

Einfache lineare Regression

Übungen

Aufgabe 1

Gegeben ist folgende gepaarte Stichprobe:

x	3	9	18
y	12	5	4

- ▶ Bestimme die Gleichung der Regressionsgeraden.
- ▶ Skizziere das Streudiagramm und die Regressionsgerade.
- ▶ Berechne den Korrelationskoeffizienten.

Aufgabe 1

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
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

Aufgabe 1

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
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

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

Aufgabe 1

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
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

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

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

Aufgabe 1

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
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

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

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

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

Aufgabe 1

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
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

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

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

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

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

Aufgabe 1

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
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

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

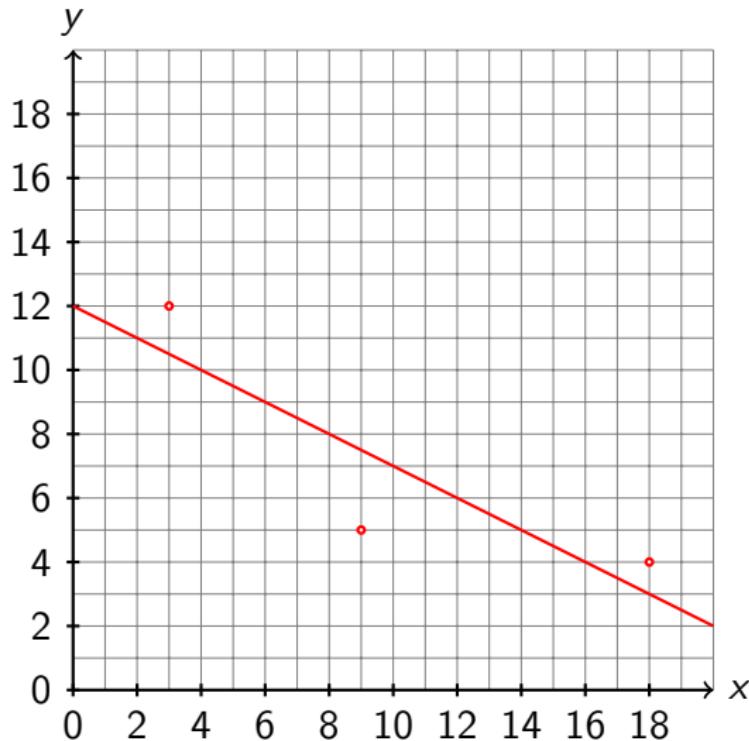
$$a = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} = \frac{-57}{114} = -\frac{1}{2}$$

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

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

Korrelation:

$$r_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \cdot \sum(y_i - \bar{y})^2}} = \frac{-57}{\sqrt{4332}} = -0.865$$



Aufgabe 2

Gegeben ist folgende gepaarte Stichprobe:

x	9	13	1	5
y	14	15	4	7

- ▶ Bestimme die Gleichung der Regressionsgeraden.
- ▶ Skizziere das Streudiagramm und die Regressionsgerade.
- ▶ Berechne den Korrelationskoeffizienten.

Aufgabe 2

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
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

Aufgabe 2

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
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

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

Aufgabe 2

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
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

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

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

Aufgabe 2

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
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

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

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

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

Aufgabe 2

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
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

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

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

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

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

Aufgabe 2

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
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

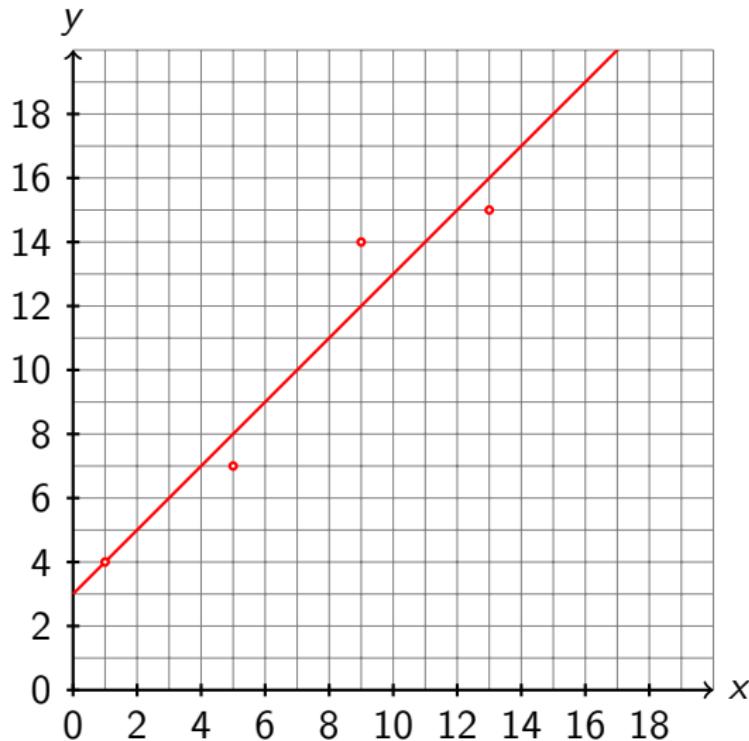
$$\bar{x} = 7, \bar{y} = 10$$

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

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

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

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



Aufgabe 3

Gegeben ist folgende gepaarte Stichprobe:

x	16	13	9	10
y	9	12	14	13

- ▶ Bestimme die Gleichung der Regressionsgeraden.
- ▶ Skizziere das Streudiagramm und die Regressionsgerade.
- ▶ Berechne den Korrelationskoeffizienten.

Aufgabe 3

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
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

Aufgabe 3

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
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

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

Aufgabe 3

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
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

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

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

Aufgabe 3

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
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

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

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

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

Aufgabe 3

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
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

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

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

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

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

Aufgabe 3

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
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

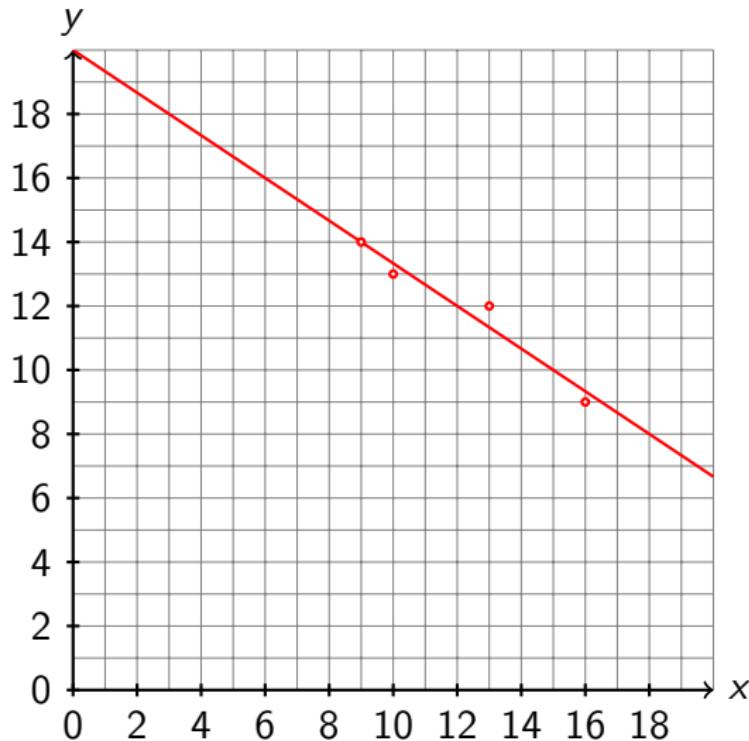
$$\bar{x} = 12, \bar{y} = 12$$

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

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

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

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



Aufgabe 4

Gegeben ist folgende gepaarte Stichprobe:

x	7	3	6	4
y	17	13	15	11

- ▶ Bestimme die Gleichung der Regressionsgeraden.
- ▶ Skizziere das Streudiagramm und die Regressionsgerade.
- ▶ Berechne den Korrelationskoeffizienten.

Aufgabe 4

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
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

Aufgabe 4

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
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

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

Aufgabe 4

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
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

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

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

Aufgabe 4

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
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

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

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

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

Aufgabe 4

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
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

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

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

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

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

Aufgabe 4

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
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

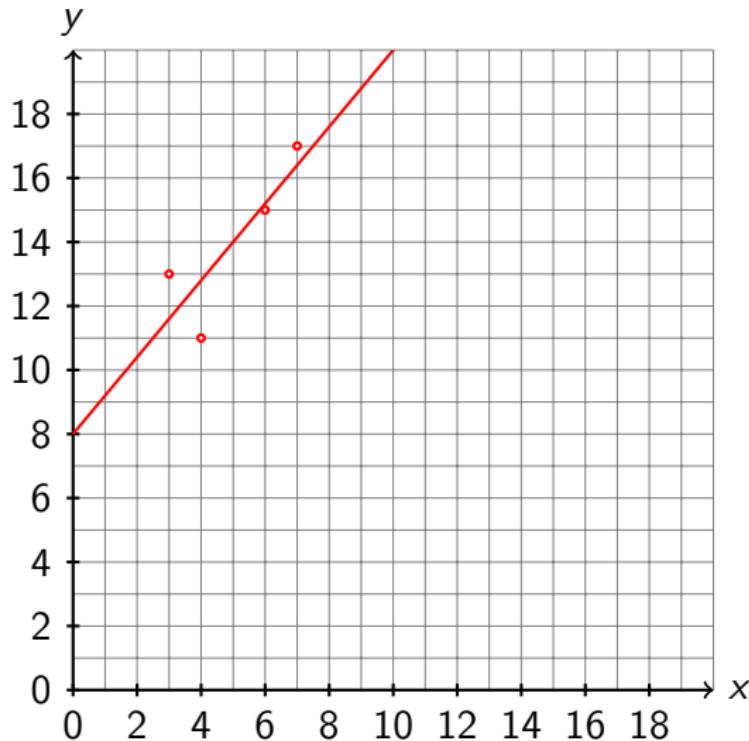
$$\bar{x} = 5, \bar{y} = 14$$

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

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

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

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



Aufgabe 5

- (a) Berechne die empirische Kovarianz $s_{xy} = \text{cov}(x, y)$ für die Wertepaare $(2, 4), (3, 5), (6, 2), (1, 9), (3, 5)$.
- (b) Gilt $s_{xy} = s_{yx}$? Begründe die Antwort.

Aufgabe 5

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

Aufgabe 5

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

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

Aufgabe 5

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

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

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

Aufgabe 5

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

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

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

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

Aufgabe 5

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

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

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

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

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

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

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

Dasselbe gilt auch für die Kovarianz von Grundgesamtheiten.

Aufgabe 6

Gegeben: $\bar{x} = 5$, $\bar{y} = 3$, $s_x = 2\sqrt{3}$, $s_y = 1$, $s_{xy} = 3$

Gesucht: Gleichung g der Regressionsgeraden und Korrelationskoeffizient.

Aufgabe 6

Aufgabe 7

Für vier Messreihen wurden die Korrelationskoeffizienten ermittelt.
Ordne diese Werte den folgenden Grafiken zu.

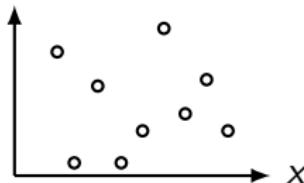
$$r_{xy} = 0.946$$

$$r_{xy} = -0.996$$

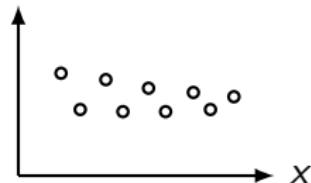
$$r_{xy} = 0.021$$

$$r_{xy} = -0.344$$

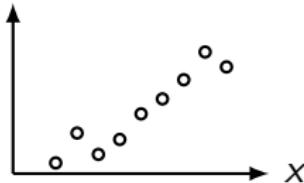
(a) y



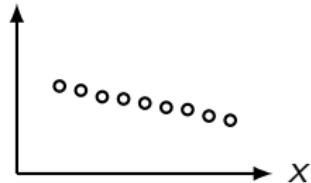
(b) y



(c) y

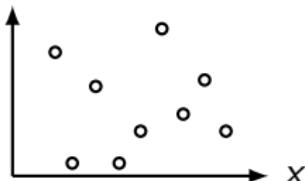


(d) y

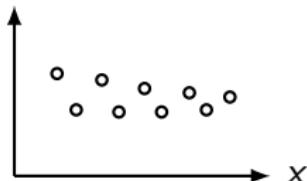


Aufgabe 7

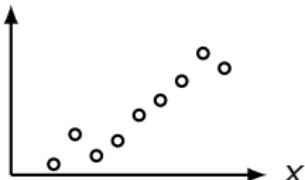
(a) $y \quad r_{xy} = 0.021$



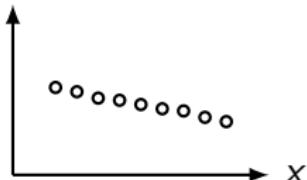
(b) $y \quad r_{xy} = -0.344$



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



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



Aufgabe 8

Rex Boggs von der Glenmore State High School, Rockhampton, Queensland, Australia hat untersucht, wie sich das Gewicht eines Stücks Seife im Laufe der Zeit verändert.

Datum	Tag	Gewicht (g)	Datum	Tag	Gewicht (g)
30.8.1999	0	124	10.9.1999	11	58
31.8.1999	1	121	11.9.1999	12	50
3.9.1999	4	103	16.9.1999	17	27
4.9.1999	5	96	18.9.1999	19	16
5.9.1999	6	90	19.9.1999	20	12
6.9.1999	7	84	20.9.1999	21	8
7.9.1999	8	78	21.9.1999	22	6
8.9.1999	9	71			

Am 22.9.1999 zerbrach die Seife in zwei Teile und eines davon verschwand im Abfluss.

- Berechne Regressionsgrade und Korrelationskoeffizient der Datenpaare (d, m) , wobei d für den Tag und m für die Masse des Seifenstücks (in Gramm) steht.
- Schätze das Gewicht der Seife am Dienstag, 14.9.1999.

Aufgabe 8

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

Aufgabe 8

- (a) $m(d) = -5.575 \cdot d + 123.1$; $r = 0.9953$ (gut!)
- (b) 14.9.1999 ist Tag 14

Aufgabe 8

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

(b) 14.9.1999 ist Tag 14

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

Aufgabe 9

- ▶ Gib die folgenden vier Datensätze jeweils in ein separates Arbeitsblatt eines Tabellenkalkulationsprogramms ein.
- ▶ Stelle die Datensatz in einem Streudiagramm dar.
- ▶ Führe eine lineare Regression durch. Was stellst du fest?

(a)	x	y
	10	8.04
	8	6.95
	13	7.58
	9	8.81
	11	8.33
	14	9.96
	6	7.24
	4	4.26
	12	10.84
	7	4.82
	5	5.68

(b)	x	y
	10	9.14
	8	8.14
	13	8.74
	9	8.77
	11	9.26
	14	8.1
	6	6.13
	4	3.1
	12	9.13
	7	7.26
	5	4.74

(c)	x	y
	10	7.46
	8	6.77
	13	12.74
	9	7.11
	11	7.81
	14	8.84
	6	6.08
	4	5.39
	12	8.15
	7	6.42
	5	5.73

(d)	x	y
	8	6.58
	8	5.76
	8	7.71
	8	8.84
	8	8.47
	8	7.04
	8	5.25
	19	12.5
	8	5.56
	8	7.91
	8	6.89

Aufgabe 9

