

# Functions and Control Structures

## Übungsaufgabe 4\*

Gegeben seien zwei numerische Variablen  $x_1$  und  $x_2$ , sowie eine Faktorvariable  $g$ .

- a) Schreiben Sie eine Funktion 'stats\_group', die das arithmetische Mittel und die Varianz der Variablen  $x_1$  und  $x_2$  sowie die Korrelation und die Kovarianz zwischen den Variablen  $x_1$  und  $x_2$  berechnet. Diese sechs Kennzahlen sollen jeweils für die gesamten Beobachtungen und zusätzlich getrennt nach den Kategorien, die durch den Faktor  $g$  definiert sind, berechnet werden. Die Grundstruktur der Funktion soll folgendermaßen aussehen:

```
stats_group <- function( ..... ){  
  .....  
}
```

Die Funktion soll außerdem folgende Spezifikationen erfüllen:

- Eingabe der Funktion sind die Variablen  $x_1$ ,  $x_2$  und  $g$ , sowie weitere, erforderliche Parameter.
- Prüfen Sie die Eingaben für  $x_1$ ,  $x_2$  und  $g$  auf Richtigkeit. Die Funktion soll abbrechen, falls mindestens eine der Eingaben nicht vom richtigen Typ ist. Außerdem soll die Funktion nur ein Ergebnis liefern, falls alle drei Vektoren gleich lang sind.
- Die Funktion soll so flexibel gestaltet sein, dass die Kovarianz bzw. Korrelationen nach Pearson, Kendall oder Spearman berechnet werden kann.
- Die Funktion soll so flexibel gestaltet sein, dass beim Aufruf der Funktion global entschieden werden kann, wie an den relevanten Stellen im Funktionskörper mit fehlenden Werten umgegangen wird.
- Die Funktion soll überprüfen, ob die Korrelationen zwischen  $x_1$  und  $x_2$ , getrennt nach den Kategorien, alle dieselbe Richtung aufweisen, wie die Korrelation auf den gesamten Beobachtungen, oder nicht. In beiden Fällen soll die Funktion eine geeignete Ausgabe in der Konsole anzeigen.

Rückgabe der Funktion soll eine Liste sein mit folgenden Elementen:

- Anzahl an Beobachtungen (Skalar)
- Anzahl an Kategorien des Faktors  $g$  (Skalar)
- Anzahl an Beobachtungen je Kategorie (Vektor)
- Berechnete Kennzahlen für alle Beobachtungen (Vektor)
- Berechnete Kennzahlen getrennt nach Kategorien (Matrix)

Achten Sie bei allen Elementen auf sinnvolle Bezeichnungen.

- b) Testen Sie Ihre Funktion mit geeigneten Aufrufen, so dass die gesamte Funktionalität Ihrer Funktion überprüft wird. Beispiele für mögliche Eingaben (inkl. Kombinationen) sind:

- **Für  $x_1$ :**

```
x1 <- rnorm(30, mean=0, sd=1)
```

```
x1 <- c(1:29, NA)
```

```
x1 <- c(letters[1:15], letters[1:15])  
x1 <- data.frame(var1=1:30, var2=30:1)  
x1 <- rnorm(29, mean=0, sd=1)
```

- **Für x2:**

```
x2 <- rnorm(30, mean=0, sd=1)  
x2 <- c(NA,1:28,NA)  
x2 <- c(letters[1:10], letters[1:10], letters[1:10])  
x2 <- data.frame(var1=1:30, var2=30:1)  
x2 <- rnorm(25, mean=0, sd=1)
```

- **Für g:**

```
g <- factor(rep((1:3), times=10), levels=1:3, labels=c("a", "b", "c"))  
g <- c(rep(1,10), rep(2,10), rep(3,10))  
g <- factor(c(rep(1,9), rep(2,8), rep(3,9)), levels=1:3)  
g <- factor(sample(1:3, 30, replace=TRUE), levels=1:3)
```

*\* Durch Lösen dieser Übungsaufgabe kann ein Teil der Prüfungsleistung erbracht werden.*

**Abgabetermin: 07.06.2016/08.06.2016**