

## Lösungen zu den Übungsaufgaben in Kapitel 6

(1) Bestimmen Sie den Median der drei folgenden gruppierten intervallskalierten Messwertreihen. Angegeben sind die Kategorienmitten.

a) 1; 5; 2; 5; 3; 6; 1; 5; 7; 7

b) 31; 25; 16; 22; 24; 30

c) 101; 104; 103; 104; 110; 104; 108; 104; 102; 104

Formal lässt sich der Median bei gruppierten Daten gemäß Formel F 6.9 wie folgt bestimmen:

$$Md = c_{l-1} + \frac{0,5 - \sum_{j=1}^{l-1} h_j}{h_l} \cdot d_l$$

wobei

- $j$ : Merkmalsklasse,
- $l$ : Medianklasse (Klasse, in der  $kh_j$  erstmals größer als 0,5 ist),
- $c_{l-1}$ : untere Klassengrenze der Medianklasse,
- $c_l$ : obere Klassengrenze der Medianklasse,
- $h_j$ : relative Häufigkeit der Klasse  $j$ ,
- $h_l$ : relative Häufigkeit der Medianklasse  $l$ ,
- $d_l = c_l - c_{l-1}$ : Breite der Medianklasse  $l$ .

Die Messwertreihe (a) lautet in geordneter Form  $\{1; 1; 2; 3; 5; 5; 5; 6; 7; 7\}$ . Die Medianklasse ist diejenige Klasse, in der  $kh_j$  erstmals größer als 0,5 ist. Das ist der Fall in der Klasse  $l = 5$  mit einer Untergrenze von  $c_{l-1} = 4,5$  und einer Obergrenze von  $c_l = 5,5$ . Unterhalb der Medianklasse liegen vier weitere Merkmalsklassen mit den relativen Häufigkeiten  $h_1 = 0,2$ ;  $h_2 = 0,1$ ;  $h_3 = 0,1$  und  $h_4 = 0$  (denn der Wert 4 kommt ja hier nicht vor). Die Medianklasse selbst hat eine relative Häufigkeit von  $h_l = h_5 = 0,3$ . Die Kategorienbreite beträgt in allen Klassen  $d = 1$ . Setzt man diese Werte in Gleichung F 6.9 ein, erhält man

$$Md = 4,5 + \frac{0,5 - 0,4}{0,3} \cdot 1 = 4,833$$

Die Messwertreihe (b) lautet in geordneter Form  $\{16; 22; 24; 25; 30; 31\}$ . Die Medianklasse hat eine Untergrenze von  $c_{l-1} = 24,5$  und eine Obergrenze von  $c_l = 25,5$ . Unterhalb der

Medianklasse liegen neun weitere Merkmalsklassen. Die erste Merkmalsklasse weist die Klassenmitte 16 auf, die zweite die Klassenmitte 17 etc. Die relativen Häufigkeiten lauten dann  $h_1 = 1/6; h_2 = 0; h_3 = 0; h_4 = 0; h_4 = 0; h_5 = 0; h_6 = 0; h_7 = 1/6; h_8 = 0$  und  $h_9 = 1/6$ . Die Medianklasse selbst hat eine relative Häufigkeit von  $h_l = h_{10} = 1/6$ . Die Kategorienbreite beträgt in allen Klassen  $d = 1$ . Setzt man diese Werte in Gleichung F 6.9 ein, erhält man

$$Md = 24,5 + \frac{0,5 - 0,5}{0,17} \cdot 1 = 24,5$$

Die Messwertreihe (c) lautet in geordneter Form  $\{101; 102; 103; 104; 104; 104; 104; 104; 108; 110\}$ . Die Medianklasse hat eine Untergrenze von  $c_{l-1} = 103,5$  und einer Obergrenze von  $c_l = 104,5$ . Unterhalb der Medianklasse liegen drei weitere Merkmalsklassen mit den relativen Häufigkeiten  $h_1 = 0,1; h_2 = 0,1$  und  $h_3 = 0,1$ . Die Medianklasse selbst hat eine relative Häufigkeit von  $h_l = h_4 = 0,5$ . Die Kategorienbreite beträgt in allen Klassen  $d = 1$ . Setzt man diese Werte in Gleichung F 6.9 ein, erhält man

$$Md = 103,5 + \frac{0,5 - 0,3}{0,5} \cdot 1 = 103,9$$

**(2) Demonstrieren Sie die Eigenschaften des Mittelwertes an einer kleinen Messwertreihe eigener Wahl!**

Beispiel: Die Messwertreihe  $\{1; 2; 2; 3; 6; 10\}$ . Der Mittelwert ist bei dieser Messwertreihe  $\bar{x} = 4$ . Die Eigenschaften des Mittelwerts lauten:

- (1) Die Summe der Abweichungen aller Messwerte vom Mittelwert beträgt stets Null.

$$\sum_{m=1}^n (x_m - \bar{x}) = 0$$

In unserem Beispiel lauten die Abweichungen  $(x_m - \bar{x}) = (1 - 4) = -3; (2 - 4) = -2; (2 - 4) = -2; (3 - 4) = -1; (6 - 4) = 2; (10 - 4) = 6$ . Die Summe dieser Abweichungen ergibt  $-3 - 2 - 2 - 1 + 2 + 6 = 0$ .

- (2) Die Summe der *quadrierten* Abweichungen der Messwerte vom Mittelwert ist stets kleiner als die Summe der quadrierten Abweichungen von irgendeinem anderen Wert.

$$\sum_{m=1}^n (x_m - \bar{x})^2 = \min.$$

Die quadrierten Abweichungen lauten  $(x_m - \bar{x})^2 = (1 - 4)^2 = 9; (2 - 4)^2 = 4; (2 - 4)^2 = 4;$   
 $(3 - 4)^2 = 1; (6 - 4)^2 = 4; (10 - 4)^2 = 36$ . Die Summe dieser Abweichungen ergibt QS =  
 $9 + 4 + 4 + 1 + 4 + 36 = 58$ . Es gibt keinen anderen Wert (als 4), für den gelten würde,  
 dass die Quadratsumme kleiner ist als 58.

- (3) Wird zur Variablen  $X$  (d.h. zu jedem Messwert  $x_m$ ) eine Konstante  $a$  addiert, verändert sich der Mittelwert additiv um eben diese Konstante  $a$ . Formal:

$$y_m = x_m + a \quad \Rightarrow \quad \bar{y} = \bar{x} + a$$

Würde man in unserem Beispiel zu jedem Messwert eine Konstante  $a = 5$  hinzuaddieren, so würde die Messwertreihe lauten  $\{6; 7; 7; 8; 11; 15\}$ . Deren Mittelwert wäre dann  $\bar{y} = 9$ . Das entspricht  $\bar{y} = \bar{x} + a = 4 + 5 = 9$ .

- (4) Wird die Variable  $X$  (d.h. jeder Messwert  $x_m$ ) mit einer Konstante  $b$  multipliziert, verändert sich der Mittelwert multiplikativ um eben diese Konstante  $b$ . Formal:

$$y_m = b \cdot x_m \quad \Rightarrow \quad \bar{y} = b \cdot \bar{x}$$

Würde man in unserem Beispiel jeden Messwert mit  $b = 3$  multiplizieren, so würde die Messwertreihe lauten  $\{3; 6; 6; 9; 18; 30\}$ . Deren Mittelwert wäre dann  $\bar{y} = 12$ . Das entspricht  $\bar{y} = b \cdot \bar{x} = 3 \cdot 4 = 12$ .

- (3) Berechnen Sie die Varianz und die Standardabweichung der Messwerte in der vierten Zeile der Urliste in Tabelle 6.8!**

Die Messwerte in dieser Zeile lauten  $\{89; 63; 55; 93; 71; 81; 72; 68; 75; 93\}$ . Die empirische Varianz dieser Messwertreihe beträgt  $s_x^2 = 148,8$ , die empirische Standardabweichung  $s_x = 12,20$ .

- (4) Vollziehen Sie die Eigenschaften von Varianz und Standardabweichung an einer kleinen Messwertreihe eigener Wahl nach!**

Beispiel: Die Messwertreihe  $\{1; 2; 2; 4; 4; 5\}$ . Die empirische Varianz dieser Messwertreihe beträgt  $s_x^2 = 2$ . Die Eigenschaften der Varianz (und der Standardabweichung) lauten:

- (1) Die Varianz reagiert sensibel auf Ausreißer.

Würde man in der obigen Messwertreihe den Wert 5 durch den Wert 10 austauschen, würde die Varianz  $s_X^2 = 8,81$  betragen.

- (2) Wird zur Variable  $X$  (d.h. zu jedem Messwert  $x_m$ ) eine Konstante  $a$  addiert, bleiben die Varianz und die Standardabweichung davon gänzlich unberührt.

$$y_m = x_m + a \quad \Rightarrow \quad s_Y^2 = s_X^2$$

Würde man in unserem Beispiel zu jedem Messwert eine Konstante  $a = 5$  hinzuaddieren, so würde die Messwertreihe lauten  $\{6; 7; 7; 9; 9; 10\}$ . Die empirische Varianz dieser Messwertreihe wäre ebenfalls  $s_Y^2 = s_X^2 = 2$ .

- (3) Wird die Variable  $X$  (d.h. jeder Messwert  $x_m$ ) mit einer Konstante  $b$  multipliziert, verändert sich die Varianz um den Faktor  $b^2$ :

$$y_m = b \cdot x_m \quad \Rightarrow \quad s_Y^2 = b^2 \cdot s_X^2$$

Würde man in unserem Beispiel jeden Messwert mit  $b = 3$  multiplizieren, so würde die Messwertreihe lauten  $\{3; 6; 6; 12; 12; 15\}$ . Die empirische Varianz dieser Messwertreihe wäre dann  $s_Y^2 = 18$ . Dies entspricht  $s_Y^2 = b^2 \cdot s_X^2 = 3^2 \cdot 2 = 9 \cdot 2 = 18$ .

- (5) Transformieren Sie die folgenden Messwerte in z-Werte:  $x_1 = 0$ ;  $x_2 = 1$ ;  $x_3 = 2$ ;  $x_4 = 2$ ;  $x_5 = 5$ .**

Der Mittelwert dieser Messwertreihe ist  $\bar{x} = 2$ , die empirische Standardabweichung  $s_X = 1,67$ . In z-Werte transformiert lautet die Messwertreihe also  $\{-1,20; -0,60; 0; 0; 1,79\}$ .

- (6) In den Online-Materialien sind die Messwerte der Urliste aus Tabelle 6.8 in der Datei mit dem Namen „daten61.txt“ abgelegt. Lesen Sie diese Datei in ein Tabellenkalkulationsprogramm (z.B. Excel) oder ein Statistikprogramm (z.B. SPSS oder R) ein und finden Sie die Prozeduren, mit denen Mittelwert, Varianz und Standardabweichung berechnet werden! Welche Werte werden für diese Kennwerte angegeben?**

Der Mittelwert dieser Messwertreihe beträgt  $\bar{x} = 66,73$ , die empirische Varianz  $s_X^2 = 251,58$  und die empirische Standardabweichung  $s_X = 15,86$ .