

1. Übungsblatt zum Wiederholungskurs
Deskriptive Statistik und Wahrscheinlichkeitsrechnung WS 2023/24

Aufgabe 1

Geben Sie bei den folgenden Merkmalen an, welches Skalierungsniveau (nominalskaliert, ordinalskaliert oder kardinalskaliert) vorliegt, ob das Merkmal qualitativ oder quantitativ ist, und gegebenenfalls, ob es sich eher um ein diskretes oder stetiges Merkmal handelt.

- (a) Anzahl der Vereinsmitgliedschaften
- (b) Netto-Jahresgehalt 2023
- (c) Augenfarbe
- (d) Weltranglistenplatz einer Fußballnationalmannschaft
- (e) Persönliche Bestleistung im Kugelstoßen

Aufgabe 2

Bei einer Umfrage wurden 50 Autofahrer befragt, an wieviel Tagen in der Woche sie ihr Auto benutzen. Das Ergebnis der Umfrage ist die folgende (bereits aufsteigend sortierte) Urliste:

1, 1, 1, 1, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 5, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 7, 7, 7, 7, 7

- (a) Geben Sie die Menge M der vorstellbaren Merkmalswerte und die Menge A der aufgetretenen Merkmalswerte an.
- (b) Erstellen Sie eine Tabelle mit den absoluten und relativen Häufigkeiten.
- (c) Erstellen Sie ein Stabdiagramm.
- (d) Stellen Sie die zugehörige empirische Verteilungsfunktion auf und zeichnen Sie diese.
- (e) Wie viele Autofahrer nutzen ihr Fahrzeug mehr als 2, aber höchstens 5 Tage in der Woche? Beschreiben Sie drei verschiedene Möglichkeiten, das Ergebnis zu erhalten.
- (f) Wie groß ist der Anteil der Autofahrer, die ihr Fahrzeug höchstens an 4 Tagen in der Woche nutzen?

Aufgabe 3

Zu einem erhobenen Merkmal sei die folgende empirische Verteilungsfunktion gegeben:

$$F(x) = \begin{cases} 0.00 & \text{für } x < 0 \\ 0.08 & \text{für } 0 \leq x < 2 \\ 0.20 & \text{für } 2 \leq x < 4 \\ 0.50 & \text{für } 4 \leq x < 6 \\ 0.68 & \text{für } 6 \leq x < 8 \\ 0.84 & \text{für } 8 \leq x < 10 \\ 0.92 & \text{für } 10 \leq x < 12 \\ 0.94 & \text{für } 12 \leq x < 14 \\ 0.98 & \text{für } 14 \leq x < 16 \\ 1.00 & \text{für } x \geq 16 \end{cases}$$

- Geben Sie die Menge A der Merkmalsausprägungen an.
- Erstellen Sie eine Tabelle der relativen Häufigkeiten.
- Wie viele Merkmalsträger müssen gemäß der vorliegenden empirischen Verteilungsfunktion F mindestens in der statistischen Masse enthalten gewesen sein?
- Ergänzen Sie mit der Zusatzinformation, dass die Urliste $n = 50$ Einträge umfasst, die Tabelle aus Teil (b) um die absoluten Häufigkeiten.

Aufgabe 4

Zu einem kardinalskalierten Merkmal sei die folgende (zur einfacheren Bearbeitung der Aufgabe bereits sortierte) Urliste der Länge $n = 50$ gegeben:

7.41, 8.47, 22.37, 23.94, 25.99, 34.60, 39.62, 39.88, 41.32, 42.61, 43.25, 43.82, 50.24, 51.93, 54.29, 54.58, 55.85, 55.87, 56.30, 57.21, 57.62, 58.50, 60.37, 60.41, 60.49, 61.37, 61.46, 63.31, 64.65, 64.91, 67.78, 68.19, 68.24, 69.63, 69.95, 72.26, 73.01, 73.48, 75.02, 75.60, 75.99, 76.13, 76.20, 77.52, 77.96, 78.16, 78.49, 78.85, 78.85, 79.87

- Führen Sie eine Klassierung der erhobenen Daten auf Grundlage der Klassen $K_1 = (0, 15]$, $K_2 = (15, 30]$, $K_3 = (30, 45]$, $K_4 = (45, 60]$, $K_5 = (60, 70]$, $K_6 = (70, 80]$ durch. Geben Sie insbesondere die jeweiligen Klassenbreiten, Klassenmitten, absoluten und relativen Klassenhäufigkeiten, Häufigkeitsdichten sowie die Werte der empirischen Verteilungsfunktion an den Klassengrenzen an.
- Zeichnen Sie das zugehörige Histogramm.
- Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.
- Bestimmen Sie (aus der Urliste) die Anzahl von Merkmalswerten zwischen 15 und 50. Welche Näherung für diese Anzahl können Sie aus der in Teil (c) aufgestellten empirischen Verteilungsfunktion berechnen?