Florian Wickelmaier
Forschungsmethoden und Mathematische Psychologie
Fachbereich Psychologie
Universität Tübingen

Hausaufgaben
Programming in R

Erzeugen Sie ein lauffähiges und kommentiertes R-Skript (Dateiendung `.R`), in dem Sie die folgenden Aufgaben bearbeiten.

a) (6 Pkt.) Erzeugen Sie in einer for-Schleife acht individuelle Dataframes mit jeweils folgender Struktur:

```
> str(df.ind)

'data.frame':        5 obs. of  5 variables:
 $ id : num  3 3 3 3 3
 $ x1 : int  1 0 1 1 0
 $ x2 : int  5 2 13 8 2
 $ rt1: num  601 535 508 621 519
 $ rt2: num  669 655 684 619 629
```

Die Variable `id` gibt die Vp-Nummer (1–8) an, `x1` und `x2` sind zufällig aus Poissonverteilungen mit $\lambda_1 = 2$ und $\lambda_2 = 7$ gezogen, `rt1` und `rt2` aus Normalverteilungen mit $\mu_1 = 620$, $\mu_2 = 690$ und $\sigma = 70$. Schreiben Sie die Dataframes als Textdateien (`subj1.txt`, ..., `subj8.txt`) in den Ordner `results`. Hinweis: Dieser Ordner muss *nicht* mit einem R-Befehl, sondern kann auch in der Shell oder im Dateimanager erzeugt werden.

Lesen Sie in einer zweiten Schleife (oder mit `lapply()`) die acht Dataframes wieder ein und kombinieren Sie diese zu einem Gesamt-Dataframe.

b) (5 Pkt.) Schreiben Sie eine Funktion, die ein statistisches Verfahren implementiert. Dieses soll *entweder*

- eine einfache lineare Regression sein, $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$; Ihre Funktion sollte den Parametervektor $\boldsymbol{\beta} = (\beta_0, \beta_1)'$ schätzen, indem Sie z. B. $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ verwenden oder die Schätzer aus der Formelsammlung; *oder*
- ein Unabhängigkeitstest für eine Vierfeldertafel, der sowohl die Pearson-χ^2-Statistik

$$X^2 = \sum_{cells} \frac{(x_{ij} - \hat{\mu}_{ij})^2}{\hat{\mu}_{ij}}$$

als auch die Likelihood-Verhältnis-Statistik

$$G^2 = 2 \sum_{cells} x_{ij} \log \frac{x_{ij}}{\hat{\mu}_{ij}}$$

berechnet. Dabei sind x_{ij} jeweils die beobachteten Häufigkeiten und $\hat{\mu}_{ij} = x_{i+}x_{+j}/x_{++}$ die erwarteten Häufigkeiten.

Weisen Sie dem Rückgabewert Ihrer Funktion eine neue Klasse zu. Schreiben Sie eine Printmethode für diese Klasse (für Hinweise siehe `stats:::print.lm` bzw. `stats:::print.htest`). Rufen Sie zum Test die Funktion mit Beispieldaten Ihrer Wahl auf.