

2 Univariate deskriptive Statistik

Aufgabe 7

Der Besitzer des Kinos *Cinemia* macht sich Gedanken über die Wirtschaftlichkeit seines Hauses. An 100 Tagen zählt er daher die Anzahl der Besucher.

Besucherzahl (a_j)	41	42	43	44	45	46	47	48	49	50	51
Anzahl der Tage (h_j)	1	9	13	13	20	15	10	7	5	4	3

- (a) Wie heißt das untersuchte Merkmal und wie ist es skaliert?
- (b) Berechnen Sie die relativen und kumulierten relativen Häufigkeiten.
- (c) Stellen Sie die relativen Häufigkeiten in einem Stabdiagramm dar.
- (d) Zeichnen Sie die empirische Verteilungsfunktion.
- (e) Wie groß ist der Anteil der Tage, an denen der Kinobesitzer weniger als 270 Euro einnimmt? Dabei wird ein Einheitspreis von 6 Euro erhoben.
- (f) An wieviel Prozent der Tage kommen maximal 50, mindestens aber 45 Besucher?

Aufgabe 8

Für die Daten x_1, \dots, x_n soll ein Histogramm mit k Klassen und Klassengrenzen $c_0 < c_1 < \dots < c_k$, $c_0 \leq \min_i x_i$, $\max_i x_i \leq c_k$ betrachtet werden. Für die abzutragende Höhe \tilde{f}_j der j -ten Klasse ($j = 1, \dots, k$) mit Klassenbreite $d_j = d$ für alle j gilt

$$\tilde{f}_j(x) = \begin{cases} f_j/d, & x \in [c_{j-1}, c_j), \\ 0, & \text{sonst.} \end{cases}$$

- (a) Zeigen Sie, dass unter Berücksichtigung aller k Klassen für die abzutragende Höhe an einer beliebigen Stelle $x \in \mathbb{R}$ gilt

$$\tilde{f}(x) = \frac{1}{nd} \sum_{i=1}^n \sum_{j=1}^k 1_{\{x_i \in [c_{j-1}, c_j)\}} 1_{\{x \in [c_{j-1}, c_j)\}}, \text{ mit } 1_{\{x \in A\}} = \begin{cases} 1, & x \in A, \\ 0, & \text{sonst.} \end{cases}$$

- (b) Zeigen Sie, dass

$$\int_{-\infty}^{\infty} \tilde{f}(x) dx = 1.$$

Aufgabe 9* (20 Punkte)

Unter den Teilnehmern der Veranstaltung “Deskriptive Statistik” im Wintersemester 2014/2015 wurde eine Umfrage durchgeführt. Unter anderem wurde das Geschlecht und die Mathenote aus der Schule erhoben. Der in der folgenden Tabellen abgedruckte Teildatensatz mit 30 Befragten soll für eine erste Analyse mit Hilfe graphischer Verfahren dargestellt werden.

Person i	Geschlecht	Mathenote	Person i	Geschlecht	Mathenote
1	2	3	2	1	2
3	1	1	4	1	2
5	2	2	6	1	1
7	2	2	8	1	2
9	2	3	10	1	3
11	1	2	12	2	1
13	2	3	14	2	1
15	1	3	16	2	3
17	2	2	18	1	1
19	2	1	20	2	1
21	2	2	22	1	2
23	1	3	24	1	5
25	1	2	26	2	1
27	1	3	28	2	3
29	2	4	30	1	1

- Bestimmen Sie die Grundgesamtheit dieser Erhebung und die statistischen Einheiten! Durch welche Identifikationsmerkmale ist die Grundgesamtheit festgelegt?
- Bestimmen Sie Methode, Datenform und Umfang der Erhebung!
- Benennen Sie die erhobenen Merkmale dieses Teildatensatzes und bestimmen Sie den zugehörigen Merkmalsraum.
- Erstellen Sie eine Häufigkeitstabelle für das Merkmal *Mathenote*, bestehend aus den absoluten, relativen und kumulierten (relativen) Häufigkeiten.
- Unterteilen Sie die Stichprobe nach Geschlecht (1: weiblich, 2: männlich). Bestimmen Sie nun für beide Gruppen getrennt die relativen Häufigkeiten des Merkmals *Mathenote* und interpretieren Sie das Ergebnis.
- Erstellen Sie die empirische Verteilungsfunktion des Merkmals *Mathenote* der Absolvent(inn)en. Welche Note war die schlechteste unter den besten 20% der Absolvent(inn)en. Welche Note war die beste unter den schlechtesten 20% der Befragten?

Aufgabe 10

25 Studenten der Universität X in X-hausen wurden im Juni 2012 nach dem Studiengang, der Zahl der Geschwister und nach dem Einkommen befragt. Das Ergebnis war:

i	Name	Studiengang	Anzahl der Geschwister	Einkommen (in Euro)
1	Martin A.	Mathematik	0	472
2	Ute A.	Informatik	1	403
3	Wilhelm A.	Statistik	0	698
4	Kurt B.	Statistik	1	349
5	Sven B.	Physik	1	380
6	Mike D.	Physik	2	327
7	Klaus D.	Informatik	2	323
8	Theo E.	Mathematik	1	416
9	Jean F.	Physik	1	398
10	Elvira G.	Statistik	0	543
11	Karl H.	Statistik	0	629
12	Andreas K.	Mathematik	1	350
13	Thomas K.	Statistik	0	435
14	Chris L.	Informatik	3	328
15	Uwe L.	Physik	2	327
16	Axel M.	Statistik	0	435
17	Maria M.	Statistik	1	349
18	Ruth M.	Informatik	0	315
19	Bärbel N.	Statistik	1	349
20	Armin R.	Statistik	2	349
21	Christa R.	Mathematik	1	337
22	Bernd S.	Statistik	1	736
23	Claudia S.	Informatik	3	406
24	Edmund T.	Mathematik	0	337
25	Claudia W.	Physik	1	327

- (a) Wie ist das Merkmal „Studiengang“ skaliert? Ermitteln Sie aus den Daten die absoluten und relativen Häufigkeiten. Stellen Sie die relativen Häufigkeiten in einem Säulendiagramm dar.
- (b) Wie ist das Merkmal „Anzahl der Geschwister“ skaliert? Ermitteln Sie die absoluten und relativen Häufigkeiten sowie die kumulierten Häufigkeiten. Stellen Sie die relativen Häufigkeiten in einem Kreisdiagramm dar.
- (c) Beantworten Sie folgende Fragen:
- Wieviele Studenten haben höchstens zwei Geschwister?
 - Wieviel Prozent der Studenten haben mindestens zwei Geschwister?
 - Wieviel Prozent der Studenten haben ein oder zwei Geschwister?
- (d) Wie ist das Merkmal „Einkommen“ skaliert? Stellen Sie die Häufigkeitsverteilung in einem Stammblatt-Diagramm dar. Interpretieren Sie das Diagramm.