

# Inhalt

- 1 Das lineare gemischte Modell
- 2 Likelihood-Schätzung für lineare gemischte Modelle
- 3 Likelihood-Inferenz im linearen gemischten Modell
- 4 Bayes-Schätzung für lineare gemischte Modelle
- 5 Additive gemischte Modelle
- 6 Das generalisierte lineare gemischte Modell
- 7 Likelihood-Schätzung für generalisierte lineare gemischte Modelle
  - Laplace-Approximation und P-IRLS
  - Adaptive Gauss-Hermite Quadratur (AGQ)
  - Penalized Quasi-Likelihood (PQL)
  - Inferenz in GLMMs

# Likelihood eines GLMM

Verteilungsannahme für  $\mathbf{y}|\mathbf{b}$ :  $y_i|\mathbf{b} \sim \text{Expo.fam.}(\boldsymbol{\theta}, \phi)$  unabhängig

Verteilungsannahme für  $\mathbf{b}|\boldsymbol{\vartheta}$ :  $\mathbf{b}|\boldsymbol{\vartheta} \sim N(\mathbf{0}, \mathbf{G}(\boldsymbol{\vartheta}))$

⇒ gemeinsame / penalisierte Likelihood:

$$L(\boldsymbol{\beta}, \phi, \boldsymbol{\vartheta}, \mathbf{b}) = f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \phi, \boldsymbol{\vartheta}) = \left( \prod_{i=1}^n f(y_i|\boldsymbol{\beta}, \phi, \mathbf{b}, \boldsymbol{\vartheta}) \right) f(\mathbf{b}|\boldsymbol{\vartheta})$$

bzw. marginale Likelihood:

$$L(\boldsymbol{\beta}, \phi, \boldsymbol{\vartheta}) = f(\mathbf{y}|\boldsymbol{\beta}, \phi, \boldsymbol{\vartheta}) = \int \left( \prod_{i=1}^n f(y_i|\boldsymbol{\beta}, \phi, \mathbf{b}, \boldsymbol{\vartheta}) \right) f(\mathbf{b}|\boldsymbol{\vartheta}) d\mathbf{b}.$$

Erinnerung: im LMM ist das Integral analytisch lösbar.

$\int f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \phi, \boldsymbol{\vartheta}) d\mathbf{b}$  ist die Dichte einer  $N(\mathbf{X}\boldsymbol{\beta}, \mathbf{ZG}(\boldsymbol{\vartheta})\mathbf{Z}' + \mathbf{R}(\phi, \boldsymbol{\vartheta}))$ -Verteilung.

Im GLMM: ??

# ML-Inferenz

Interessante Parameter: primär  $\beta$ ,  $\vartheta$ , evtl.  $\phi$   
 $\Rightarrow$  Problem: finde  $\operatorname{argmax} L(\beta, \vartheta, \phi)$ , wobei

$$L(\beta, \vartheta, \phi) = \int \left( \prod_{i=1}^n f(y_i | \beta, \phi, \mathbf{b}, \vartheta) \right) f(\mathbf{b} | \vartheta) d\mathbf{b}$$

$$= \int \left( \prod_{i=1}^n \exp \left( \frac{y_i \theta_i - b(\theta_i)}{\phi} - c(y_i, \phi) \right) \right) |\mathbf{G}(\vartheta)|^{-1/2} \exp \left( -\frac{1}{2} \mathbf{b}' \mathbf{G}(\vartheta)^{-1} \mathbf{b} \right) d\mathbf{b}$$

mit  $\theta_i = (b')^{-1}(g^{-1}(\eta_i))$ ,  $\eta_i = \mathbf{x}'_i \beta + \mathbf{z}'_i \mathbf{b}$ .

Hoch-dimensionales Integral, i.A. nicht analytisch lösbar  
 $\Rightarrow$  iterative Optimierung einer Approximation der Likelihood

# Laplace-Approximation

**Problem:** Löse  $r$ -dimensionales Integral  $H = \int \exp(h(\boldsymbol{\theta})) d\boldsymbol{\theta}$

**Ansatz:**

- bestimme  $\boldsymbol{\theta}_0 = \arg \max h(\boldsymbol{\theta})$

- quadratische Taylor-Entwicklung von  $h(\boldsymbol{\theta})$  um  $\boldsymbol{\theta}_0$ :

$$h(\boldsymbol{\theta}) \approx h(\boldsymbol{\theta}_0) + \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\theta}_0)' \underbrace{\left( \frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} h(\boldsymbol{\theta}_0) \right)}_{=-\mathbf{P}} (\boldsymbol{\theta} - \boldsymbol{\theta}_0)$$

- $\Rightarrow \int \exp(h(\boldsymbol{\theta})) d\boldsymbol{\theta} \approx \int \exp\left(h(\boldsymbol{\theta}_0) - \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\theta}_0)' \mathbf{P} (\boldsymbol{\theta} - \boldsymbol{\theta}_0)\right) d\boldsymbol{\theta}$   
wie bei  $N(\boldsymbol{\theta}_0, \mathbf{P}^{-1})$

- $\Rightarrow H \approx \exp(h(\boldsymbol{\theta}_0)) \underbrace{\left( (2\pi)^{r/2} |\mathbf{P}|^{-1/2} \right)}_{1/\text{Normierung der } N(\boldsymbol{\theta}_0, \mathbf{P}^{-1})}$

# GLMM-Likelihood: Laplace-Approximation

Verwende eine Laplace-Approximation, mit Entwicklung um den Maximierer  $\hat{\mathbf{b}}$  von  $L(\boldsymbol{\beta}, \phi, \boldsymbol{\vartheta}, \mathbf{b}) = f(\mathbf{y}|\boldsymbol{\beta}, \phi, \mathbf{b})f(\mathbf{b}|\boldsymbol{\vartheta})$ ,

$$\begin{aligned}
 \log(L(\boldsymbol{\beta}, \phi, \boldsymbol{\vartheta})) &= \log \left( \int f(\mathbf{y}|\boldsymbol{\beta}, \phi, \mathbf{b})f(\mathbf{b}|\boldsymbol{\vartheta})d\mathbf{b} \right) \\
 &= \log \left( \int L(\boldsymbol{\beta}, \phi, \mathbf{b})(2\pi)^{-r/2} |\mathbf{G}(\boldsymbol{\vartheta})|^{-\frac{1}{2}} \exp \left( -\frac{1}{2} \mathbf{b}' \mathbf{G}(\boldsymbol{\vartheta})^{-1} \mathbf{b} \right) d\mathbf{b} \right) \\
 &\approx \log(L(\boldsymbol{\beta}, \hat{\mathbf{b}}, \phi)) - \frac{r}{2} \log(2\pi) - \frac{1}{2} \log |\mathbf{G}(\boldsymbol{\vartheta})| - \frac{1}{2} \hat{\mathbf{b}}' \mathbf{G}(\boldsymbol{\vartheta})^{-1} \hat{\mathbf{b}} \\
 &\quad + \log \left( \int \exp \left( -\frac{1}{2} (\mathbf{b} - \hat{\mathbf{b}})' \mathcal{I}(\hat{\mathbf{b}}) (\mathbf{b} - \hat{\mathbf{b}}) \right) d\mathbf{b} \right) \\
 &= l(\boldsymbol{\beta}, \hat{\mathbf{b}}, \phi) - \frac{1}{2} \log |\mathcal{I}(\hat{\mathbf{b}})| - \frac{1}{2} \log |\mathbf{G}(\boldsymbol{\vartheta})| - \frac{1}{2} \hat{\mathbf{b}}' \mathbf{G}(\boldsymbol{\vartheta})^{-1} \hat{\mathbf{b}}.
 \end{aligned}$$

Dabei ist  $\mathcal{I}(\mathbf{b}) = -E \left( \frac{\partial^2}{\partial \mathbf{b} \partial \mathbf{b}'} l(\boldsymbol{\beta}, \phi, \boldsymbol{\vartheta}, \mathbf{b}) \right)$  (mit zusätzlichem Erwartungswert), Herleitung siehe S. 166.

# Schaukel-Algorithmus

In der Laplace-Approximation wird der Maximierer  $\hat{\mathbf{b}}$  von  $L(\boldsymbol{\beta}, \phi, \boldsymbol{\vartheta}, \mathbf{b})$  benötigt.  
⇒ iterativer, zweistufiger Schaukel-Algorithmus:

- 1 Für gegebene  $\boldsymbol{\beta}, \phi, \boldsymbol{\vartheta}$  bestimme  $\hat{\mathbf{b}} = \arg \max_{\mathbf{b}} L(\boldsymbol{\beta}, \phi, \boldsymbol{\vartheta}, \mathbf{b})$  über penalisierten IRLS-Algorithmus (P-IRLS).
- 2 Maximiere Laplace-Approximation  $\tilde{L}(\boldsymbol{\beta}, \phi, \boldsymbol{\vartheta})$  von  $L(\boldsymbol{\beta}, \phi, \boldsymbol{\vartheta})$  in  $\hat{\mathbf{b}}$  mit numerischer Optimierung (Pseudo-Newton-Algorithmen wie BFGS).

Iteriere bis zur Konvergenz der Devianz  $-2 \log L(\boldsymbol{\beta}, \phi, \boldsymbol{\vartheta}, \mathbf{b})$ .

Dieser Schaukel-Algorithmus ist im R-Paket [lme4](#) implementiert.

# Grundidee: IRLS-Algorithmus

- IRLS = Fisher-Scoring für GLM
  - IRLS: Iteratively **R**e-Weighted **L**east **S**quares
  - Führt GLM-Schätzproblem auf iterierte gewichtete KQ-Schätzung zurück.
- Fisher-Scoring zur Lösung des Score-Gleichungssystems  $s(\boldsymbol{\theta}) \stackrel{!}{=} \mathbf{0}$ 
  - lineare Taylor-Entwicklung  $s(\boldsymbol{\theta}) \approx s(\boldsymbol{\theta}_0) - \mathcal{J}(\boldsymbol{\theta}_0)(\boldsymbol{\theta} - \boldsymbol{\theta}_0) \stackrel{!}{=} \mathbf{0}$  mit der beobachteten Informationsmatrix  $\mathcal{J}(\boldsymbol{\theta})$ .
  - Ersetzen von  $\mathcal{J}(\boldsymbol{\theta})$  (Newton-Raphson) durch die erwartete Informationsmatrix  $\mathcal{I}(\boldsymbol{\theta})$  (Fisher-Scoring) liefert

$$\mathcal{I}(\boldsymbol{\theta}_0)\boldsymbol{\theta} = \mathcal{I}(\boldsymbol{\theta}_0)\boldsymbol{\theta}_0 + s(\boldsymbol{\theta}_0).$$

# $s(\mathbf{b})$ und $\mathcal{I}(\mathbf{b})$ für den kanonischen Link

Für den kanonischen Link ist  $\boldsymbol{\theta} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b}$ .

$$\begin{aligned}\Rightarrow s(\mathbf{b}) &= \frac{\partial}{\partial \mathbf{b}} l(\boldsymbol{\beta}, \phi, \boldsymbol{\vartheta}, \mathbf{b}) = \frac{\partial}{\partial \mathbf{b}} \left( \text{const} + \frac{\boldsymbol{\theta}'\mathbf{y} - b(\boldsymbol{\theta})'\mathbf{1}}{\phi} - \frac{1}{2}\mathbf{b}'\mathbf{G}(\boldsymbol{\vartheta})^{-1}\mathbf{b} \right) \\ &= \frac{1}{\phi} (\mathbf{Z}'\mathbf{y} - \mathbf{Z}' \text{diag}(b'(\boldsymbol{\theta}))\mathbf{1}) - \mathbf{G}(\boldsymbol{\vartheta})^{-1}\mathbf{b} \\ &= \frac{1}{\phi} \mathbf{Z}'(\mathbf{y} - \boldsymbol{\mu}) - \mathbf{G}(\boldsymbol{\vartheta})^{-1}\mathbf{b} \quad \text{und}\end{aligned}$$

$$\begin{aligned}\mathcal{I}(\mathbf{b}) &= -\text{E} \left( \frac{\partial^2}{\partial \mathbf{b} \partial \mathbf{b}'} l(\boldsymbol{\beta}, \phi, \boldsymbol{\vartheta}, \mathbf{b}) \right) = - \left( \frac{\partial}{\partial \mathbf{b}'} s(\mathbf{b}) \right) \\ &= \frac{1}{\phi} \mathbf{Z} \text{diag}(b''(\boldsymbol{\theta}))\mathbf{Z}' + \mathbf{G}(\boldsymbol{\vartheta})^{-1} =: \mathbf{Z}\mathbf{W}\mathbf{Z}' + \mathbf{G}(\boldsymbol{\vartheta})^{-1}\end{aligned}$$



# P-IRLS

Fisher-Scoring ausgehend von Startwert  $\mathbf{b}_0$ :

$$\mathcal{I}(\mathbf{b}_0)\mathbf{b} = \mathcal{I}(\mathbf{b}_0)\mathbf{b}_0 + s(\mathbf{b}_0)$$

mit im GLMM:  $s(\mathbf{b}) = \frac{1}{\phi} \mathbf{Z}'(\mathbf{y} - \boldsymbol{\mu}) - \mathbf{G}(\boldsymbol{\vartheta})^{-1} \mathbf{b};$

$$\mathcal{I}(\mathbf{b}) = \mathbf{Z}'\mathbf{W}\mathbf{Z} + \mathbf{G}(\boldsymbol{\vartheta})^{-1}$$

liefert mit  $\mathbf{W}_0 = \mathbf{W}(\mathbf{b}_0)$ ,  $\boldsymbol{\mu}_0 = \boldsymbol{\mu}(\mathbf{b}_0)$ :

$$(\mathbf{Z}'\mathbf{W}_0\mathbf{Z} + \mathbf{G}(\boldsymbol{\vartheta})^{-1})\mathbf{b} = (\mathbf{Z}'\mathbf{W}_0\mathbf{Z} + \mathbf{G}(\boldsymbol{\vartheta})^{-1})\mathbf{b}_0 + \frac{1}{\phi} \mathbf{Z}'(\mathbf{y} - \boldsymbol{\mu}_0) - \mathbf{G}(\boldsymbol{\vartheta})^{-1} \mathbf{b}_0$$

$$\Leftrightarrow (\mathbf{Z}'\mathbf{W}_0\mathbf{Z} + \mathbf{G}(\boldsymbol{\vartheta})^{-1})\mathbf{b} = \underbrace{\mathbf{Z}'\mathbf{W}_0 \left( \mathbf{Z}\mathbf{b}_0 + \frac{1}{\phi} \mathbf{W}_0^{-1}(\mathbf{y} - \boldsymbol{\mu}_0) \right)}_{\text{working response } \tilde{\mathbf{y}}}$$

$\Rightarrow$  Schätzugleichung eines LMM mit bekanntem  $\mathbf{W}_0$ ,  $\mathbf{G}(\boldsymbol{\vartheta})$ , vgl. (14):

$$\tilde{\mathbf{y}}|\mathbf{b} \sim N(\mathbf{Z}\mathbf{b}, \mathbf{W}_0^{-1}); \mathbf{b} \sim N(\mathbf{0}, \mathbf{G}(\boldsymbol{\vartheta}))$$

# P-IRLS

P-IRLS-Algorithmus iteriert folgende Schritte bis zur Konvergenz von  $\hat{\mathbf{b}}$ :

- i. setze Iterationswert  $\mathbf{b}_0 = \hat{\mathbf{b}}^{(k)}$ , berechne mit  $\beta$ ,  $\mathbf{b}_0$  die working responses  $\tilde{\mathbf{y}}$  und Gewichte  $\mathbf{W}_0$
- ii. berechne  $\hat{\mathbf{b}}^{(k+1)}$  als Lösung des daraus abgeleiteten gewichteten, penalisierten KQ-Problems

# (Adaptive) Gauss'sche Quadratur

- Laplace-Approximation recht schnell, aber ungenau besonders für kleine Clustergrößen und starke „Diskretheit“ (logistische Regression schlechter als Poisson-Regression).
- Gauss-Quadratur genauere Methode, um  $\int f(\mathbf{y}|\boldsymbol{\beta}, \phi, \mathbf{b})f(\mathbf{b}|\boldsymbol{\vartheta})d\mathbf{b}$  zu approximieren, aber wesentlich rechenaufwändiger.
- Benutzt orthonormalisierte zufällige Effekte  $\mathbf{b}^* = \mathbf{G}(\boldsymbol{\vartheta})^{-1/2}\mathbf{b}$ .
- $\int f(\mathbf{y}|\boldsymbol{\beta}, \phi, \boldsymbol{\vartheta}, \mathbf{b}^*)f(\mathbf{b}^*)d\mathbf{b}^* \approx \sum_{k=1}^Q w_k f(\mathbf{y}|\boldsymbol{\beta}, \phi, \boldsymbol{\vartheta}, \mathbf{b}_k^*)$
- Die Stützstellen  $\mathbf{b}_k^*$  ergeben sich aus den Nullstellen des Q-ten Hermite-Polynoms und bestimmen auch die Gewichte  $w_k$ .

# (Adaptive) Gauss'sche Quadratur

- Genauigkeit der Approximation steigt mit wachsendem  $Q$   
⇒ erhöhe  $Q$  solange bis keine Änderung in Schätzung mehr zu beobachten
- in `lme4` nur implementiert für Modelle mit *einer* einzigen Gruppierungsvariable (option: `nAGQ`)
- Laplace-Approximation ergibt sich als Spezialfall  $Q = 1$ .

# Penalized Quasi-Likelihood (PQL):

Ähnliche Idee wie P-IRLS: Approximiere  $\mathbf{y}$  durch  $\boldsymbol{\mu} = E(\mathbf{y}|\mathbf{b})$  plus Fehler mit Varianz  $\text{Var}(\mathbf{y}|\mathbf{b})$ . Eine Taylor-Approximation von  $\boldsymbol{\mu}$  um die aktuellen Werte  $\boldsymbol{\beta}_0$  und  $\mathbf{b}_0$  und Umsortieren ergibt

$$\tilde{\mathbf{y}} := \mathbf{X}\boldsymbol{\beta}_0 + \mathbf{Z}\mathbf{b}_0 + \frac{1}{\phi} \mathbf{W}_0^{-1}(\mathbf{y} - \boldsymbol{\mu}_0) = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b} + \boldsymbol{\varepsilon}.$$

Algorithmus:

- 1 Für gegebene Werte  $\boldsymbol{\beta}_0 = \hat{\boldsymbol{\beta}}^{(k)}$ ,  $\mathbf{b}_0 = \hat{\mathbf{b}}^{(k)}$  bestimme *working responses*  $\tilde{\mathbf{y}}$ .
- 2 Bestimme  $\hat{\boldsymbol{\beta}}^{(k+1)}$ ,  $\hat{\mathbf{b}}^{(k+1)}$  und  $\hat{\boldsymbol{\vartheta}}^{(k+1)}$ ,  $\hat{\phi}^{(k+1)}$  aus dem LMM  $\tilde{\mathbf{y}} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b} + \boldsymbol{\varepsilon}$  mit  $\mathbf{b} \sim N(\mathbf{0}, \mathbf{G}(\boldsymbol{\vartheta}))$  und  $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{W}_0^{-1})$

Iteriere bis zur Konvergenz.

# PQL

- Alternative Herleitung für PQL: ergibt sich ebenfalls bei Fisher-Scoring für die gemeinsame/penalisierte log-Likelihood (die im LMM die BLUPs ergibt). Daher auch mögliche Schätzart bei GAMMs.
- in R implementiert in Funktion `glmmPQL` im Paket MASS (`glmmPQL` wird auch von Funktion `gamm` im Paket `mgcv` benutzt)
- benutzt iterierte ML- oder REML-Schätzung
- Schätzung der Varianzparameter nach unten verzerrt, v.a. falls  $n_i$  klein. Bessere Approximation, je näher Daten an Normalverteilung (z.B. Poisson-Verteilung mit größeren  $\lambda$ 's)
- Konvergenz nicht garantiert
- Kein AIC, BIC oder Devianz berechenbar

# Vergleich glmmPQL/lme vs. (g)lmer

	MASS::glmmPQL / nlme::lme	lme4::(g)lmer
Daten	nur genestete Daten; grosse Datensätze/Modelle oft nicht zu fitten	genestete & gekreuzte Datenstrukturen; auch riesige Datensätze werden gefittet
(G)AMs Cov( $\mathbf{b}_i$ )	via mgcv::gamm sehr breite, erweiterbare Klasse von Kovarianzstrukturen (s. nlme::pdMat)	via gamm4 nur unstrukturierte oder diagonale Kovarianzen (in gamm4 beliebige Präzisionsmatrizen $\mathbf{P}$ mit $\mathbf{G}(\vartheta) = \vartheta \mathbf{P}^{-1}$ )
Cov( $\varepsilon$ )	sehr breite, erweiterbare Klasse von Kovarianzstrukturen (s. nlme::varFunc)	Cov( $\varepsilon$ ) = $\sigma^2 \mathbf{I}_n$ oder $\sigma^2 \text{diag}(\mathbf{w}^{-1})$ mit bekannten Gewichten $\mathbf{w}$
Stabilität	sehr instabil für komplexe Strukturen von $\mathbf{b}$	sehr stabil für LMMs; für GLMMs ab $> 3$ zuf. Effekten oft kritisch
Speed	relativ langsam	LA sehr schnell (benutzt Sparse-Matrix-Algorithmen); AGQ deutlich langsamer

# Vorhersage der zufälligen Effekte

Die beste Vorhersage für  $\mathbf{b}$  (minimaler mean squared error of prediction  $E(\mathbf{b} - \hat{\mathbf{b}})^2$ ) wäre wieder

$$\hat{\mathbf{b}} = E(\mathbf{b}|\mathbf{y}).$$

In der Praxis würde man  $\hat{\beta}$  und  $\hat{\vartheta}$  einsetzen. Allerdings erfordert

$$\hat{\mathbf{b}} = \int \mathbf{b}f(\mathbf{b}|\mathbf{y}, \beta, \vartheta)d\mathbf{b} = \frac{\int \mathbf{b}f(\mathbf{y}|\mathbf{b}, \beta, \phi)f(\mathbf{b}|\vartheta)d\mathbf{b}}{\int f(\mathbf{y}|\mathbf{b}, \beta, \phi)f(\mathbf{b}|\vartheta)d\mathbf{b}}$$

wieder numerische Integration.

Alternativ zum Posteriori-EW wird häufig (z.B. in lme4 als Teil von P-IRLS) der Posteriori-Modus berechnet, der  $f(\mathbf{b}|\mathbf{y}, \beta, \vartheta) \propto f(\mathbf{y}|\mathbf{b}, \beta, \phi)f(\mathbf{b}|\vartheta)$  maximiert. Die beiden unterscheiden sich i.A. außer im LMM unter NV.

PQL schätzt  $\mathbf{b}$  direkt im Schätzalgorithmus.



# Hypothesen-Tests für $\beta$

- Wegen der Maximum-Likelihood-Schätzung von  $\beta$  können prinzipiell Wald-, Likelihood-Quotienten- (LQT) oder Score-Tests verwendet werden mit einer entsprechenden  $\chi^2$ -Verteilung als Referenzverteilung.
- Die Güte der Approximation hängt ab von
  - der Güte der Approximation an die Likelihood in der Schätzung
  - der Güte der asymptotischen Approximation, die nur für Longitudinal/Clusterdaten bei  $N \rightarrow \infty$  greift.
- PQL-Schätzung beruht auf der Likelihood von Pseudodaten und daher ist kein Likelihood-Quotiententest möglich. Inferenz für PQL wird meist basiert auf dem LMM in der PQL-Schätzung. Allerdings ist  $\hat{\beta}$  i.A. nicht konsistent.

# Inferenz für $D$

## Hypothesen-Tests

- Für das longitudinale/Cluster-GLMM gelten die gleichen asymptotischen Ergebnisse für Tests von Parametern in  $D$  wie im LMM (Rand des Parameterraums).
- Es ist keine exakte Verteilung vorhanden.

## Modellselektion

- Ein konditionales AIC für Poisson-, binomial- oder normalverteilte Zielgrößen ist im R-Paket `cAIC4` implementiert, siehe auch Saefken, Kneib, van Waveren & Greven (2014). A unifying approach to the estimation of the conditional Akaike information in generalized linear mixed models. *Electronic Journal Statistics*, 8, 201-225.

# Konfidenzintervalle

- Generell gilt für einen Parameter  $\theta$ :  
Ein Wert  $\theta_0$  ist im  $(1 - \alpha)\%$ -Konfidenzintervall für  $\theta \Leftrightarrow$  Die Hypothese  $H_0 : \theta = \theta_0$  gegen  $H_A : \theta \neq \theta_0$  wird zum Level  $\alpha$  nicht verworfen.  
(Sofern Test und Konfidenzintervall auf der gleichen Statistik beruhen.)
- Im GLMM: Für Parameter  $\theta$  in  $\beta$  oder  $\mathbf{D}$ , betrachte ein Gitter von  $\theta_0$ -Werten um den Schätzer  $\hat{\theta}$ .
- Konstruiere das Konfidenzintervall für  $\theta$  so, dass alle Werte auf dem Gitter enthalten sind, für die ein LQT mit Referenzverteilung  $\chi_1^2$   $H_0$  nicht ablehnt.
- Dieser Ansatz funktioniert auch bei nicht-symmetrischen Verteilungen von Schätzern, jedoch nicht bei Parametern  $\theta_0$  in  $\mathbf{D}$  nahe des Randes des Parameterraums.
- Der Ansatz ist in `lme4` implementiert in der Funktion `confint`. Alternativen in `confint` sind Wald- (für  $\beta$ ) und Bootstrap-basierte Konfidenzintervalle.