Linear Mixed Effects Models

Basic idea: Individuals in population are assumed to have their own subject-specific mean response trajectories over time.

- Allow a subset of parameters to vary from one individual to another, thereby accounting for sources of natural heterogeneity in the population.
- Fixed effects: population mean, shared by all the subjects
- Random effects: subject-specific effects
- Appropriate even for unbalanced design

Random Intercept Model

Let Y_{ij} and X_{ij} be the outcomes and covariates for the ith individual, measured at time t_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n_i$. Consider the random intercept model

$$Y_{ij} = \beta_0 + X_{ij}^T \beta + b_i + \epsilon_{ij}$$

- $b_i \sim N(0, \sigma_b^2)$ is the random subject effect
- $\epsilon_{ij} \sim N(0, \sigma^2)$ are within-subject measurement errors
- b_i and ϵ_{ij} are assumed to be independent of each other.

Note the random intercept model describes the mean response trajectory over time for the ith subject, i.e. the *conditional* mean of Y_{ij} given the subject-specific effect b_i:

$$\mathbb{E}(Y_{ij}|b_i) = \beta_0 + X_{ij}^T \beta + b_i$$

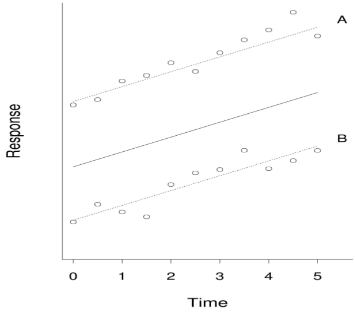
Integrating over b_i, the mean response profile in the population, i.e., *marginal* mean of Y_{ij}:

$$\mathbb{E}(Y_{ij}) = \beta_0 + X_{ij}^T \beta$$

- If X_{ij} is time, β describes patterns of change in the mean response over time in the population of interest.
- b_i represents the ith individual's deviation from the population mean intercept, after the effect of covariates have been accounted for.

Example

Consider a simple linear model $Y_{ij} = (\beta_1 + b_i) + \beta_2 t_{ij} + \epsilon_{ij}$



- Overall mean response over time in the population changes linearly with time
- Subject-specific mean responses for two specific individuals deviate from the population trend

For the random intercept model, the marginal variance of each response is given by

$$\text{var}(Y_{ij}) = \text{var}(b_i + \epsilon_{ij}) = \sigma_b^2 + \sigma^2$$

Similarly, the marginal covariance between any pair of response, Y_{ij} and Y_{ik}, is given by

$$\text{cov}(Y_{ij}, Y_{ik}) = \text{cov}(b_i + \epsilon_{ij}, b_i + \epsilon_{ik}) = \text{cov}(b_i, b_i) = \sigma_b^2$$

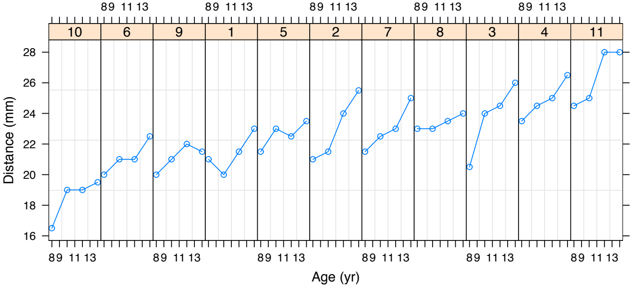
Namely, the introduction of a random intercept, b_i, induces correlation among the repeated measurements in longitudinal data. The induced covariance pattern is

$$\text{cov}(\boldsymbol{Y}_i) = \begin{pmatrix} \sigma_b^2 + \sigma^2 & \sigma_b^2 & \cdots & \sigma_b^2 \\ \sigma_b^2 & \sigma_b^2 + \sigma^2 & \cdots & \sigma_b^2 \\ \cdots & & & \\ \sigma_b^2 & \sigma_b^2 & \cdots & \sigma_b^2 + \sigma^2 \end{pmatrix}$$

Example

Orthodontic Measurements

- Orthodontic measurements were taken from 27 children (16 boys and 11 girls) every two years from age 8 to 14.
- The following table presents the data for 11 girls.
- One question of interest is the prediction of individual *growth curve*.



Random Intercept and Slope Model

Consider the following random intercept and random slope model

$$Y_{ij} = \beta_1 + \beta_2 t_{ij} + b_{1i} + b_{2i} t_{ij} + \epsilon_{ij}$$

- $b_{1i} \sim N(0, g_{11})$ and $b_{2i} \sim N(0, g_{22})$ are the random intercept and random slope with $\text{cov}(b_{1i}, b_{2i}) = g_{12}$.
- $\epsilon_{ij} \sim N(0, \sigma^2)$ are within-subject measurement errors
- $\boldsymbol{b}_i = (b_{1i}, b_{2i})^T$ and ϵ_{ij} are assumed to be independent

- The *conditional* mean of Y_{ij} given the subject-specific effect $\boldsymbol{b}_i$:

$$\mathbb{E}(Y_{ij} \mid \boldsymbol{b}_i) = \beta_1 + \beta_2 t_{ij} + b_{1i} + b_{2i} t_{ij}$$

- The mean response profile in the population, i.e.,*marginal* mean of Y_{ij}:

$$\mathbb{E}(Y_{ij}) = \beta_1 + \beta_2 t_{ij}$$

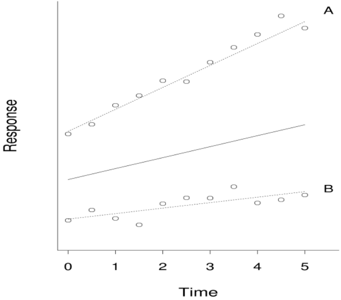
- The covariance of $\boldsymbol{Y}_i$:

$$\text{var}(Y_{ij}) = g_{11} + 2t_{ij}g_{12} + t_{ij}^2 g_{22} + \sigma^2$$

$$\text{cov}(Y_{ij}, Y_{ik}) = g_{11} + (t_{ij} + t_{ik})g_{12} + t_{ij}t_{ik}g_{22}$$

Example

Consider the model $Y_{ij} = \beta_1 + \beta_2 t_{ij} + b_{1i} + b_{2i} t_{ij} + \epsilon_{ij}$



- Subjects vary not only in their baseline level of response, but also in terms of the changes in their response over time.

Mixed Effects Model

A general formula for mixed effects model:

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i, \quad i = 1, \cdots, m$$

$\boldsymbol{Y}_i$: $(n_i \times 1)$ response vector

$\boldsymbol{X}_i$: $(n_i \times p)$ design matrix for fixed effects

$\boldsymbol{\beta}$: $(p \times 1)$ regression coefficients for fixed effects

$\boldsymbol{Z}_i$: $(n_i \times q)$ design matrix for random effects

$\boldsymbol{b}_i$: $(q \times 1)$ random effects

$\boldsymbol{\epsilon}_i$: $(n_i \times 1)$ error vector

Distributional assumptions: $\boldsymbol{b}_i$ and $\boldsymbol{\epsilon}_i$ are independent with

$$\boldsymbol{b}_i \sim N(\mathbf{0}, G), \ \boldsymbol{\epsilon}_i \sim N(\mathbf{0}, \sigma^2 I)$$

- The columns of $\boldsymbol{Z}_i$ are typically a subset of the columns of $\boldsymbol{X}_i$. In particular, $Z_i = \mathbf{1}_i$ corresponds to the random intercept model.
- The marginal mean of $\boldsymbol{Y}_i$ is $\mathbb{E}(\boldsymbol{Y}_i) = \mathbb{E}(\mathbb{E}(\boldsymbol{Y}_i|\boldsymbol{b}_i)) = \boldsymbol{X}_i\boldsymbol{\beta}$. Thus the coefficients $\boldsymbol{\beta}$ has the population mean interpretation.
- The marginal covariance of $\boldsymbol{Y}_i$ is $\boldsymbol{Z}_i G \boldsymbol{Z}_i^T + \sigma^2 \boldsymbol{I}$.
- Unknown parameters $\boldsymbol{\beta}, G, \sigma^2$ can be estimated from ML or REML.
- Inference about the mean and variance parameters

Best Linear Unbiased Predictor

- In many applications, it is of interest to *predict* random effects $\boldsymbol{b}_i$, and furthermore, the predicted response profile for the ith subject.
- Using maximum likelihood, the prediction of $\boldsymbol{b}_i$, say $\hat{\boldsymbol{b}}_i$, is given by

$$\hat{\boldsymbol{b}}_i = \mathbb{E}(\boldsymbol{b}_i | \boldsymbol{Y}_i; \hat{\beta}, \hat{G}, \hat{\sigma}^2)$$

- Under the Gaussian assumption,

$$\hat{\boldsymbol{b}}_i = \hat{G} \boldsymbol{Z}_i^T \hat{\Sigma}_i^{-1} (\boldsymbol{Y}_i - \boldsymbol{X}_i \hat{\beta})$$

  where $\Sigma_i = \text{cov}(\boldsymbol{Y}_i | \boldsymbol{X}_i) = \boldsymbol{Z}_i G \boldsymbol{Z}_i^T + \sigma^2 \boldsymbol{I}$.
- This is known as the "best linear unbiased predictor" (or BLUP)

- When the unknown covariance parameters have been replaced by their ML or REML estimates, the resulting predictor is often referred to as the "Empirical BLUP" or the "Empirical Bayes" (EB) estimator.
- The ith predicted response profile is

$$\hat{\boldsymbol{Y}}_i = \boldsymbol{X}_i\hat{\beta} + \boldsymbol{Z}_i\hat{\boldsymbol{b}}_i = (\hat{\sigma}^2\hat{\Sigma}_i^{-1})\boldsymbol{X}_i\hat{\beta} + (I - \hat{\sigma}^2\hat{\Sigma}_i^{-1})\boldsymbol{Y}_i$$

- In general, BLUP shrinks predictions towards population-averaged mean.