MATH43515: Multilevel Modelling

Lecture 8: Generalized Linear Mixed Models

Module Convenor / Tutor: Andy Golightly

Durham University

This lecture does not follow any particular book, but makes use of some aspects of

Aitkin, M., Francis, B., Hinde, J., and Darnell, R.
Statistical Modelling in R (2009).

... also available via the library.

Outline (Lecture 8)

- Combining multilevel models and generalized linear models
- Example: The betablocker data (logistic mixed model)
- Software considerations

Recall...

- We know, for Gaussian response models ("linear models"), how to account for data with hierarchical (multilevel) structure
- We know, for one-level models, how to account for data with non-Gaussian response.

Gap: Data with multilevel structure and non-Gaussian response.

Models which do this job are referred to as generalized linear mixed models, a combination of mixed models ($\equiv$ multilevel models) and generalized linear models.

Recall the earlier defined random-intercept-and-slope model (Lecture 4), for two levels (i=lower, j=upper):

$$y_{ij} = a + bx_{ij} + u_j + v_j x_{ij} + \epsilon_{ij}$$

We worked out that

$$E(y_{ij}|x_{ij}) = a + bx_{ij}$$

The more relevant quantity for the current context is

$$E(y_{ij}|x_{ij}, u_j, v_j) = a + bx_{ij} + u_j + v_j x_{ij} \tag{1}$$

Thinking now of a situation with p predictor variables, we would have observations $x_{ij1}, \ldots, x_{ijp}$, yielding $\boldsymbol{x}_{ij} = (1, x_{ij1}, \ldots, x_{ijp})^T$, and a corresponding upper-level vector of random effects $\boldsymbol{z}_j = (u_j, v_{j1}, \ldots v_{jp})^T$.

Hence, equation (1) can be written as

$$E(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{z}_j) = \boldsymbol{x}_{ij}^T\boldsymbol{\beta} + \boldsymbol{x}_{ij}^T\boldsymbol{z}_j$$

which makes the mixed model character of the multilevel model explicit.

Recall also the generalized linear model (at one level), with

$$E(y_i|\boldsymbol{x}_i) = h(\boldsymbol{x}_i^T\boldsymbol{\beta}).$$

From this, we define a model of type

$$E(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{z}_j) = h(\boldsymbol{x}_{ij}^T\boldsymbol{\beta} + \boldsymbol{x}_{ij}^T\boldsymbol{z}_j)$$

for some response function h (link function $g = h^{-1}$). The result is a generalized linear mixed model (GLMM).

Large multi-centre study (meta-analysis of several clinical trials) to check the effectiveness of beta-blockers in reducing mortality after myocardial infarction (beta-blockers "block" the release of certain hormones, to reduce blood pressure).

Our data set has 44 rows: One for each of treatment and control, for each of 22 centres.

For each centre and treatment, `r` patients died out of `n` treated.

Betablocker data

```r
betablok <-
  read.table('https://andygolightly.github.io/teaching/MATH43515/betablok.dat')
```

Add trial and centre information:

```r
names(betablok) <- c('r','n')
betablok$treat <- factor(gl(2,1),labels=c(0,1))
betablok$center <- gl(22,2)
```

```r
head(betablok)
```

```
##    r   n treat center
## 1  3  39     0      1
## 2  3  38     1      1
## 3 14 116     0      2
## 4  7 114     1      2
## 5 11  93     0      3
## 6  5  69     1      3
```

```r
dim(betablok)
```

```
## [1] 44  4
```

Firstly, ignore two-level structure and fit a GLM, with Binomial response and logit link:

```
betablok.glm <- glm(cbind(r,(n-r))~treat, data=betablok, family=binomial)
summary(betablok.glm)
```

```
##
## Call:
## glm(formula = cbind(r, (n - r)) ~ treat, family = binomial, data = beta
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -5.3160  -1.4916  -0.1341   1.7067   5.8564
##
## Coefficients:
```

```
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.19711    0.03359 -65.417  < 2e-16 ***
## treat1      -0.25737    0.04942  -5.207 1.91e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 332.99  on 43  degrees of freedom
## Residual deviance: 305.76  on 42  degrees of freedom
## AIC: 527.19
##
## Number of Fisher Scoring iterations: 4
```

Betablocker data: logistic mixed model

Now, recognize the two-level structure, with centers in the upper level:

```
require(lme4)
betablok.glmer <- glmer(cbind(r,(n-r)) ~ treat + (1|center),
     data=betablok, family=binomial)
summary(betablok.glmer)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: cbind(r, (n - r)) ~ treat + (1 | center)
##    Data: betablok
##
##      AIC      BIC   logLik deviance df.resid
##    324.4    329.8   -159.2    318.4       41
##
## Scaled residuals:
```

Betablocker data: logistic mixed model

```
##     Min      1Q  Median      3Q     Max
## -1.8876 -0.5129  0.0605  0.4969  1.8623
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  center (Intercept) 0.2362   0.486
## Number of obs: 44, groups:  center, 22
##
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.19618    0.11292 -19.450  < 2e-16 ***
## treat1      -0.26091    0.04982  -5.237 1.63e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr)
## treat1 -0.205
```

We have seen that, in this particular case, accounting for the multilevel structure did not make much difference – but it could have!

The treatment of myocardial infarction patients with beta-blockers reduces the log-odds of patient death by the value 0.26. That is, after treatment, the odds of dying against not dying have reduced by the factor $\exp(-0.26) = 0.77$, i.e. 77% of the untreated value.

Summary

Generalized linear mixed models are in principle simple to apply: Just use `glmer` (or `glmmTMB`) with the respective `family` argument.

We have seen the R package **lme4** (functions `lmer` and `glmer`). Other R packages for random effect modelling in R include

- **nlme** (for non-linear models)
- **npmlreg** (for non-Gaussian random effect distributions),
- **gamlss** (very general modelling framework, incl. smoothing).