



2. Übungsblatt zur „Einführung in die Stochastik“

Aufgabe 6

(3 Punkte)

Gegeben sei die Messreihe

4, 5, 1, 3.7, 2.3, 4, 1.3, 2.7, 5, 3.3

- Bestimmen sie das Mittel, den Median, die Spannweite, die Varianz, die Standardabweichung und den Interquartilabstand dieser Messreihe.
- Zeichnen sie ein Histogramm dieser Messreihe bzgl. der Partition $\{[1, 2.5), [2.5, 5]\}$ des Intervalls $[1, 5]$.

Aufgabe 7

(3 Punkte)

Durch das in Abbildung 1 abgebildete Histogramm und den Boxplot wird eine Messreihe bestehend aus den Ankunftszeiten von 40 zufällig ausgewählten Studenten bei der Vorlesung “Statistik I für WiWi” am 26.10.2001 beschrieben:

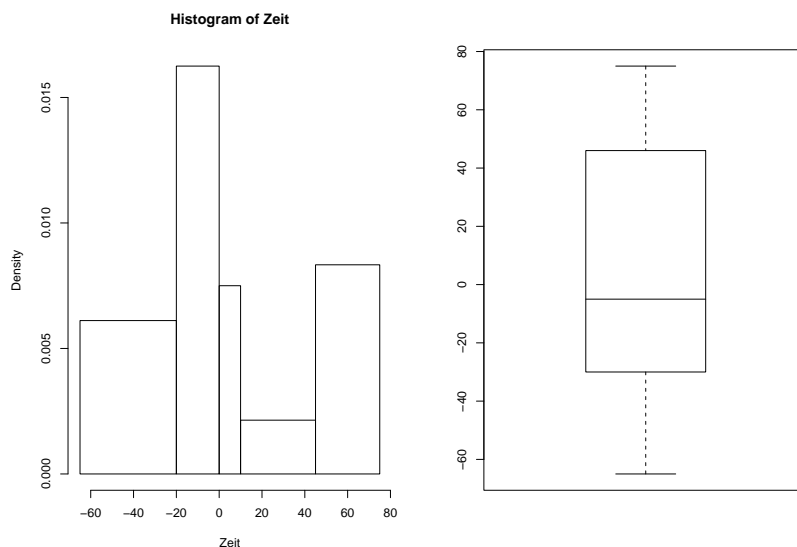


Abbildung 1: Abbildung zu Aufgabe 7.

- Wie kann man aus dem Histogramm approximativ ablesen, wieviele der 40 Studenten nicht mehr als 20 Minuten zu früh und gleichzeitig nicht mehr als 10 Minuten zu spät kamen?
- Wie groß ist approximativ der Median und der IQR dieser Messreihe ?

Aufgabe 8

(3 Punkte)

In der folgenden Tabelle sind die Ausgaben pro Student (in Euro) und die Arbeitslosenquote (in Prozent) in den sechs neuen Bundesländern im Jahr 2001 angegeben.

	Ausgaben pro Student (in Euro)	Arbeitslosenquote (in Prozent)
Berlin	8100	17.9
Brandenburg	6600	18.8
M.-V.	8700	19.6
Sachsen	8700	19
Sachsen-Anhalt	9900	20.9
Thüringen	8800	16.5

- Zeichnen Sie ein Streudiagramm (Scatterplot) der Daten, wobei sie als x -Wert die Ausgaben pro Student und als y -Wert die Arbeitslosenquote verwenden.
- Bestimmen Sie mit Hilfe der in diesem Kapitel hergeleiteten allgemeinen Formel die zugehörige Regressionsgerade und zeichnen Sie diese in das Streudiagramm aus a) ein.
- Inwieweit ändert sich das Resultat in b), wenn man den zu Sachsen-Anhalt gehörenden Datenpunkt weglässt?

Aufgabe 9

(3 Punkte)

- (a) Seien $x_1, y_1, \dots, x_n, y_n \in \mathbb{R}$, $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ und $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$. Zeigen Sie:

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

und

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

- Berechnen Sie die Korrelation der Daten aus Aufgabe 8.
- Was folgt aus b) für die Steigung der zugehörigen Regressionsgeraden ?
- Inwieweit ändert sich das Ergebnis aus b), wenn man vor Beginn der Berechnung der Korrelation die Ausgaben pro Student in Dollar und die Arbeitslosenquote in Promille umrechnet? Begründen Sie ihre Antwort.