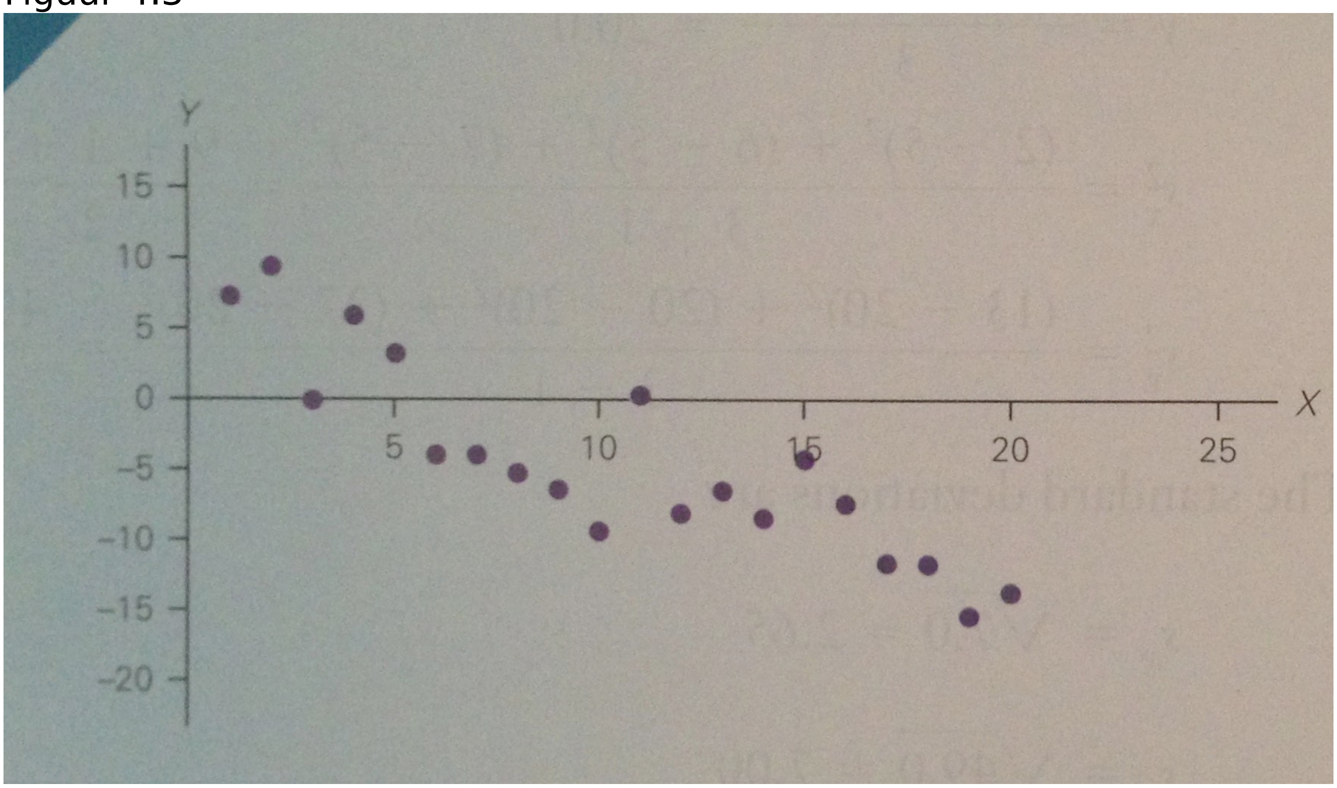
It is now easier to see the strength of the linear relationship between X and Y.

Bijlage 4.9:

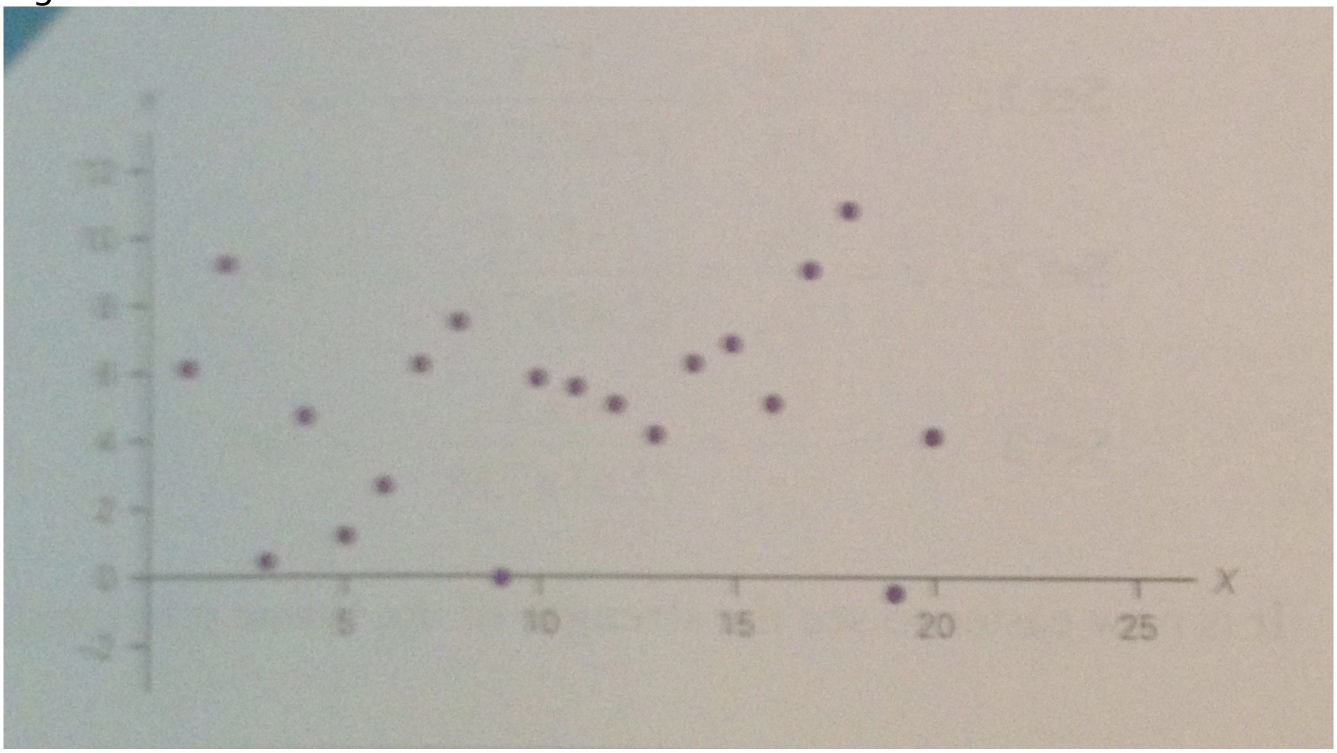
Figuur 4.2



Figuur 4.3



Figuur 4.4



Bijlage 4.10

$$\hat{y} = b_0 + b_1 x$$

Bijlage 4.11

Least Squares Line Coefficients

$$b_1 = \frac{s_{xy}}{s_x^2}$$

$$b_0 = \bar{y} - b_1 \bar{x}$$