Generalized Linear Mixed Model

Biostatistics 653

Applied Statistics III: Longitudinal Analysis

GLMM

- We now discuss the incorporation of random effects into generalized linear models. The basic idea is that we assume there is natural heterogeneity across individuals in a subset of the regression coefficients. A subset of these coefficients are assumed to vary across subjects according to some distribution.
- Conditional on the random effects, we assume the responses for a single subject are independent observations from a distribution in the exponential family.

Review: LMM

We can consider the linear mixed effects model in two steps.

- Assume Y_i has a normal distribution that depends both on population level effects, β, and individual-specific effects, b_i, with

$$Y_i = X_i\beta + Z_i b_i + \epsilon_i, \epsilon_i \sim N(0, R_i)$$

- Assume b_i are independent across subjects and $b_i \sim N(0, D)$.

- The response of subject i at occasion j differs from the population mean $X_{ij}^T\beta$ by a subject effect, b_i, and a within-subject error, ϵ_{ij}.

Review: LMM

- Note that $R_i = Cov(\epsilon_i)$ describes the covariance among observations when we focus on the response prole of a single subject. It is the covariance of subject i's deviation from his or her mean prole $X_i\beta + Z_i b_i$. Often, we assume $R_i = \sigma^2 I$, which is termed the conditional independence assumption.

- Also recall that the vector of regression parameters β are the fixed effects, which are assumed to be the same for all subjects. However, the subject-specific random effects b_i are used to describe the mean response prole of a single individual (when combined with the fixed effects).

Review: LMM

- Finally, recall that in the mixed model

$$Y_i = X_i\beta + Z_i b_i + \epsilon_i,$$
$$E(Y_i|b_i) = X_i\beta + Z_i b_i,$$
$$E(Y_i) = X_i\beta$$

- Similarly,

$$V(Y_i|b_i) = V(\epsilon_i) = R_i = \sigma^2 I$$
$$V(Y_i) = V(Z_i b_i) + V(\epsilon_i) + 2Cov(Z_i b_i, \epsilon_i) = Z_i D Z_i^T + R_i$$

- So we see that the introduction of the random effects, b_i, has induced correlation (marginally) among the Y_i.

GLMM

We can also consider the generalized linear mixed model in two steps.

- Assume the conditional distribution of each Y_{ij}, conditional on b_i, belongs to the exponential family with conditional mean

$$g\left(E\left[Y_{ij}\middle|b_i\right]\right) = X_{ij}^T\beta + Z_{ij}^T b_i$$

where g(.) is a known link function.

- Assume the b_i are independent across subjects with $b_i \sim N(0, D)$. We also assume that given b_i, the responses $Y_{i1}, \cdots, Y_{in}$ are mutually independent.

GLMM Examples

- Binary logistic model with random intercepts:

$$logit\left(E\left[Y_{ij}\middle|b_i\right]\right) = X_{ij}^T\beta + b_i$$

where $b_i \sim N(0, \sigma^2)$.

- Random coefficients Poisson regression:

$$log\left(E\left[Y_{ij}\middle|b_i\right]\right) = X_{ij}^T\beta + Z_{ij}^T b_i$$

where $X_{ij} = Z_{ij} = [1, t_{ij}]$ (random slopes and intercepts) and $b_i \sim N(0, D)$

Interpretation of GLMM

- Mixed effects models are most useful when the scientific objective is to make inferences about individuals. The primary focus is on the individual and on the influence of covariates on the individual. Regression parameters, β, measure the direct influence of covariates on the responses of heterogeneous individuals.

- For example, in the model

$$logit\left(E\left[Y_{ij}\middle|b_i\right]\right) = X_{ij}^T\beta + b_i$$

where $b_i \sim N(0, \sigma^2)$. Each element of measures the change in the log odds of a 'positive' response per unit change in the respective covariate, for a specific subject who has an underlying propensity to respond positively given by b_i.

Interpretation of GLMM

- Note also that with a non-linear link function, a "non-linear contrast of the averages" is not equal to the "average of non-linear contrasts," so that the parameters do not in general have population-average interpretations (as they would in a linear mixed effects model, which has identity link). So while

$$g\left(E\left(Y_{ij}\middle|X_{ij},b_i\right)\right)=X_{ij}^T\beta+Z_{ij}^Tb_i$$

when g(.) is non-linear (e.g. log or logit), then

$$g\left(E\left(Y_{ij}\middle|X_{ij}\right)\right)\neq X_{ij}^T\beta$$

for all β when averaged over the distribution of the random effects.

Marginal Moments

- In GLMM, the marginal mean of Y_{ij} is

$$E(Y_{ij}) = E\left(E(Y_{ij}|b_i)\right) = E(\mu_{ij}^b) = E_{b_i}(g^{-1}(X_{ij}^T\beta + Z_{ij}^T b_i))$$

- The marginal variance is

$$V(Y_{ij}) = E\left(V(Y_{ij}|b_i)\right) + V\left(E(Y_{ij}|b_i)\right)$$
$$= E_{b_i}\left(V(Y_{ij}|b_i)\right) + V_{b_i}\left(g^{-1}(X_{ij}^T\beta + Z_{ij}^T b_i)\right)$$

Marginal Moments

- The marginal covariance is

$$Cov(Y_{ij}, Y_{ik}) = E\left(Cov(Y_{ij}, Y_{ik}|b_i)\right) + Cov\left(E(Y_{ij}|b_i), E(Y_{ik}|b_i)\right)$$
$$= 0 + Cov\left(g^{-1}(X_{ij}^T\beta + Z_{ij}^T b_i), g^{-1}(X_{ik}^T\beta + Z_{ik}^T b_i)\right)$$
$$= Cov\left(g^{-1}(X_{ij}^T\beta + Z_{ij}^T b_i), g^{-1}(X_{ik}^T\beta + Z_{ik}^T b_i)\right)$$

GLMM Example

- A study on seizure collected data from 59 epileptics. The number of epileptic seizures were recorded during a baseline period of 8 weeks. After 8 weeks, patients were randomized to receive either treatment (progabide) or placebo. The number of seizures was recorded in four consecutive two-week periods after the treatment.

GLMM Example

We denote:

- Y_{ij}: the number of seizures on patient i at occasion j
- $time_{ij}$: time at occasion j (i.e. $\mathrm{t_{ij}} = 0, 2, 4, 6, 8$ for $j = 0,1,2,3,4$, respectively)
- d_{ij}: the observation period before occasion j (i.e. $\mathrm{d}_{ij} = 8, 2, 2, 2, 2$ for $j = 0,1,2,3,4$, respectively)
- t_i: a binary indicator on whether i'th patient received treatment
- x_{ij}: a binary indicator on whether j'th observation is made at the baseline (i.e. $x_{ij} = 1$ if $j = 0$; and $x_{ij} = 0$ otherwise)

where $i = 1, \cdots, 59; j = 0,1,2,3,4$

GLMM Example

- We consider the following generalized linear mixed model

$$Y_{ij}|b_i \sim Poi(d_{ij}\lambda_{ij})$$

$$\log(\lambda_{ij}) = \beta_0 + t_i\beta_1 + x_{ij}\beta_2 + t_i x_{ij}\beta_3 + b_i$$

$$b_i \sim N(0, \sigma^2)$$

GLMM Example

Parameter interpretation:

- $\exp(\beta_0)$ is the expected rate of seizures for a typical individual in the placebo group at time 0 (or for a specific individual who has an random intercept of b_i)
- $\exp(\beta_1)$ is the ratio of expected seizure rate in the treatment versus the placebo group for a typical individual at time 0
- $\exp(\beta_2)$ is the ratio of expected seizure rate at time j=1,2,3,4 as compared to j=0 in the placebo group for a typical individual
- $\exp(\beta_3)$ is the ratio of expected seizure rate at time j=1,2,3,4 as compared to j=0 in the treatment group, relative to the ratio of expected seizure rate at time j=1,2,3,4 as compared to j=0 in the placebo group, for a typical individual

GLMM Example

- To simplify the algebra, we now consider a reduced model with

$$\log(\lambda_{ij}) = \beta_0 + t_i\beta_1 + b_i$$

- With the model, we can compute the conditional mean as

$$\mu_{ij}^b = E(Y_{ij}|b_i) = d_{ij}\exp(\beta_0 + t_i\beta_1 + b_i)$$

- And the marginal mean as

$$E(Y_{ij}) = E_{b_i}\left(E(Y_{ij}|b_i)\right) = E_{b_i}\left(d_{ij}\exp(\beta_0 + t_i\beta_1 + b_i)\right)$$
$$= d_{ij}e^{\beta_0+t_i\beta_1}E_{b_i}(e_i^b) = d_{ij}e^{\beta_0+t_i\beta_1+\sigma^2/2}$$

GLMM Example

- The marginal median is

$$d_{ij}e^{\beta_0+t_i\beta_1}$$

- The marginal variance is

$$V(Y_{ij}) = E(\mu_{ij}^b) + V(\mu_{ij}^b) = E(Y_{ij}) + d_{ij}^2 e^{2\beta_0+2t_i\beta_1} V_{b_i}(e^{b_i})$$
$$= E(Y_{ij}) + E(Y_{ij})^2 (e^{\sigma^2} - 1) = E(Y_{ij})(1 + E(Y_{ij}) \times \kappa)$$

where $\kappa = e^{\sigma^2} - 1 > 0$, enabling the model to account for extra-Poisson variation

GLMM Example

- The marginal covariance is

$$Cov(Y_{ij}, Y_{ik}) = Cov(d_{ij}e^{\beta_0+t_i\beta_1+b_i}, d_{ik}e^{\beta_0+t_i\beta_1+b_i})$$
$$= d_{ij}d_{ik}e^{2\beta_0+2t_i\beta_1}V(e^{b_i}) = E(Y_{ij})E(Y_{ik})\kappa$$

- Hence, for individual i we have the variance covariance matrix as

$$\begin{pmatrix} \mu_{i1}+\mu_{i1}^2\kappa & \mu_{i1}\mu_{i2}\kappa & \cdots & \mu_{i1}\mu_{in}\kappa \\ \mu_{i2}\mu_{i1}\kappa & \mu_{i2}+\mu_{i2}^2\kappa & \cdots & \mu_{i2}\mu_{in}\kappa \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{in}\mu_{i1}\kappa & \mu_{in}\mu_{i2}\kappa & \cdots & \mu_{in}+\mu_{in}^2\kappa \end{pmatrix}$$

where $\kappa = e^{\sigma^2} - 1 > 0$, and a single parameter σ^2 controls both excess-Poisson variability and dependence.

Relationship of β in GEE and GLMM

- Recall that in GEE, we have

$$E(Y_{ij}) = \mu_{ij} = g^{-1}(X_{ij}^T\beta)$$

- In GLMM, we have

$$E(Y_{ij}) = E_{b_i}(\mu_{ij}^b) = E_{b_i}\left(g^{-1}(X_{ij}^T\beta + Z_{ij}^T b_i)\right) \neq g^{-1}(X_{ij}^T\beta)$$

i.e. generally cannot switch the order of E and g^{-1}

- Because of this difference, β in GEE has a population-average interpretation, while β in GLMM has a subject-specific interpretation, which is always conditional on subject-specific random effects b_i.

Relationship of β in GEE and GLMM

- For normal data with a linear link, in GLMM we have

$$E(Y_{ij}) = E_{b_i}(\mu_{ij}^b) = E_{b_i}(X_{ij}^T\beta + Z_{ij}^T b_i) = X_{ij}^T\beta$$

- Therefore, in this case, β in GEE and GLMM have the same interpretation.

Relationship of β in GEE and GLMM

- For binary data with a probit link, in GLMM we have

$$E(Y_{ij}) = E_{b_i}(\mu_{ij}^b) = E_{b_i}\left(\Phi\left(X_{ij}^T\beta + Z_{ij}^T b_i\right)\right)$$
$$= \Phi\left(\frac{1}{\sqrt{1 + Z_{ij}^T D Z_{ij}}} X_{ij}^T\beta\right)$$

- Therefore,

$$\Phi^{-1}\left(E(Y_{ij})\right) = \frac{1}{\sqrt{1 + Z_{ij}^T D Z_{ij}}} X_{ij}^T\beta < X_{ij}^T\beta, if\ D \neq 0$$

- Hence, β_{GEE} in GEE is attenuated compared to β_{GLMM} in GLMM.

Relationship of β in GEE and GLMM

- For binary data with a logit link, in GLMM we have

$$E(Y_{ij}) = E_{b_i}(\mu_{ij}^b) = E_{b_i}\left(\frac{e^{X_{ij}^T\beta + Z_{ij}^T b_i}}{e^{X_{ij}^T\beta + Z_{ij}^T b_i} + 1}\right)$$

$$\approx F\left(\frac{1}{\sqrt{1 + c^2 Z_{ij}^T D Z_{ij}}} X_{ij}^T\beta\right)$$

where F is the logistic CDF and $c = \frac{16\sqrt{3}}{15}\pi \approx 0.58$.

- Therefore,

$$logit\left(E(Y_{ij})\right) \approx \frac{1}{\sqrt{1 + c^2 Z_{ij}^T D Z_{ij}}} X_{ij}^T\beta < X_{ij}^T\beta, if\ D \neq 0$$

- Hence, β_{GEE} in GEE is attenuated compared to β_{GLMM} in GLMM.

Relationship of β in GEE and GLMM

- For count data with a log link, in GLMM we have

$$E(Y_{ij}) = E_{b_i}(\mu_{ij}^b) = E_{b_i}\left(e^{X_{ij}^T\beta + Z_{ij}^T b_i}\right) = e^{X_{ij}^T\beta + Z_{ij}^T D Z_{ij}}$$

- Therefore,

$$log\left(E(Y_{ij})\right) = X_{ij}^T\beta + Z_{ij}^T D Z_{ij}$$

- Hence, β_{GEE} in GEE differs from β_{GLMM} in GLMM by an intercept, if X and Z do not overlap except for the intercept (otherwise, $Z_{ij}^T D Z_{ij}$ will contribute to other β estimates).