

1. Zusatzübungsblatt zur Vorlesung
Deskriptive Statistik und Wahrscheinlichkeitsrechnung SS 2021

Aufgabe 1 $(1 + 5 + 3 + 2 + 1 + 5 + 3 = 20 \text{ Punkte})$

Bei einer Umfrage wurden 40 Personen darüber befragt, wie viele Mobiltelefone sie in den vergangenen 5 Jahren gekauft haben. Das Ergebnis der Umfrage ist die folgende (bereits aufsteigend sortierte) Urliste:

0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 5, 5, 5, 5, 6

- (a) Geben Sie die Menge A der aufgetretenen Merkmalswerte an.
- (b) Erstellen Sie eine Tabelle mit den absoluten und relativen Häufigkeiten.
- (c) Stellen Sie die zugehörige empirische Verteilungsfunktion auf.
- (d) Wie viele Personen haben in den vergangenen 5 Jahren mehr als 1, aber höchstens 5 Mobiltelefone gekauft?
- (e) Wie groß ist der Anteil der Personen, die in den vergangenen 5 Jahren mehr als 3 Mobiltelefone gekauft haben?
- (f) Berechnen Sie das arithmetische Mittel und die empirische Standardabweichung des erhobenen Merkmals „Anzahl der in den vergangenen 5 Jahren gekauften Mobiltelefone“.
- (g) Geben Sie das untere Quartil, den Median sowie das obere Quartil des Merkmals „Anzahl der in den vergangenen 5 Jahren gekauften Mobiltelefone“ an. Verwenden Sie bei Bedarf die aus der Vorlesung bekannten Konventionen für kardinalskalierte Merkmale, um Eindeutigkeit zu erreichen.

Aufgabe 2 $(8 + 3 + 4 + 5 + 3 + 11 = 34 \text{ Punkte})$

Zu einem kardinalskalierten Merkmal sei die folgende (zur einfacheren Bearbeitung der Aufgabe bereits sortierte) Urliste der Länge $n = 40$ gegeben:

3.92, 4.46, 4.77, 6.92, 9.90, 10.54, 11.94, 12.23, 12.87, 12.92, 13.82, 14.40, 15.28, 16.44, 16.82, 17.44, 17.71, 17.73, 19.80, 21.12, 21.13, 22.42, 22.88, 22.91, 23.37, 23.40, 25.36, 25.78, 26.10, 26.17, 26.37, 27.02, 30.94, 32.79, 34.75, 35.50, 38.43, 42.50, 44.53, 47.64

- (a) Führen Sie eine Klassierung der erhobenen Daten auf Grundlage der Klassen

$$K_1 = (0, 10], K_2 = (10, 20], K_3 = (20, 25], K_4 = (25, 30], K_5 = (30, 40], K_6 = (40, 50]$$

durch. Geben Sie insbesondere die jeweiligen Klassenbreiten, Klassenmitten, absoluten und relativen Klassenhäufigkeiten, Häufigkeitsdichten sowie die Werte der empirischen Verteilungsfunktion an den Klassengrenzen an.

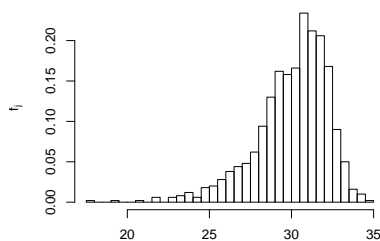
- (b) Berechnen Sie aus den klassierten Daten den (approximativen) arithmetischen Mittelwert der Daten. Wie groß ist die relative Abweichung vom tatsächlichen (aus der Urliste bestimmten) Mittelwert von 21.526?
- (c) Bestimmen Sie aus den klassierten Daten (näherungsweise) die empirische Standardabweichung.
- (d) Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.
- (e) Bestimmen Sie (aus der Urliste) die Anzahl von Merkmalswerten zwischen 20 und 35. Welche Näherung für diese Anzahl können Sie aus der in Teil (d) aufgestellten empirischen Verteilungsfunktion berechnen?
- (f) Bestimmen Sie unteres Quartil, Median und oberes Quartil sowohl exakt aus der Urliste (eindeutig mit den aus der Vorlesung bekannten Konventionen für kardinalskalierte Merkmale) als auch mit Hilfe der approximativen Verteilungsfunktion für die klassierten Daten. Berechnen Sie jeweils auch den Interquartilsabstand.

Aufgabe 3 (6 Punkte)

Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.

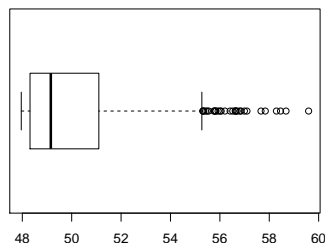
Richtige Antworten geben +3 Punkte, falsche Antworten und nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).

1. Kreuzen Sie an, auf welche Merkmalseigenschaften das folgende Histogramm *am ehesten* hindeutet:



- (a) leptokurtisch und rechtssteil
- (b) leptokurtisch und linkssteil
- (c) platykurtisch und rechtssteil
- (d) platykurtisch und linkssteil

2. Kreuzen Sie an, auf welche Merkmalseigenschaften der folgende Box-Plot *am ehesten* hindeutet:



- (a) leptokurtisch und rechtssteil
- (b) leptokurtisch und linkssteil
- (c) platykurtisch und rechtssteil
- (d) platykurtisch und linkssteil

Aufgabe 4 (1 + 2 + 3 + 2 + 12 + 5 = 25 Punkte)

Bei den 130 Schülerinnen und Schülern einer (fiktiven) Klassenstufe wurde — getrennt nach Geschlecht (X) — das monatliche Taschengeld Y (in €) erhoben. Das Ergebnis ist nach Klassierung des Merkmals „Taschengeld“ in der folgenden Tabelle der gemeinsamen absoluten Häufigkeiten zusammengefasst:

| $X \setminus Y$ | (0, 20] | (20, 40] | (40, 60] | $h_{i.}$ |
|-----------------|---------|----------|----------|----------|
| männlich | 11 | 27 | 12 | |
| weiblich | 19 | 48 | 13 | |
| $h_{.j}$ | | | | |

- (a) Geben Sie die Menge der Merkmalsausprägungen A von X bzw. B von Y an.
- (b) Ergänzen Sie die Tabelle der gemeinsamen Häufigkeiten in den vorgesehenen Feldern um die jeweiligen Randhäufigkeiten.
- (c) Erstellen Sie eine Tabelle mit den bedingten (relativen) Häufigkeiten von $Y|X = a$ für alle $a \in A$.
- (d) Sind X und Y unabhängig? Begründen Sie Ihre Antwort!
- (e) Berechnen Sie den korrigierten Pearsonschen Kontingenzkoeffizient von X und Y .
- (f) Bekommen (auf Basis der klassierten Daten) im Mittel Mädchen oder Jungen in dieser Klassenstufe mehr Taschengeld?

Bitte notieren Sie (nach Bearbeitung aller Aufgaben) die Lösungen zu Aufgabe 3 sowie zu Teil (b) der Aufgabe 4 nochmals auf Ihre Lösungsblätter, falls Sie diese Aufgaben(teile) auf dem Übungsblatt bearbeitet haben und dieses nicht mit abgeben möchten!