Einfache lineare Regression **Lösungen+** **Übungen**

Aufgabe 1

x_i	$x_i - \overline{x}$	$(x_i - \overline{x})^2$	y_i	$y_i - \overline{y}$	$(y_i - \overline{y})^2$	$(x_i - \overline{x})(y_i - \overline{y})$
3	−7	49	12	5	25	−35
9	−1	1	5	−2	4	2
18	8	64	4	−3	9	−24
30	0	114	21	0	38	−57

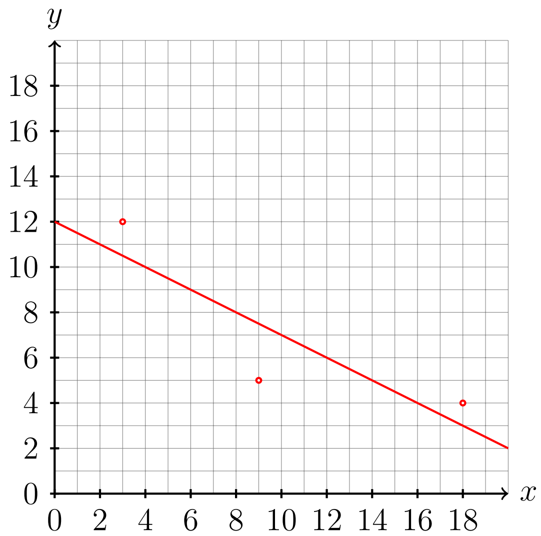
$\overline{x} = 10$, $\overline{y} = 7$

$$a = \frac{\sum(x_i - \overline{x})(y_i - \overline{y})}{\sum(x_i - \overline{x})^2} = \frac{-57}{114} = -\frac{1}{2}$$

$b = \overline{y} - a \cdot \overline{x} = 12$

Regressionsgerade: $y = -\frac{1}{2} \cdot x + 12$

Korrelation: $r_{xy} = \dfrac{\sum(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum(x_i - \overline{x})^2 \cdot \sum(y_i - \overline{y})^2}} = \dfrac{-57}{\sqrt{4332}} = -0.865$



Aufgabe 2

x_i	$x_i - \overline{x}$	$(x_i - \overline{x})^2$	y_i	$y_i - \overline{y}$	$(y_i - \overline{y})^2$	$(x_i - \overline{x})(y_i - \overline{y})$
9	2	4	14	4	16	8
13	6	36	15	5	25	30
1	-6	36	4	-6	36	36
5	-2	4	7	-3	9	6
28	0	80	40	0	86	80

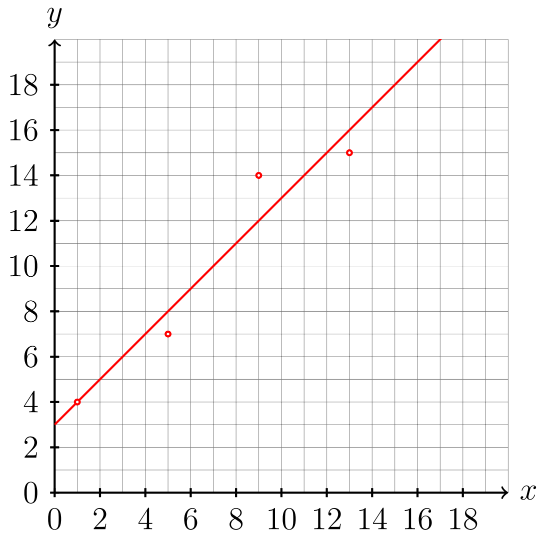
$\overline{x} = 7$, $\overline{y} = 10$

$$a = \frac{\sum(x_i - \overline{x})(y_i - \overline{y})}{\sum(x_i - \overline{x})^2} = \frac{80}{80} = 1$$

$b = \overline{y} - a \cdot \overline{x} = 3$

Regressionsgerade: $y = 1 \cdot x + 3$

Korrelation: $r_{xy} = \frac{\sum(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum(x_i - \overline{x})^2 \cdot \sum(y_i - \overline{y})^2}} = \frac{80}{\sqrt{6880}} = 0.964$



Aufgabe 3

x_i	$x_i - \overline{x}$	$(x_i - \overline{x})^2$	y_i	$y_i - \overline{y}$	$(y_i - \overline{y})^2$	$(x_i - \overline{x})(y_i - \overline{y})$
16	4	16	9	-3	9	-12
13	1	1	12	0	0	0
9	-3	9	14	2	4	-6
10	-2	4	13	1	1	-2
48	0	30	48	0	14	-20

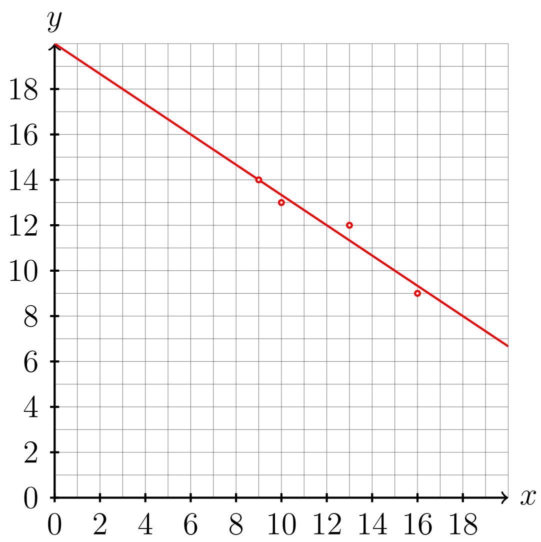
$\overline{x} = 12$, $\overline{y} = 12$

$$a = \frac{\sum(x_i - \overline{x})(y_i - \overline{y})}{\sum(x_i - \overline{x})^2} = \frac{-20}{30} = -\frac{2}{3}$$

$b = \overline{y} - a \cdot \overline{x} = 20$

Regressionsgerade: $y = -\frac{2}{3} \cdot x + 20$

Korrelation: $r_{xy} = \frac{\sum(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum(x_i - \overline{x})^2 \cdot \sum(y_i - \overline{y})^2}} = \frac{-20}{\sqrt{420}} = -0.975$



Aufgabe 4

x_i	$x_i - \overline{x}$	$(x_i - \overline{x})^2$	y_i	$y_i - \overline{y}$	$(y_i - \overline{y})^2$	$(x_i - \overline{x})(y_i - \overline{y})$
7	2	4	17	3	9	6
3	-2	4	13	-1	1	2
6	1	1	15	1	1	1
4	-1	1	11	-3	9	3
20	0	10	56	0	20	12

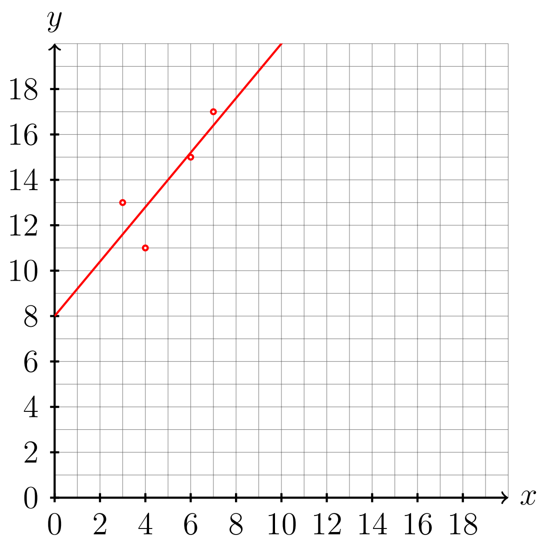
$\overline{x} = 5$, $\overline{y} = 14$

$$a = \frac{\sum(x_i - \overline{x})(y_i - \overline{y})}{\sum(x_i - \overline{x})^2} = \frac{12}{10} = \frac{6}{5}$$

$b = \overline{y} - a \cdot \overline{x} = 8$

Regressionsgerade: $y = \frac{6}{5} \cdot x + 8$

Korrelation: $r_{xy} = \frac{\sum(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum(x_i - \overline{x})^2 \cdot \sum(y_i - \overline{y})^2}} = \frac{12}{\sqrt{200}} = 0.849$



Aufgabe 5

(a) Wertepaare: $(2,4)$, $(3,5)$, $(6,2)$, $(1,9)$, $(3,5)$

$x = (2,3,6,1,3) \quad \Rightarrow \quad \overline{x} = 3$

$y = (4,5,2,9,5) \quad \Rightarrow \quad \overline{y} = 5$

$$s_{xy} = \frac{(-1)\cdot(-1) + 0\cdot 0 + 3\cdot(-3) + (-2)\cdot 4 + 0\cdot 0}{4} = -4$$

(b) Die (empirische) Covarianz ist kommutativ, denn es gilt

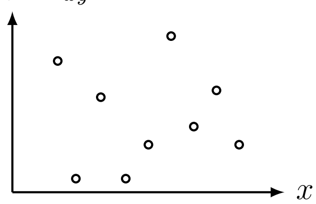
$$s_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}) = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \overline{y})(x_i - \overline{x}) = s_{yx}$$

Dasselbe gilt auch für die Kovarianz von Grundgesamtheiten.

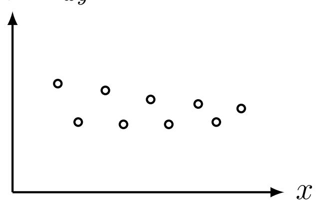
Aufgabe 6

Aufgabe 7

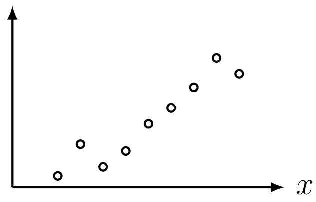
(a) y $\quad r_{xy} = 0.021$



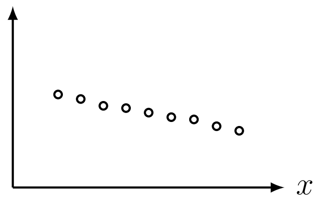
(b) y $\quad r_{xy} = -0.344$



(c) y $\quad r_{xy} = 0.946$



(d) y $\quad r_{xy} = -0.996$



Aufgabe 8

(a) $m(d) = -5.575 \cdot d + 123.1$; $r = 0.9953$ (gut!)

(b) 14.9.1999 ist Tag 14

$m(14) = -5.575 \cdot 14 + 123.1 \approx 45\,\mathrm{g}$

Aufgabe 9