M8: Generalized linear mixed models GLMM

TMA4315
15.11.2018

Aim: analyse correlated observations with non-normal response

GLM: (Y_i, x_i) $i = 1, .., n$ independent pairs

1. Random component: Y_i from exponential family
   Beaches: $Y_i \sim \text{Poisson}(\mu_i)$, $\mu_i = E(Y_i)$
2. Systematic component: linear predictor $\eta_i = x_i^T \beta$
3. Link/response function

$$\eta_i = g(\mu_i) \quad \text{or} \quad \mu_i = h(\eta_i)$$

Beaches: $\eta_i = \ln(\mu_i)$, $\mu_i = \exp(\eta_i)$

a) Population model: black curve: $\hat{\mu} = \exp(\hat{\eta}) = \exp(\hat{\beta}_0 + \hat{\beta}_1 x)$

b) Each beach fixed intercept: $\hat{\mu}_{ij} = \exp(\hat{\eta}_{ij}) = \exp(\hat{\beta}_0 + \hat{\beta}_{\text{beach}, i} + \hat{\beta}_1 x_{ij})$

GLMM: we add a random component to the linear predictor and let $\gamma_i \sim N(0, Q)$ as for LMM.
+ distribution of Y_{ij} is conditional on γ_i.

cluster i
$i = 1, .., m$

measure within cluster
$j = 1, .., n_i$

GLMM

Distributional assumptions

$$Y_{ij} \mid \gamma_i, x_{ij}, u_{ij}, \beta, \cancel{Q}$$

parameters

response cluster i obs j

$i = 1, \dots, m$
$j = 1, \dots, n_i$

random effects

$(q+1) \times 1$

elements (rows) of design matrix

are conditionally independent

for all i and j

from $f(y_{ij} \mid \gamma_i, x_{ij}, u_{ij}, \beta, \cancel{Q})$

exponential family

Structural assumptions

$$\mu_{ij} = E(Y_{ij} \mid \gamma_i, x_{ij}, u_{ij}, \beta) \quad \text{conditional mean}$$

$$\eta_{ij} = x_{ij}^T \beta + u_{ij}^T \gamma_i$$

$$\eta_{ij} = g(\mu_{ij}) \quad \text{and} \quad \mu_{ij} = h(\eta_{ij})$$

NB: now conditional mean is linked to linear predictor

Distributional assumptions on random effects

γ_i i.i.d $\qquad \gamma_i \sim N(0, Q)$

$i = 1, \dots, m \qquad Q: (q+1) \times (q+1)$

How to simulate data (to be analysed with GLMM)?

1) Decide on β and elements of Q

β_0, β_1 random intercept τ_0^2

2) Draw $\gamma_i \sim N(0, Q)$

Choose some values

3) x_{ij}, u_{ij} ← fixed or generate

4) $\eta_{ij} = x_{ij}^T \beta + u_{ij}^T \gamma_i$ and $\mu_{ij} = h(\eta_{ij})$

5) Draw $Y_{ij} \sim f(\mu_{ij})$

Beaches: Poisson $\left(\exp\left(x_{ij}^T \beta + u_{ij}^T \gamma_i \right) \right)$

GLMM with random intercept

Distributional assumptions: $Y_{ij} \mid \gamma_i, x_{ij}, u_{ij}$

[Poisson for Beaches] $\mu_{ij} = E(Y_{ij} \mid \gamma_i, x_{ij}, u_{ij})$

Structural assumptions:

$$\eta_{ij} = \underbrace{x_{ij}^T \beta + \gamma_{0i}} \quad \text{and} \quad \eta_{ij} = \log(\mu_{ij})$$

[Beaches: $\beta_0 + \beta_1 NAP_{ij} + \gamma_{0i}$]

Distribution of $\gamma_{0i} \sim N(0, \tau_0^2)$

Q: Beaches with NAP: # parameters to estimate: $\beta_0, \beta_1, \tau_0^2$ → 3

Beeches (print-out): ML (Laplace approx)

$$\hat{\mu}_{ij} = \exp(\hat{\beta}_0 + \hat{\beta}_1 x_j^{NAP} + \hat{\gamma}_{i0}) = \exp(\hat{\beta}_0 + \hat{\beta}_1 x_j) \cdot \exp(\hat{\gamma}_{i0})$$

Not just a shift in the intercept on exp. scale

→ but a multiplicative effect

(for LMM: $\hat{\mu}_{ij} = \hat{\beta}_0 + \hat{\beta}_1 x_{ij} + \hat{\gamma}_{i0}$)

Parameter estimation

Conditional model: $f(y_{ij} \mid \gamma_i)$ and $f(\gamma_i)$

$f(y_{ij} \mid \gamma_i)$: Normal, binomial, Poisson, gamma

$f(\gamma_i)$: Normal

Joint model: $f(y_{ij}, \gamma_i) = f(y_{ij} \mid \gamma_i) \cdot f(\gamma_i)$

Marginal model: $f(y_{ij}) = \int_{\gamma_i} f(y_{ij} \mid \gamma_i)\, f(\gamma_i)\, d\gamma_i$

$f(y_{ij})$	$f(y_{ij} \mid \gamma_i)$	$f(\gamma_i)$
N	N	N
?	Poisson, binomial	N

only for special cases this is an closed form

Likelihood: $L(\beta, Q) = \prod_{i=1}^{m} f(y_i \mid \beta, Q)$

$$= \prod_{i=1}^{m} \int_{\gamma_i} \prod_{j=1}^{n_i} f(y_{ij} \mid \gamma_i, \beta) \cdot f(\gamma_i \mid Q)\, d\gamma_i$$

Inference: as for LMM (and GLM)

$$\hat{\beta} \approx N(\beta, F(\hat{\beta})^{-1})$$

and deviance for model fit and AIC for model selection.