

**Vorbereitungen:**

Laden Sie den Datensatz `orthodont` von der Homepage herunter und lesen Sie ihn in R ein.

**Aufgabe 1:** Die `reshape`-Funktion

- a) Der eingelesene Datensatz befindet sich im `long`-Format. Nutzen Sie die `reshape`-Funktion, um den Datensatz in das `wide`-Format umzuformatieren.
- b) Bringen Sie den Datensatz aus dem `wide`-Format mit Hilfe der `reshape`-Funktion nun wieder zurück in das `long`-Format.

**Aufgabe 2:** Das `groupedData`-Format

- a) Legen Sie eine neue Spalte `InBehandlung` in `orthodont` an. Weisen Sie der Spalte per `sample`-Funktion zufällig die Werte 0 (Subjekt befindet sich aktuell nicht in Behandlung) oder 1 (aktuell in Behandlung) zu. Pro Person muss die Variable dabei nicht zeitkonstant sein!
- b) Formatieren Sie den `orthodont`-Datensatz zu einem neuen Datensatz `orthodont2` im `groupedData`-Format um. Übergeben Sie die Variablen `Sex` und `InBehandlung` jeweils entweder an das `outer`- oder an das `inner`-Argument der Funktion.  
Hinweis: Die Funktion `groupedData` ist Teil des `nlme`-Pakets.
- c) Die Formatierung von Datensätzen im `groupedData`-Format erleichtert u.a. die Erstellung von Plots.  
Hinweis: `?plot.nfnGroupedData`
  - i) Wenden Sie die `plot`-Funktion auf `orthodont2` an. Vergleichen Sie das Ergebnis mit dem Ergebnis, wenn man `plot` auf `orthodont` anwendet.
  - ii) Plotten Sie die Verläufe der Subjekte getrennt nach Geschlechtern.
  - iii) Plotten Sie die Verläufe der Subjekte unter Darstellung des Behandlungsstatus.

### Aufgabe 3: `lme`, `lmer`, `gam` und `gamm`

In dieser Aufgabe soll ein Modell auf den `orthodont`-Daten geschätzt werden. Die gemessenen Distanzen werden dabei durch ein lineares Modell erklärt, wobei als Einflussgrößen das Alter, das Geschlecht sowie eine Interaktion zwischen den beiden Variablen einbezogen werden. Darüber hinaus wird ein Random Intercept pro Subjekt sowie ein Random Slope für das Alter einbezogen.

- a) Notieren Sie das Modell auf Beobachtungsebene. Geben Sie auch alle Modellannahmen an.
- b) Schätzen Sie das Modell unter Benutzung der angegebenen Funktionen und sehen Sie sich jeweils den `summary` der Modelle an. Kommen die Modelle zu vergleichbaren Ergebnissen?
  - i) Verwenden Sie die `lme`-Funktion aus dem `nlme`-Paket
  - ii) Verwenden Sie die `lmer`-Funktion aus dem `lme4`-Paket
  - iii) Verwenden Sie die `gam`-Funktion aus dem `mgcv`-Paket
  - iv) Verwenden Sie die `gamm`-Funktion aus dem `mgcv`-Paket
- c) Nützliche Funktionen bei der Arbeit mit `lme`-Modellen sind `getVarCov` sowie `cov2cor` aus dem `nlme`-Paket. Sehen Sie sich in der Hilfe zu `getVarCov` die Argumente der Funktion an und testen Sie die Optionen für das Argument `type` an dem in b) gefitteten Modell.
  - i) Warum sind die konditionalen Kovarianzmatrizen diagonal?
  - ii) Berechnen Sie die marginale Korrelationsmatrix. Wie sind die geschätzten Korrelationen zu interpretieren?
  - iii) Berechnen Sie die konditionalen und marginalen Kovarianzen für verschiedene Subjekte. Wieso erhalten Sie immer das gleiche Ergebnis?
- d) Extrahieren Sie jeweils die angegebenen Parameter aus den in a) erzeugten Modellen. Beachten Sie, dass die Modell-Objekte pro Paket unterschiedlich implementiert sind und deshalb meist unterschiedliche Funktionen verwendet werden müssen. Kommen die Modelle zu vergleichbaren Ergebnissen?

Hinweis: Hilfreiche Funktionen sind u.a. `ranef`, `VarCorr`, `getVarCov`, `cov2cor`, `gam.vcomp`

- i) `lme`: Extrahieren Sie die geschätzten Random Effects, die Kovarianzmatrix der Random Effects sowie die geschätzte Residualvarianz. Erzeugen Sie aus der Kovarianzmatrix eine Korrelationsmatrix der Random Effects.
- ii) `lmer`: Extrahieren Sie die geschätzten Random Effects, die Varianzen der Random Effects, deren Korrelation sowie die geschätzte Residualvarianz.
- iii) `gam`: Extrahieren Sie die geschätzten Random Effects, die geschätzten Varianzen der Random Effects sowie die geschätzte Residualvarianz. Wieso ist eine Extrahierung der Kovarianzen der Random Effects hier nicht notwendig?
- iv) `gamm`: Extrahieren Sie die geschätzten Random Effects, die geschätzten Varianzen der Random Effects sowie die geschätzte Residualvarianz.