

## Blatt 10

Abgabe bis Dienstag, 04. Juli 2023, 23:59 Uhr  
Jede komplett richtig gelöste Aufgabe ergibt 4 Punkte.

### Aufgaben

**37. Skalentypen.**

Angenommen, in einer Gruppe von Informatik-Studierenden werden (im Rahmen der Bevölkerungsforschung) verschiedene Merkmale erhoben. Geben Sie für die nachstehenden Merkmale jeweils den zugehörigen Skalentyp des Merkmals an und begründen Sie Ihre Antworten.

- (a) Anzahl Fachsemester
- (b) Bundesland des Erstwohnsitzes  
(bzw. Land des Erstwohnsitzes bei Erstwohnsitz im Ausland)
- (c) Durchschnittsnote bei den bislang bestandenen Modulprüfungen
- (d) verfügbares Monats-Nettoeinkommen (in Euro und Cent)

**38. Beschreibende Statistik.**

Bei der Messung der Körpergröße von 20 männlichen Schülern ergaben sich die folgenden Werte (in cm):

149 147 158 165 153 153 168 158 163 159  
177 175 163 170 162 162 170 153 147 157

- (a) Zeichnen Sie die empirische Verteilungsfunktion der angegebenen Messreihe (von Hand).
- (b) Zeichnen Sie ein Histogramm der angegebenen Messreihe mit der folgenden Klasseneinteilung:  $(145, 150]$ ,  $(150, 155]$ ,  $\dots$ ,  $(175, 180]$ .
- (c) Zeichnen Sie ein gleitendes Histogramm der angegebenen Messreihe.
- (d) Berechnen Sie die empirische Schiefe des Merkmals „Körpergröße von männlichen Schülern“ anhand der angegebenen Messreihe und interpretieren Sie Ihr Ergebnis anhand der erstellten Graphiken.

**39. Programmieraufgabe: Beschreibende Statistik.**

Der in dem Foliensatz zur deskriptiven Statistik besprochene Datensatz zum „Old Faithful“-Geysir ist in R mit dem Namen `faithful` enthalten.

Veranschaulichen Sie sich die univariaten Verteilungen der beiden Variablen „Eruptionsdauer“ und „Wartezeit bis zum Ausbruch“ anhand geeigneter Kennzahlen und Graphiken.

Hinweis:

Bitte schicken Sie Ihren R-Code auch elektronisch an den für Sie zuständigen Übungsgruppenleiter.

40. **Multiple Select-Aufgabe.**

Betrachten Sie die folgenden Aussagen über empirische Verteilungen und empirische Verteilungsfunktionen. Dazu seien  $Y_1, \dots, Y_n$  reellwertige, identisch verteilte Zufallsvariablen mit Verteilungsfunktion  $F$  von  $Y_1$  und mit empirischer Verteilungsfunktion  $\hat{F}_n$  von  $Y_1, \dots, Y_n$ .

- a) Falls  $Y_1, \dots, Y_n$  nicht stochastisch unabhängig sind, so kann es (mit nicht vernachlässigbarer Wahrscheinlichkeit) passieren, dass  $\hat{F}_n$  die wahre Verteilungsfunktion  $F$  selbst für großes  $n$  nicht präzise approximiert.
- b) Falls  $Y_1, \dots, Y_n$  nominalskaliert sind, so lassen sich  $F$  und  $\hat{F}_n$  nicht sinnvoll interpretieren.
- c) Falls  $Y_1, \dots, Y_n$  dichotom sind, so lässt sich  $\hat{F}_n$  nicht sinnvoll interpretieren.
- d) Falls  $Y_1, \dots, Y_n$  intervallskaliert sind, so lässt sich aus  $\hat{F}_n$  ein Histogramm der empirischen Verteilung von  $Y_1, \dots, Y_n$  ableiten.

Ermitteln Sie die richtige Kombination korrekter Aussagen.

Hinweise:

Das korrekte Ermitteln des Wahrheitsgehaltes der Aussagen ergibt jeweils einen Punkt. Um Raten nicht zu belohnen, werden nur begründete Antworten gewertet.