

**Vorlesung: Programmieren in Anwendungen**  
**Übungsblatt 5**

[http://softech.informatik.uni-kl.de/Homepage/PIA\\_SS14](http://softech.informatik.uni-kl.de/Homepage/PIA_SS14)

## Zufallsvariablen und Wahrscheinlichkeitsverteilungen

1. Erstellen Sie Plots für die Dichtefunktion der Exponentialverteilung mit verschiedenen Parameterwerten in einer Grafik.
2. Erzeugen Sie Stichproben der Größe 10, 100, 1000, und 1000 für die Standardnormalverteilung. Berechnen Sie für den jeweiligen Stichprobenumfang Punktschätzer für den Mittelwert und die Varianz sowie Intervallschätzer für den Mittelwert. Wie verändern sich diese Schätzwerte bei steigender Stichprobengröße?

## Hypothesentests

### Arzneimittelstudie

Gegeben seien die folgenden Daten aus einer klinischen Studie zur Wirksamkeit zweier Medikamente.

	Medikament A	Medikament B
Behandlungserfolg	6	4
kein Behandlungserfolg	12	9

1. Erstellen Sie eine vollständige Kontingenztabelle für die Medikamentenstudie.
2. Verfassen Sie eine formale Hypothese, um zu testen, ob Medikament A besser wirkt als Medikament B.
3. Warum ist der  $\chi^2$ -Unabhängigkeitstest nicht anwendbar? Welche Alternativen gibt es?
4. Verwenden Sie den Fisher-Test, um die Hypothese zu testen. Zu welchem Ergebnis kommen Sie?

### Der PISA-Test

Das *Programme for International Student Assessment*, kurz PISA, ist eine im dreijährigen Turnus durchgeführte, standardisierte Bewertung von 15-jährigen Schülern der OECD Staaten. Auf der Homepage der Vorlesung finden Sie den Datensatz `PISA.csv` mit den Ergebnissen einiger ausgewählter Staaten, getrennt nach dem Geschlecht (Variable `Sex` kodiert mit 1=Female, 2=Male den Eintrag für Schülerinnen und Schüler, `Perc_SexXX` gibt den jeweiligen prozentualen Anteil an). Die weiteren Einträge sind:

**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

- Lesen Sie die Daten in R ein und verschaffen Sie sich einen Überblick.
- Wandeln Sie die Werte der Spalte `Sex` in geeignete Factors um.
- Untersuchen Sie mit einem geeigneten Test, ob sich die PISA-Bewertungen im Laufe der Jahre signifikant verändert haben.
  - Visualisieren Sie die Daten dazu mit einer geeigneten Graphik.
  - Stellen Sie nun zunächst passende Hypothesen auf.
  - Testen Sie Ihre Hypothesen mit einem geeigneten Test