

Aufgabe 1

In dieser Aufgabe sollen die Ergebnisse der verschiedenen Schätzmethoden für GLMMs verglichen werden.

(a) Simulieren Sie einen Datensatz in R für folgendes Modell:

$$y_{ij}|x_{ij}, b_i \sim \text{Ber}(\pi_{ij}), \text{ unabhängig verteilt}$$

$$\pi_{ij} = P(y_{ij} = 1|x_{ij}, b_i) = g^{-1}(x_{ij}\beta + b_i), \quad i = 1, \dots, m; \quad j = 1, \dots, n_i$$

mit

- $g^{-1}(\eta) = \frac{\exp(\eta)}{1+\exp(\eta)}$
- $x_{ij} \stackrel{\text{iid}}{\sim} U(-1, 1)$, d.h. x_{ij} ist gleichverteilt auf dem Intervall $[-1, 1]$
- $\beta = 2$
- $b_i \stackrel{\text{iid}}{\sim} N(0, 1)$
- $m = 50$ Subjekte
- $n_i = 5$ Beobachtungen pro Subjekt

Dabei soll der Datensatz so erzeugt werden, dass Sie Parameter, wie Anzahl der Subjekte oder Größe von β , einfach verändern können.

(b) Schätzen Sie das Modell auf den simulierten Daten mit den 3 Schätzmethoden, die Sie aus der Vorlesung kennen, also mit

- (i) Penalized Quasi-Likelihood (PQL), Funktion `glmmPQL()` aus R-Paket `MASS`
- (ii) Laplace-Approximation, Funktion `glmer()` aus R-Paket `lme4`
- (iii) (Adaptive) Gaußsche Quadratur, Funktion `glmer()` aus R-Paket `lme4`

Vergleichen Sie die Ergebnisse aus den verschiedenen Schätzmethoden, insbesondere die Schätzungen der festen Effekte, sowie die Schätzung der Varianz der zufälligen Effekte.

(c) Erhöhen Sie den Stichprobenumfang und/oder die Zahl der Beobachtungen pro Subjekt. Schätzen Sie das Modell erneut mit den verschiedenen Methoden und betrachten Sie, wie sich die Schätzungen verändern.