

# Musterlösungen zum Aufgabenblatt 04: Streuungsmaße

## Aufgabe 4.1.

Diese Aufgabe bezieht sich auf den Datenbestand aus der Aufgabe 3.2 (Variable „Vorbereitungszeit“).

- Bestimmen Sie die Spannweite und die mittlere lineare Abweichung, die Standardabweichung und die Varianz dieser Häufigkeitsverteilung.
- Bestimmen Sie die Quartile aus der Tabelle und graphisch aus dem Verlauf der Verteilungsfunktion. Berechnen Sie für beide Varianten den Quartilsabstand sowie den Semi-quartilsabstand. Vergleichen Sie die Ergebnisse.
- Eine Befragung an der TU Dortmund ergab bezüglich der Frage „Wie viel Zeit verwenden Sie pro Woche für die Vorbereitung auf Ihre Kurse in der Uni?“ ein arithmetisches Mittel von 8,8 Stunden mit einer Standardabweichung von 3,1 Stunden. Vergleichen Sie die Verteilungen in Bochum und in Dortmund anhand der Variationskoeffizienten.

(a) Hier liegt eine Häufigkeitsverteilung einer stetigen Variablen vor. Dementsprechend werden die Formeln für eine Häufigkeitsverteilung angewendet.

In der Aufgabe 3.2 wurde bereits bestimmt:  $AM = \frac{\sum x_i f_i}{\sum f_i} = \frac{544}{80} = 6,8$

Zeitintervall	$f_i$	$x_i$ (Klassenmitte)	$x_i - \bar{x}$	$ x_i - \bar{x} $	$ x_i - \bar{x}  * f_i$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 * f_i$
0 bis unter 5 Stunden	18	2,5	-4,3	4,3	77,4	18,49	332,82
5 bis unter 8 Stunden	30	6,5	-0,3	0,3	9	0,09	2,7
8 bis unter 9 Stunden	16	8,5	1,7	1,7	27,2	2,89	46,24
9 bis unter 12 Stunden	16	10,5	3,7	3,7	59,2	13,69	219,04
<b>Summe</b>	<b>80</b>				<b>172,8</b>		<b>600,8</b>

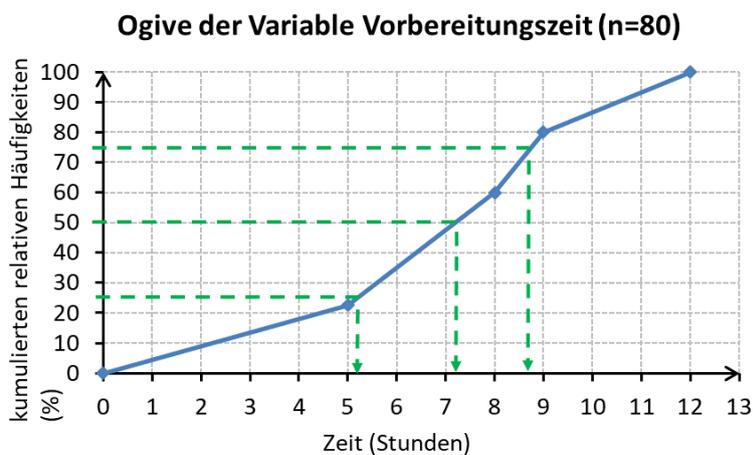
- Bestimmung der Spannweite:  $Spannweite = x_{max} - x_{min} = 12 - 0 = 12$
- Bestimmung der mittleren linearen Abweichung:  $AD = \frac{\sum |x_i - \bar{x}| * f_i}{\sum f_i} = \frac{172,8}{80} = 2,16$
- Bestimmung der Varianz:  $var_x = \frac{\sum (x_i - \bar{x})^2 * f_i}{n} = \frac{600,8}{80} = 7,51$
- Bestimmung der Standardabweichung:  $SD = \sqrt{var_x} = \sqrt{7,51} \approx 2,74$

(b) Die **Bestimmung der Quartile aus der Tabelle** erfolgt anhand der kumulierten relativen Häufigkeiten, die bereits in der Aufgabe 3.2 bestimmt wurden. Um die Quartile zu quantifizieren, wird die Klassenmitte ( $x_i$ ) angegeben, **nicht** die gesamte Klasse. Die Klassenmitte repräsentiert die Klasse.

Zeitintervall	$f_i$	Klassenmitte $x_i$	$f'_i = \text{in \%}$	$F'_i = \text{in \%}$
0 bis unter 5 Stunden	18	2,5	22,5 %	22,5 %
5 bis unter 8 Stunden	30	6,5	37,5 %	60 %
8 bis unter 9 Stunden	16	8,5	20 %	80 %
9 bis unter 12 Stunden	16	10,5	20 %	100 %
<b>Summe</b>	<b>80</b>		<b>100 %</b>	

- **Erstes Quartil:** Q1 liegt bei  $F'_i = 25 \%$ . Der Wert  $F'_i = 25 \%$  befindet sich bei 6,5 Stunden
- **Zweites Quartil:** Q2 liegt bei  $F'_i = 50 \%$ . Der Wert  $F'_i = 50 \%$  befindet sich bei 6,5 Stunden. Das zweite Quartil ist außerdem der Median und wurde bereits in der Aufgabe 3.2 bestimmt.
- **Drittes Quartil:** Q3 liegt bei  $F'_i = 75 \%$ . Der Wert  $F'_i = 75 \%$  befindet sich bei 8,5 Stunden.
- **Quartilsabstand:**  $QA = Q3 - Q1 = 8,5 - 6,5 = 2$  Stunden
- **Semiquartilsabstand:**  $SQA = \frac{Q3 - Q1}{2} = 1$  Stunden

**Bestimmung der Quartile graphisch aus der Ogive** erfolgt anhand der Zeichnung, die auch in der Aufgabe 3.2 verwendet wurde. Für die Bestimmung siehe grüne gestrichelte Linien.



- $Q1 \approx 5,2$  Stunden
- $Q2 = \text{Median} \approx 7,2$  Stunden
- $Q3 \approx 8,7$  Stunden
- $QA = Q3 - Q1 = 8,7 - 5,2 = 3,5$  Stunden
- $SQA = \frac{Q3 - Q1}{2} = 1,75$  Stunden

**Vergleich:** Die Werte, die aus der Ogive graphisch bestimmt wurden, sind genauer (vgl. die Argumentation bei der Aufgabe 3.2). Die genauen Werte sind kleiner, daher kann man sagen, dass in diesem Beispiel die Streuung bei der tabellarischen Bestimmung überschätzt wird.

(c) Der **Variationskoeffizient**:  $VK = \frac{s_x}{\bar{x}} * 100\%$

- $VK (\text{Dortmund}) = \frac{3,1}{8,8} * 100\% \approx 35,23\%$
- $VK (\text{Bochum}) = \frac{2,74}{6,8} * 100\% \approx 40,30\%$  (\*)

**Vergleich:** Dortmund hat, absolut gesehen, eine größere Streuung als Bochum (3,1 vs. 2,74). Aber, in der Relation zum Mittelwert, hat Dortmund eine geringere Streuung (35,23 % vs. 40,3 %) als Bochum.

(\*) Anmerkung: Wenn mit gerundeten Werten gerechnet wird, dann  $VK (\text{Bochum}) = \frac{2,74}{6,8} * 100\% \approx 40,29\%$ . Die Abweichungen, die auf das Runden zurückgeführt werden können, stellen kein Problem für die Lösung der Aufgabe dar.

**Der Rest der Datei steht Ihnen nicht zur Verfügung.**