

1 Objectives

We introduce some examples of linear mixed models (LMM). Note that random effects naively have **zero** ground mean.

2 Clustered Data: One-way ANOVA with Random Effects

$$y_{ij} = \mu + \alpha_i + \epsilon_{ij}, i = 1, m, j = 1, n_i \quad (1)$$

where, μ is unknown overall mean, α_i is random effects, e.g., class effects, follows $N(0, \sigma^2)$ and ϵ_{ij} are errors following i.i.d. $N(0, \tau^2)$ for student j in class i . α s and ϵ s are independent. We have some observations:

- 1) **Conditional** on each random effects α_i , y_{ij} s are **independent** (Why?)
- 2) y_{ij} s with common random effects are correlated (Why?)

$$Cov(y_{i,j_1}, y_{i,j_2}) = \sigma^2 + \tau^2, (j_1 = j_2) \quad (2)$$

and

$$Cov(y_{i,j_1}, y_{i,j_2}) = \sigma^2, (j_1 \neq j_2) \quad (3)$$

3)

$$Corr(y_{i,j_1}, y_{i,j_2}) = \frac{1}{1 + \tau^2/\sigma^2}, (j_1 \neq j_2) \quad (4)$$

- 4) The **larger** the variance of random effects, the **larger** the within-group correlation (Why?)
- 5) Random effects could describe the correlation structure and important!

3 Longitudinal Data

Molenberghs and Verbeke (2005) gave two two extensions of univariate linear models to longitudinal (repeated) measures.

3.1 Marginal Multivariate Model: Classical Repeated Measures (♠)

Each component of a subject is modeled using univariate linear model, a marginal covariance matrix is used to model the association structure.

$$\begin{aligned} Y_i &= (Y_{i,1}, Y_{i,2}, \dots, Y_{i,n_i}) \\ Y_{i,j} &= \beta_0 + \beta_1 t_{i,j} + \epsilon_{i,j} \\ \epsilon_i &= (\epsilon_{i,1}, \dots, \epsilon_{i,n_i}) \\ \text{i.e., } Y_i &\sim N(X_i \beta, V_i) \end{aligned}$$

Remark

- 1) If $V_i = \sigma^2 I_{n_i}$, we have univariate linear regression model, the repeated $Y_{i,j}$ is independent, or uncorrelated (Why?).
- 2) In balanced case, i.e., $n_i \equiv n$, covariance model $V_i = V$ leads to **classical multivariate regression** model.
- 3) Special V :

First-order autoregression:

$$\text{Cov}(Y_{i,j}, Y_{i,k}) = \sigma^2 \rho^{|t_{i,j} - t_{i,k}|}$$

Compound symmetry:

$$\text{Cov}(Y_{i,j}, Y_{i,k}) = \sigma^2 + \gamma \delta_{jk}, \quad \gamma > -\sigma^2, \quad \text{where } \delta_{jk} = 1 \text{ for } j = k \text{ and zero otherwise.}$$

Note that, they are homogeneous covariance structures where variance of each $Y_{i,j}$ is equal. This extension has been seen in traditional linear regression textbooks as **repeated measures**.

3.2 Linear Mixed Model (LMM)

3.2.1 Random effects model

- 1) For each subject, the responses can be modeled by a linear regression model.
- 2) Subject-specific regression coefficients could be random.

For example, for longitudinal data, we assume $Y_{i,j}$ measured at time point $t_{i,j}$, follows model

$$Y_{i,j} = \beta_{i,0} + \beta_{i,1} t_{i,j} + \epsilon_{i,j} \tag{5}$$

Model Assumptions

- 1) As in marginal multivariate model, we assume $\epsilon_i = (\epsilon_{i,1}, \dots, \epsilon_{i,n_i})$ follow $N(0, \Sigma_i)$.
- 2) We assume **subject-specific** regression coefficients $(\beta_{i,0}, \beta_{i,1})$ follow a distribution, say bivariate normal $N((\beta_0, \beta_1), D)$. This is a Bayesian theme!
- 3) We rewrite random effects model as

$$Y_{i,j} = (\beta_0 + b_{i,0}) + (\beta_1 + b_{i,1})t_{i,j} + \epsilon_{i,j} \quad (5)$$

where, $(b_{i,0}, b_{i,1})$ follow a normal bivariate normal $N((0,0), D)$, e.g., random effects with zero mean.

4 General Linear Mixed Model (LMM)

4.1 Hierarchical Structure

- 1) Denote Y_i to be repeated measures for the i -th subject
- 2) $Y_i | b_i \sim N(X_i \beta + Z_i b_i, \Sigma_i)$
- 3) $b_i \sim N(0, D)$ (prior in Bayesian approach?)

where,

- 1) fixed effects design matrix $X_i: n_i \times p$
- 2) random effects design matrix $Z_i: n_i \times q$
- 3) **fixed effects** β : p -dimensional
- 4) **random effects** b_i : q -dimensional

Mixed effects = fixed effects + random effects.

4.2 Marginal Formulation

The marginal distribution of Y_i is

$$f_i(y_i) = \int f_i(y_i | b_i) f(b_i) db_i \sim N(X_i \beta, V_i = Z_i D Z_i' + \Sigma_i) \quad (5)$$

Remark

- 1) Different random-effects may lead to identical marginal distributions.
- 2) Some marginal models could not be derived from mixed model. For example, in the compound symmetric covariance structure.

5 Compound Sampling Models

5.1 Hierarchical Structure within Bayesian Framework

Bradley and Louis (2000) Chapter 3.2.. Compound sampling applies to multi-site clinical trials, small-area disease rate estimation, longitudinal studies among others:

$$Y_i|\theta_i \sim f_i(y_i|\theta_i)(\text{independent}) \quad (5)$$

and

$$\theta_i \sim G(\cdot), (\text{i.i.d.}) \quad (5)$$

where $i = 1, 2, \dots, k$.

Remark: Y_i are marginally independent, i.e.,

$$m_G(y) = \prod_{i=1}^k m_G(y_i)$$

5.2 Parametric Empirical Bayesian (EB) Point Estimation

The posterior distribution of θ_i given data and hyperparameter η is

$$p(\theta_i|y_i, \eta) = \frac{f_i(y_i|\theta_i)g(\theta_i|\eta)}{m_i(y_i|\eta)} \quad (5)$$

We use marginal distribution of all data $m(y|\eta)$ to estimate unknown parameter η : $\hat{\eta} = \hat{\eta}(y)$ and further to estimate $\hat{\theta}_i$.

5.3 Example

$$Y_i|\theta_i \sim N(\theta_i, \sigma^2), i = 1, 2, \dots, k; (\text{i.i.d.}) \quad (5)$$

and

$$\theta_i|\mu \sim N(\mu, \tau^2), i = 1, 2, \dots, k; (\text{i.i.d.}) \quad (5)$$

where σ^2 and τ^2 are assumed known, we have

1) The distribution:

$$m(y|\mu) = \frac{1}{[2\pi(\sigma^2+\tau^2)]^{n/2}} \exp\left[-\frac{1}{2(\sigma^2+\tau^2)} \sum_{i=1}^k (y_i - \mu)^2\right]$$

2) The marginal MLE: $\hat{\mu} = \bar{y}$.

3) From (5.2)

$$p(\theta_i|y_i, \hat{\mu}) = N(B\hat{\mu} + (1 - B)y_i, (1 - B)\sigma^2) \quad (5)$$

where, $B = \sigma^2/(\sigma^2 + \tau^2)$. The PEB point estimate of θ_i is

$$\hat{\theta}_i^\mu = \bar{y} + (1 - B)(y_i - \bar{y}).$$

This is a shrinkage estimator.

6 Multilevel Models

Goldstein (1995).

- 1) Schooling systems is an example of **hierarchical structure**, students are grouped within schools, schools are grouped within education boards....
- 2) Multilevel model has advantages:
 - a) Obtain statistically efficient estimate of regression coefficients (fixed effects).
 - b) Provides correct standard errors for significance test, previous result may be wrongly significant using too small standard errors.
- ...

We have education boards, schools, students. How to write down random effects models?

7 References

- ◇ Carlin P. Bradley and Louis A. Thomas (2000), Bayes and Empirical Bayes Methods for Data Analysis. 2nd Edition. Chapman & Hall.
- ◇ Geert Molenberghs and Geert Verbeke (2005), Models for Discrete Longitudinal Data. Springer.
- ◇ Harvey Goldstein (1995), Multilevel Models, 2nd Edition, Edward Arnold.