

Aufgabe 1: Modellwahl

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

Fitten Sie ein LMM, in welchem Sie die `zufriedenheit` durch lineare Effekte von `Jahr` und `Einkommen`, sowie einem Random Intercept pro `Subjekt` erklären. Verwenden Sie das Modell als Ausgangsmodell für eine in dieser Aufgabe durchzuführende Modellselektion.

- a) Notieren Sie sich die allgemeine Formel der marginalen sowie der konditionalen logarithmierten ML-Likelihood eines Linearen Gemischten Modells in Matrixnotation.
- b) Welche Arten von Modellen sind anhand des `mAIC` vergleichbar? Welche anhand des `cAIC`? Gibt es für die Anwendung des `mAICs/cAICs` Voraussetzungen?
- c) Beantworten Sie die folgenden Fragen mithilfe einer AIC-Modellselektion:

Hinweis I: Mithilfe der im Internet bereitgestellten Funktion `compute_cAIC` können Sie das konditionale AIC für `lme`-Modelle berechnen.

Hinweis II: Verwenden Sie bei der Schätzung des `lme`-Modells die Einstellung „`control = lmeControl(opt = 'optim')`“, um mit der Verwendung des numerisch stabileren Optimierungsalgorithmus `optim()` dem vorzubeugen, dass es in einer der Bootstrap-Iterationen zur Berechnung der `cAIC`-Freiheitsgrade zu Konvergenzproblemen kommt.

- i) Ist die Aufnahme eines quadratischen `Jahr`-Effektes sinnvoll?
- ii) Ist die Aufnahme eines Random Slopes für `Einkommen` sinnvoll?

Aufgabe 2: GLMMs und Marginale Modelle

Vorbereitung: Laden Sie den Datensatz `epil` aus dem `MASS`-Paket mittels dem Befehl `data(epil)` in R. Im Datensatz sind die Anzahlen an epileptischen Anfällen innerhalb von 2-Wochen-Intervallen für 59 Epileptiker über die Zeit gesammelt. Die Probanden sind in eine Behandlungsgruppe und eine Placebogruppe aufgeteilt.

Ein Mediziner kommt auf Sie zu und beauftragt Sie zu untersuchen, ob das `Treatment` (Variable `trt`) die Wahrscheinlichkeit beeinflusst, dass eine Person mehr als drei Anfälle in 14 Tagen hat. Nehmen Sie als Kontrollvariablen die Variablen `lbase` (logarithmierte Anzahl Anfälle in 8-wöchiger Baseline-Periode) und `age` (Alter) sowie einen Random Intercept pro `subject` auf.

- a) Fitten Sie mithilfe der Funktion `glmmPQL` aus dem `MASS`-Paket ein GLMM, um die Fragestellung zu untersuchen. Interpretieren Sie den `Treatment`- und den `Alter`effekt!

Hinweis: Die binäre Zielgröße müssen Sie aus der Anzahl an Anfällen (Variable `y`) neu erstellen.

- b) Fitten Sie mithilfe der Funktion `gee` aus dem gleichnamigen Paket ein Marginales Modell, um die Fragestellung zu untersuchen. Nehmen Sie dafür eine unstrukturierte Korrelationsstruktur an. Interpretieren Sie den Treatment- und den Alterseffekt!