

Dieses Aufgabenblatt soll Sie in die Schätzung Linearer Gemischter Modelle (Linear Mixed Models) einführen. Die zu bearbeitenden Aufgaben beziehen sich auf die Inhalte der dritten und vierten Vorlesungsfolien.

**Aufgabe 1:**

In dieser Aufgabe beschäftigen wir uns erneut mit dem Datensatz `rats` (siehe Blatt 1, Aufg. 1). Verwenden Sie den auf der Homepage verfügbaren Code, um `rats` in R einzulesen, sowie für die anschließende Analyse vorzubereiten.

- a) Fitten Sie ein lineares Modell basierend auf allen Daten mit einem linearen Trend in `logT` und einem festen Effekt für jede Ratte. Welche Nachteile sehen Sie in diesem Vorgehen?
- b) Berechnen Sie nun mithilfe der Funktion `lme()` ein lineares gemischtes Modell mit linearem Trend in `logT` und subjekt-spezifischen zufälligen Intercepts.
  - (i) Formulieren Sie das zugrundeliegende Modell für den Responsevektor  $\mathbf{Y}_i$  der  $i$ -ten Ratte und geben Sie die Dimensionen aller Komponenten an.
  - (ii) Wie groß ist die geschätzte marginale Korrelation zwischen zwei Messungen an derselben Ratte? *Hinweis:* Betrachten Sie die Folien 31 – 38 aus dem dritten Foliensatz. Die Funktion `getVarCov()` könnte behilflich sein.
  - (iii) Wie groß ist die geschätzte bedingte Korrelation von zwei Messungen an der gleichen Ratte?
  - (iv) Wie groß ist die Korrelation zwischen zwei Messungen zum gleichen Zeitpunkt bei verschiedenen Ratten?
- c) Um grafisch zu überprüfen, ob auch subjekt-spezifische Trends sinnvoll wären, schätzen Sie mit der Funktion `lmList()` für jede Ratte mit mindestens 3 Messungen ein separates lineares Modell und plotten Sie mithilfe der Funktion `plot(intervals())` die Schätzer und Konfidenzintervalle für den Intercept und für `logT`. *Hinweis:* Der Datensatz muss als `groupedData`-Objekt mit Gruppierungsvariable `SUBJECT` vorliegen.
  - i) Wieso betrachten wir hier nur Ratten mit mindestens 3 Messungen?
  - ii) Wieso sind separate lineare Modelle für die einzelnen Ratten nur für eine solche Veranschaulichung geeignet?
  - iii) Da sich auch eine recht große Streuung in den Schätzern für `logT` erkennen lässt, fitten Sie nun ein lineares gemischtes Modell mit linearem Trend in `logT` und subjekt-spezifischen zufälligen Intercepts und Trends. Verwenden Sie für bessere Vergleich-

barkeit auch nur die Ratten mit mindestens 3 Messungen.

- (iv) Bestimmen Sie die geschätzte Kovarianzmatrix  $\hat{\mathbf{D}}$  der zufälligen Effekte. Wie groß ist die geschätzte Korrelation zwischen den zufälligen Intercepts und Slopes?
- (v) Wie könnte man das Modell hinsichtlich dieser Kovarianzstruktur vereinfachen? Schätzen Sie das entsprechende Modell.
- (vi) Vergleichen Sie nun die Parameterschätzungen und die gefitteten Werte der subjekt-spezifischen linearen Modelle und des linearen gemischten Modells aus v) mithilfe von `plot(compareFits())` und `plot(comparePred())`. Beschreiben Sie, was Ihnen auffällt und benennen und erklären Sie den Effekt.

d) Betrachten Sie nun das folgende Modell:

$$\begin{aligned} \text{RESPONSE}_{ij} = & \beta_0 + \beta_1 \log T_{ij} + \beta_2 \text{GROUP1}_i + \beta_3 \text{GROUP2}_i + \beta_4 \text{GROUP1}_i \cdot \log T_{ij} \\ & + \beta_5 \text{GROUP2}_i \cdot \log T_{ij} + b_{0i} + \varepsilon_{ij}. \end{aligned}$$

Kann man davon ausgehen, dass die random effects assumption hier erfüllt ist? Begründen Sie Ihre Antwort.