1. Man teste auf dem Niveau $\alpha = 5\%$ die Hypothese, dass das Gewicht australischer Straußeneier normalverteilt ist mit Erwartungswert $\mu = 110$ (in Gramm), gegen die Alternative, dass dies nicht so ist. Hierzu verwende man die folgenden Daten. Man bestimme auch den $p$-Wert des Tests.

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $x_i$ | 106 | 110 | 100 | 103 | 109 | 101 | 97 | 103 | 111 | 99 |

Hinweis: Führen Sie die Rechnungen einmal mit dem Computer aus und einmal per Hand mit Hilfe eines Taschenrechners und einer Verteilungstabelle.

2. **(einzusenden)** Nur für die Normalverteilung ist das Konfidenzintervall

$$C(x) = \left[ M(x) - \frac{\sqrt{V^*(x)}}{\sqrt{n}} Q_{1-\frac{\alpha}{2}}(t_{n-1}),\ M(x) + \frac{\sqrt{V^*(x)}}{\sqrt{n}} Q_{1-\frac{\alpha}{2}}(t_{n-1}) \right] \qquad (*)$$

für den Erwartungswert ein gutes Konfidenzintervall zum Niveau $\alpha$. Für andere Verteilungen können zwei Probleme auftreten: Entweder sind die Intervalle unnötig groß, oder (schlimmer!) sie halten das Niveau $\alpha$ nicht ein. Eventuell tritt auch beides ein. Es soll hier durch ein Computerexperiment für drei Verteilungstypen geprüft werden, welcher Fall eintritt. Ferner soll das Ergebnis mit einem binomialverteilungsbasierten Konfidenzintervall für den Median verglichen werden.

Sei im Folgenden $n$ fest, und seien $X_1, \ldots, X_n$ unabhängige Zufallsvariablen mit einer der drei Verteilungen:

(i) $\mathcal{N}_{\mu,\sigma^2}$. $\mu \in \mathbb{R}$ und $\sigma^2$ unbekannt, $\mu$ ist zu schätzen.

(ii) $\exp_\theta$. $\theta > 0$ ist zu schätzen.

(iii) Cauchy mit Größenparameter 1 und Lageparameter $a \in \mathbb{R}$. D.h. die Dichte ist $\frac{1}{\pi}\frac{1}{1+(x-a)^2}$ für $x \in \mathbb{R}$.

(a) Für $\alpha = 5\%$ und $\alpha = 2\%$ sowie für $n = 5$, $n = 15$, $n = 30$, $n = 50$, $n = 200$ simuliere man die Zufallszahlen (für (i) mit $\mu = 1$ und $\sigma^2 = 5$, für (ii) mit $\theta = 2$, für (iii) mit $a = -3$) und stelle fest, ob das Konfidenzintervall $(*)$ den Median der Verteilung enthält. Durch hinreichend häufige Wiederholung ermittle man empirisch, ob das Konfidenzintervall das geforderte Niveau $\alpha$ einhält.

(b) Seien $q^-$ und $q^+$ das $\alpha/2$-Quantil und das $1 - \alpha/2$-Quantil von $b_{n,p}$. Ferner seien $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$ die nach Größe geordneten Werte von $x_1, \ldots, x_n$. Wie in (a) bestimme man empirisch das Niveau von $C'(x) = [x_{(q^-)}, x_{(q^+)}]$. Ferner vergleiche man die mittleren Längen der Intervalle $C(X)$ und $C'(X)$.

3. Die Abnutzung von Rädern an Eisenbahnwagen wird paarweise für jede Achse gemessen. Man teste anhand der folgenden Daten von zehn Achsen die Nullhypothese, dass linke und rechte Räder in gleicher Weise abnutzen, auf dem Niveau $\alpha = 5\%$.

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| links | 2.046 | 10.888 | 5.905 | 8.986 | 14.704 | 1.252 | 2.16 | 4.005 | 12.11 | 2.502 |
| rechts | 2.135 | 12.625 | 6.8 | 8.917 | 14.144 | 3.77 | 2.582 | 0.68 | 13.667 | 2.795 |

**Abgaben bis Montag, 05.02.2018, 14:00 Uhr an**
stoch-praktikum@mathematik.uni-mainz.de