# 8. Übung am 29. Mai 2017

**UV Angewandte Statistik (405.330), Ass.-Prof. Dr. Wolfgang Trutschnig**
Link Ankreuzliste: siehe www.trutschnig.net/courses
Mit 'F' versehene Aufgaben sind freiwillig, mit * versehene Aufgaben haben einen erhöhten Schwierigkeitsgrad.
Alle Verweise beziehen sich auf das Statistik-Skriptum.

**Übungsaufgabe 36 (R)** Überprüfen Sie mittels Simulationen in R die Güte (Überdeckungswahrscheinlichkeit) des in Satz 4.7 angegebenen, exakten Konfidenzintervalls $C_n$ für den Parameter $\theta = \mu_1 - \mu_2$ zweier normalverteilter Populationen $X \sim \mathcal{N}(\mu_1, \sigma^2)$, $Y \sim \mathcal{N}(\mu_2, \sigma^2)$. Berechnen Sie $C_n(x_1, \ldots, x_n, y_1, \ldots, y_m)$ für konkrete samples und führen Sie zusätzlich t.test(x,y,var.equal = TRUE) aus - was ist zu beobachten?

**Übungsaufgabe 37** Bezeichne $C_n = [L_n, U_n]$ das in Beispiel 4.11 hergeleitete exakte Konfidenzintervall für den Parameter $\theta$ von $\mathcal{U}(0, \theta)$. Beweisen Sie, dass $[\mathbb{P}_\theta]$ die folgenden Gleichheit gilt:

$$\lim_{n\to\infty} L_n = \theta = \lim_{n\to\infty} U_n$$

Wie lässt sich das Resultat interpretieren?
Hinweis: SLLN und Satz 2.16 anwenden.

**Übungsaufgabe 38** Gegeben sei eine Stichprobe $X_1, X_2, \ldots, X_n \sim F$; $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ bezeichne, wie gewohnt, die Ordnungsstatistiken. Gini's mittlere (absolute) Differenz $G$ ist definiert durch

$$G = \frac{1}{n(n-1)} \sum_{i,j=1}^{n} |X_i - X_j|.$$

Zeigen Sie die Existenz von Konstanten $a_1, \ldots, a_n$ mit

$$G = \sum_{l=1}^{n} a_l X_{(l)}.$$

Hinweis: Überlegen Sie sich die Darstellung zuerst für den Fall $n = 3, 4$ und schließen Sie dann auf den allgemeinen Fall.

**Übungsaufgabe 39 (Fortsetzung von 38)** Berechnen Sie für den Fall, dass die Stichprobe $X_1, X_2, \ldots, X_n$ von $X \sim \mathcal{N}(\mu, \sigma^2)$ kommt den Erwartungswert $\mathbb{E}(G)$, und überprüfen Sie Ihr Resultat mittels Simulationen in R.

**Übungsaufgabe 40** $X_1, \ldots, X_n$ sei eine Stichprobe von $X \sim \mathcal{N}(\mu, \sigma^2)$, $C_n$ bezeichne das in Satz 4.5 angegebene, exakte Konfidenzintervall für $\mu$. Ausgehend von der Stichprobe möchten wir überprüfen, ob $\mu = 0$ ist, und gehen wie folgt vor: Wenn $0 \notin C_n(x_1, \ldots, x_n)$, dann verwerfen wir die Hypothese, andernfalls nicht. Wie groß ist die Wahrscheinlichkeit dafür, die Hypothese zu verwerfen obwohl sie richtig ist?