Bernd Klaus (bernd.klaus@imise.uni-leipzig.de)
Verena Zuber (verena.zuber@imise.uni-leipzig.de)

http://uni-leipzig.de/~zuber/teaching/ws09/r-kurs/

1 Aufgabe: Testen an PISA-Daten

Das *Programme for International Student Assessment*, kurz PISA, ist eine standardisierte Bewertung von (15 jährigen) Schülern unter den teilnehmenden Staaten. Ziel der Regierungen ist, eine Datenbasis zur länderübergreifenden Forschung zu ermöglichen. Im Datensatz *PISA.csv* finden Sie die Ergebnisse einiger ausgwählter OECD-Staaten, getrennt nach dem Geschlecht (Variable *sex*: 1 Female, 2 Male, Perc_Sex gibt den Anteil an). Folgende Variablen sind von Interesse:

- *R00 - R06*: Mittlerer Score zur Lesekompetenz im Jahr 2000 bzw. 2006
- *M00 - M06*: Mittlerer Score zur Kompetenz in der Mathematik im Jahr 2000 bzw. 2006
- *S00 - S06*: Mittlerer Score in den Naturwissenschaften (*science*) im Jahr 2000 bzw. 2006

1. Lesen Sie den Datensatz *PISA.csv* ein.
2. Untersuchen Sie deskriptiv, ob sich die drei PISA-Scores des Jahres 2006 im Vergleich zum Jahr 2000 verändert haben. (Gehen Sie dabei und im weiteren nicht näher auf irgendwelche Geschlechtsunterschiede ein)
3. Untersuchen Sie mit dem geeigneten Test, ob sich die drei PISA-Scores signifikant verändert haben.

2 Aufgabe: Zu den Annahmen des t-Test

1. Ziehen Sie 100 exponential-verteilte Zufallszahlen mit dem Parameter $\lambda = 0.1$ und speichern Sie diese in dem Objekt `X1`. Erstellen Sie analog ein Objekt `X2` von 100 Zufallszahlen, für die Sie erneut 100 exponential-verteilte Zufallszahlen `HX2` mit dem Parameter $\lambda = 0.1$ ziehen und anschließend alle Elemente von 20 subtrahieren. Ein Element i des Vektors `X2` berechnet sich also als `20-HX2[i]`.
2. Zeichnen Sie für die beiden Objekte in einer Graphik je den Kerndichteschätzer.
3. Führen Sie den t-Test durch, um zu untersuchen, ob die beiden Objekte unterschiedliche Mittelwerte besitzen.
4. Führen Sie den Wilcoxen Rangsummen Test durch, um zu untersuchen, ob die beiden Objekte unterschiedliche Mediane besitzen.

3 Aufgabe: Der α-Fehler

1. Ziehen Sie 100 normalverteilte Zufallszahlen mit den Parametern $\mu = 0$ und $\sigma = 1$ und speichern Sie diese in dem Objekt `X1`. Erstellen Sie analog ein Objekt `X2` mit 100 (neuen) normalverteilten Zufallszahlen mit den Parametern $\mu = 0$ und $\sigma = 1$.
2. Führen Sie den t-Test durch, um zu untersuchen, ob die beiden Objekte unterschiedliche Mittelwerte besitzen.
3. Schreiben Sie eine Funktion, die
   - die Teilaufgaben (a) und (b) n-mal durchführt
   - in jedem Schleifenschritt soll nur die t-Statistik in einem Vektor abgespeichert werden
4. Testen Sie Ihre Funktion mit $n = 1000$. Plotten Sie die Verteilung der t-Statistiken mit einem Kerndichteschätzer. Wie viele t-Statistiken sind im Betrag fälschlicherweise größer als der kritische Wert von $\kappa = 1.972$?

4 Aufgabe: Testen der Füllmenge

Die Kontrolleure des Leipziger Amt für Verbraucherschutz sind alarmiert worden, dass auf dem Weihnachtsmarkt zu wenig Glühwein in den Bechern ausgeschenkt wird. Sie entschließen sich zum Selbstversuch. An einem Abend messen sie die folgenden Werte in ihren 12 Glühweingläsern:

1	2	3	4	5	6	7	8	9	10	11	12
0.240	0.249	0.251	0.252	0.244	0.241	0.247	0.261	0.225	0.212	0.264	0.201

1. Lesen Sie den Datensatz *Weihnachtsmarkt.csv* ein.
2. Untersuchen Sie mit dem geeigneten t-Test, ob signifikant weniger als $0.25l$ in den Bechern zu finden war!

Die Kursleiter wünschen ein schönes Fest!