

Aufgaben zur Vorlesung „Werkzeuge der empirischen Forschung“

Aufg. 12) (Die Banknoten - Fortsetzung)

- d) (2 P.) Erzeugen Sie je ein vertikales Häufigkeitshistogramm für die Variable `laenge`, wobei echte und unechte Banknoten getrennt behandelt werden sollen. Die Histogramme sollen blau (`col="blue"`) und gestreift (`density=10`) sein.
- e) (3 P.) Erzeugen Sie folgende vertikale Histogramme für die Variablen `oben` und `unten`.
- e1) Beide Gruppen von Banknoten (echte und falsche) sollen in einem einzigen Histogramm zusammengefasst werden, aber mit unterschiedlicher Farbe und Schraffurdichte.
- e2) Jede Gruppe soll in ein eigenes Histogramm, beide Histogramme für eine Variable aber in einen Plot nebeneinander.
- e3) Die beiden Gruppen sollen getrennt ausgewertet werden (wie in Aufgabe 12d) und die einzelnen Balken zusätzlich unterschiedliche Muster haben.
- Hinweis:** Nutzen Sie Vektoren für `col` und `density` (siehe Beispiele `Descriptive_Statistik/Descr_Gchart...R`)
- f) (3 P.) Berechnen Sie für die Variablen `oben` und `unten` jeweils eine parametrische (Normalverteilung) und eine nichtparametrische Dichteschätzung, für beide Gruppen (echt und falsch) getrennt. Geben Sie bei der parametrischen Dichteschätzung die Schätzwerte an. Zeichnen Sie anschließend die Schätzungen mit `curve` bzw. `lines` mit verschiedenen Farben in einen Plot pro Variable.

Aufg. 13) (Toxaemia in der Schwangerschaft)

- a) (1 P.) Erzeugen Sie eindimensionale Häufigkeitstabellen für die Variablen `hyperten` und `proteinu`, jeweils nach den verschiedenen Ausprägungen der Variablen `social` und `smoking`.
- b) (2 P.) Erzeugen Sie eine gemeinsame Häufigkeitstabelle für die Variablen `hyperten` und `proteinu` und testen Sie diese auf Unabhängigkeit.
Hinweis: Für den Unabhängigkeitstest verwenden Sie die Funktion `chisq.test`.

Aufg. 14) (Über die Robustheit der Korrelationskoeffizienten, auf Papier)

Es seien jeweils die folgenden $n = 5$ Beobachtungspaare (X, Y) gegeben,

X	1	2	3	4	k
Y	1	2	3	4	5

- a) (3 P.) Bestimmen Sie die Pearson-, Spearman-, und Kendall-Korrelationskoeffizienten, jeweils für die Werte $X_n = k = 5, 10, 100, 1000$ mit der Hand. Hinweis: Überprüfen Sie Ihre Ergebnisse mit Hilfe von R.

- b) (1 P.) Zeigen Sie für den den Pearson-Korrelationskoeffizienten den Grenzwert

$$\lim_{k \rightarrow \infty} r_{XY} = \frac{1}{\sqrt{2}}$$

.

- c) (3 P.) Seien (X_i, Y_i) beliebig, sowie $X_{(n)} = k$ und $X_{(1)} = k'$. Berechnen Sie die Grenzwerte $\lim_{k \rightarrow \infty} r_{XY}$ und $\lim_{k' \rightarrow -\infty} r_{XY}$.
- d) (1 P.) Welche Grenzwerte gelten für die Spearman-, und Kendall-Korrelationskoeffizienten?
- e) (1 P.) Welche Schlussfolgerungen ziehen Sie für die Robustheit der drei Korrelationskoeffizienten?