

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

<http://uni-leipzig.de/~zuber/teaching/ws12/r-kurs/>

## 1 Aufgabe: Torwandschießen

Die Funktion `treff(rep)` (zu finden auf der Webseite der Veranstaltung) simuliert das Schießen eines ungeübten Spielers auf eine Torwand. Der Parameter `rep` steht dabei für die Anzahl der Wiederholungen.

- Wiederholen Sie das Experiment 100 mal und leiten Sie daraus eine Schätzung für die Trefferwahrscheinlichkeit des Schützen ab.
- Nun soll kein ungeübter Spieler, sondern ein Profi vom FC Barcelona auf die Torwand schießen. Wir gehen davon aus, dass sich die Trefferwahrscheinlichkeit verdoppelt. Ändern Sie die Funktion entsprechend ab.

## 2 Aufgabe: Das Geburtstagsparadoxon

Die Wahrscheinlichkeit des Ereignisses  $A_n$ , dass unter einer Anzahl von  $n$  Leuten mindestens 2 Leute am gleichen Tag Geburtstag haben, beträgt:

$$P(A_n) = 1 - \frac{\binom{365}{n}}{365^n} = 1 - \frac{365!}{(365-n)!365^n}$$

Zur praktischen Berechnung verwendet man eine Umformung mittels Logarithmus und Exponentialfunktion:

$$P(A_n) = 1 - \exp[\log(365!) - \log[(365-n)!] - n \cdot \log(365)]$$

Den Logarithmus der Fakultät erhält man in R per `lfactorial()`-Befehl, die Exponentialfunktion per `exp()`.

- Schreiben Sie eine R-Funktion, die die Wahrscheinlichkeit von  $A_n$  berechnet.
- Erstellen Sie eine Graphik, die diese Wahrscheinlichkeit von  $A_n$  in Abhängigkeit von  $n$  darstellt. Fügen Sie auch eine aussagekräftige Beschriftung hinzu.

HINWEIS: Sie können Ihrer Funktion mehrere  $n$ -Werte per Vektor übergeben. Dann erhalten Sie einen Vektor mit den jeweiligen Ergebnissen zurück. Dies ist z.B. bei der Erstellung des Graphen praktisch.

## 3 Aufgabe: Datensätze durchsuchen

- Schreiben Sie eine R-Funktion, die die Anzahl von Werten in einem Datensatz ermittelt, die in einem bestimmten offenen Intervall  $(a, b)$  liegen. Dabei sollten Sie auch auf fehlende Werte achten und Variablen ausschließen, die nicht numerisch sind (z.B. Faktoren oder Strings).
- Testen Sie die Methode anhand des *oecd*-Datensatzes.