

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

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

## 1 Aufgabe: Einfaktorielle ANOVA

Wir betrachten den Datensatz “*HotDog.csv*”. Er beschreibt den Salz- bzw. Kaloriengehalt von Hot Dogs verschiedener Marken in Abhängigkeit von der im Hotdog enthaltenen Fleischart.

- *Type*: Fleischart (“Beef”, “Meat”, “Poultry” )
  - *Calories*: Kalorien des Hot Dogs
  - *Sodium*: Salzgehalt des Hot Dogs
- (a) Lesen Sie den Datensatz mit Hilfe des `read.table()`-Befehls ein.
- (b) Untersuchen Sie mit zwei (einfaktoriellen) Varianzanalysen den Einfluss der Fleischart sowohl auf die Kalorien als auch auf den Salzgehalt.
- (c) Welches der beiden Anova-Modelle liefert einen signifikanten *F*-Wert?

## 2 Aufgabe: Mehrfaktorielle ANOVA

Wir betrachten den Datensatz “*Noise.csv*”. Er beschreibt die Auswirkung von drei Faktoren auf das Fahrgeräusch verschiedener Automodelle.

- *NOISE*: Geräuschpegel
  - *SIZE*: Größe des Autos: klein (1), mittel (2), groß (3)
  - *TYPE*: Art des Abgasfilters: (1) oder (2)
  - *SIDE*: Lage des Abgasfilters: links (1) oder rechts (2)
- (a) Lesen Sie den Datensatz mit Hilfe des `read.table()`-Befehls ein.
- (b) Erstellen Sie ein Anova-Modell mit 3 Faktoren. Welcher der 3 Haupteffekte scheint isoliert betrachtet der Wichtigste zu sein? Welcher der 3 Haupteffekte trägt am wenigsten zur Erklärung bei?
- (c) Erstellen Sie zwei weitere Anova-Modelle ohne den unwichtigsten Faktor, einmal mit und einmal ohne Interaktionseffekt. Ist die Interaktion signifikant?
- (d) Sollte das Modell aus der letzten Teilaufgabe noch weiter reduziert werden?

### 3 Aufgabe: Down-Syndrom

Trisomie 21, auch bekannt als Down-Syndrom, wird verursacht durch einen genetischen Defekt auf dem 21. Chromosom. Zur Erforschung dieser Krankheit wurde die DNA des 21. Chromosoms von Labormäusen verändert. Der Datensatz “*Mice.csv*” enthält verschiedene Daten über die Nachkommen dieser Mäuse.

- *tg*: Ist die Maus transgenetisch, d.h. ist die genetische Zusatzinformation beim Nachwuchs vorhanden, leidet dieser also unter dem Down-Syndrom (1) oder nicht (0)?
  - *sex*: Geschlecht der Maus, männlich (1) oder weiblich (0)
  - *age*: Alter der Maus
  - *weight*: Gewicht der Maus in Gramm
  - *cage*: Nummer des Käfigs in dem die Maus gelebt hat
- (a) Lesen Sie den Datensatz mit Hilfe des `read.table()`-Befehls ein und achten Sie darauf, dass **R** alle kategorialen Variablen als Faktoren behandelt.
- (b) Die Wissenschaftler vermuten, dass Mäuse mit Down-Syndrom deutlich dicker sind als ihre gesunden Artgenossen. Untersuchen Sie diese Hypothese zunächst mit geeigneten Graphiken. Berücksichtigen Sie dabei das Geschlecht.
- (c) Untersuchen Sie mit Hilfe von geeigneten ANOVA Modellen, ob diese Vermutung sich bestätigt. Achten Sie besonders darauf, dass es dabei in jeder Untergruppe genügend Beobachtungen gibt (Verwendung des Faktors *cage* ist nicht ratsam)!