

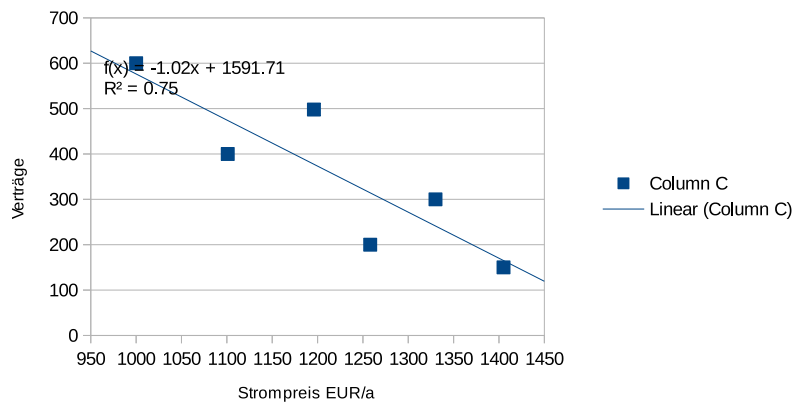
## Aufgabe 1

Welche der folgenden Aussagen sind richtig?

- (a) Richtig, die Varianz ist eine Summe quadratischer Größen.
- (b) Falsch, die Abweichung ordinaler Merkmale vom Median ist nicht definiert - also auch keine mittlere Abweichung.
- (c) Falsch, im allgemeinen sind Modus und Median natürlich verschieden.
- (d) Falsch, der Modus ist die am häufigsten in einer Stichprobe auftretende Merkmalsausprägung.
- (e) Richtig, der Quartilsabstand steigt mit der Breite der Verteilung an.

## Aufgabe 2

a) Daten mit Ausgleichsgerade



b) Anpassung einer Geraden  $y = a \cdot x + b$  über lineare Regression: Tabelle benötigter Daten

Preis	Verträge	$(x_i - \bar{x})$	$(y_i - \bar{y})$	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$
1000	600	-215	242	-52030	46225	58564
1101	400	-114	42	-4788	12996	1764
1196	498	-19	140	-2660	361	19600
1258	200	43	-158	-6794	1849	24964
1330	300	115	-58	-6670	13225	3364
1405	150	190	-208	-39520	36100	43264

Die Steigung der Geraden ist

$$a = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = -1,02$$

der Achsenabschnitt ist

$$b = \bar{y} - a\bar{x} = 1591,71$$

c) der Pearson-Koeffizient ist

$$r_{xy} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (1)$$

$$= \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} = -0,87 \quad (2)$$

Der Wert nahe eins legt einen relativ guten linearen Zusammenhang zwischen den Daten nahe.

### Aufgabe 3

Ergänzte Tabelle der absoluten Häufigkeiten:

Y	1	2	3	4	Summe
X					
1	2	4	1	9	16
2	17	34	10	85	146
3	3	6	2	15	26
Summe	22	44	13	109	188

Die beiden Merkmale sind nicht statistisch unabhängig, da die relativen Häufigkeiten in den Spalten sowie der Randspalte unterschiedlich sind.

Tabelle der relativen Häufigkeiten:

Y	1	2	3	4	Summe
X					
1	1/11	1/11	1/13	9/109	4/47
2	17/22	17/22	10/13	85/109	73/94
3	3/22	3/22	2/13	15/109	13/94
Summe	1	1	1	1	1

Für die Varianz wird die zweite Zeile der Tabelle benutzt, das arithmetische Mittel ist

$$\frac{1}{n} \sum_{i=1}^n y_i = \frac{1}{146} (17 \cdot 1 + 34 \cdot 2 + 10 \cdot 3 + 85 \cdot 4) = 3,12$$

Die Varianz

$$s^2(Y|X=2) = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2$$

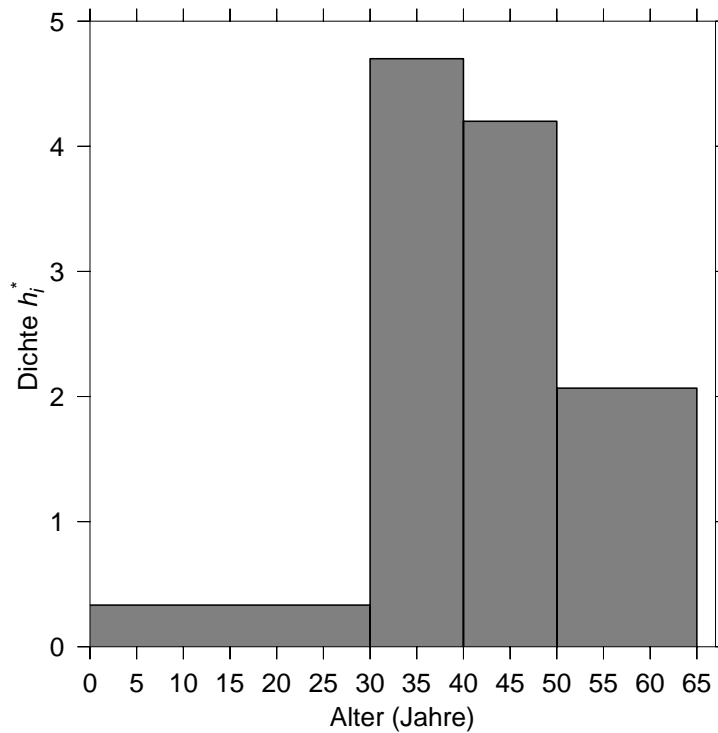
$$= \frac{1}{146} (17 \cdot (1 - 3,12)^2 + 34 \cdot (2 - 3,12)^2 + 10 \cdot (3 - 3,12)^2 + 85 \cdot (4 - 3,12)^2) = 1,27$$

#### Aufgabe 4

Daten zur Altersverteilung:

Alter in Jahren von ... bis unter ...	Absolute Häufigkeit	relative Häufigkeit $f_i$	kumulierte rel. Häufigkeit $F_i$	Klassenbreite $\Delta_i$	Dichte $h_i^* = h_i / \Delta_i$
bis 30	10	0,076	0,076	30	0,333
30 - 40	47	0,362	0,438	10	4,7
40 - 50	42	0,323	0,761	10	4,2
50 - 65	31	0,238	1	15	2,07

(a) Histogramm der Altersverteilung:  
Altersverteilung



(b) Berechnung des Durchschnittseinkommens (in tausend Credits):

$$\begin{aligned}\bar{y} &= \frac{1}{130} (10 \cdot 2,5 + 47 \cdot 4,2 + 42 \cdot 5,0 + 31 \cdot 4,9) \\ &= \frac{584,3}{130} = 4,49\end{aligned}$$

(c) Berechnung des Medians:

50% der befragten Personen werden in der 3.Klasse erreicht. Der Median lässt sich also wie folgt berechnen:

$$\begin{aligned}F(\bar{x}_Z) &= x_3^u + (x_3^o - x_3^u) \frac{F(\bar{x}_Z) - F(x_3^u)}{F(x_3^o) - F(x_3^u)} = x_3^u + (x_3^o - x_3^u) \frac{F(0,5) - F(x_3^u)}{F(x_3^o) - F(x_3^u)} \\ &= 40 + 10 \cdot \frac{0,5 - 57/130}{(99 - 57)/130} = 40 + 10 \cdot \frac{8}{42} = 41,9\end{aligned}$$

## Aufgabe 5

Tag	1	2	3	4	5	6	7
km ( $x$ )	15	16,5	17,5	18	18	20	22

a) Arithmetisches Mittel:

$$\bar{x} = \frac{15 + 16,5 + 17,5 + 18 + 18 + 20 + 22}{7} = \frac{127}{7} = 18,1429$$

b) Harmonisches Mittel:

$$\bar{x}_H = \frac{7}{\frac{1}{15} + \frac{1}{16,5} + \frac{1}{17,5} + \frac{1}{18} + \frac{1}{18} + \frac{1}{20} + \frac{1}{22}} = 17,9037$$

c) Die Durchschnittsgeschwindigkeit erhält man als Quotienten der gesamten zurückgelegten Strecke und der gesamten benötigten Zeit, also

$$\frac{15km + 16,5km + 17,5km + 18km + 18km + 20km + 22km}{7h}$$

Die Anwendung des arithmetischen Mittels ist hier korrekt.

d) Vielleicht!

## Aufgabe 6

a) Geometrisches Mittel:

$$\bar{x}_G = \sqrt[3]{(1 + 0,1) \cdot (1 + 0,15) \cdot (1 - 0,005)} - 1 = 7,97\%$$

b) Der Mittelwert berechnet sich nach

$$\bar{x} = \frac{1}{30}(10 \cdot 1 + 11 \cdot 2 + 8 \cdot 3 + 4) = \frac{60}{30} = 2,0.$$

Für die Kandidaten, die nicht bestanden haben, kann keine Note angegeben werden, sie werden zur Berechnung *nicht* herangezogen.

c) Insgesamt befragte Personen:  $100 + 1000 = 1100$ . Für die Einführung der neuen Regelung sind  $60 + 380 = 440$ . Also sind  $440/1100 = 40\%$  dafür.

d) Wenn Ted eine mittlere Geschwindigkeit von 60 km/h fahren will, benötigt er für die insgesamt 8 km Weg eine Zeit von  $8/60 \text{ h} = 8 \text{ min}$ . Da er aber für den Rückweg von 4km bereits eine Zeit von  $4/30 \text{ h} = 8 \text{ min}$  einplant, kann er die geplante Durchschnittsgeschwindigkeit nicht erreichen.