# STAT 574 Linear and Nonlinear Mixed Models

## Lecture 0: Overview

Chenceng Cai

Washington State University

# Course Information

| | |
|---|---|
| Lecture | Tu/Th 1:30 – 2:45 PM @CUE 418 |
| Office Hours | Wed 12 – 3 PM @Neill 405 or by appointment |
| Spring Break | Week 10 (Mar 10 - 14) |
| Contact | ▶ email: chencheng.cai@wsu.edu<br>▶ zoom: https://wsu.zoom.us/my/chenchengcai |
| Site | ▶ Canvas<br>▶ https://chenchengcai.com/teaching/Stat574 |

# Course Requirements

- Prerequisites: STAT 530, STAT 556, and R/Python programming.
- Textbook:
  **Mixed Models: Theory and Applications with R**, 2nd Edition.
  Eugene Demidenko. Wiley 2013.
- Recommended reading:
  **Mixed-Effects Models in S and S-PLUS**.
  José C. Pinheiro and Douglas M. Bates. Springer 2000.

# Assessment

- Homework: 40%
  Around six homework in total.
- Exam: 40%
  Two exams: mid-term and final.
- Project: 20%
  A case study project on extended topics.

# Tentative topics for the final project

- Chapter 11: Statistical Analysis of Shape
- Chapter 12: Statistical Image Analysis
- Hierarchical Bayesian models and variational inference (NLP related)
- Meta-analysis and fusion learning

# Tenative Schedule

- (2 weeks): Course overview and review on linear algebra and regression theory.
- (2 weeks): Linear mixed-effect models.
- (2 weeks): Estimation algorithms.
- (2 weeks): Marginal models.
- (2 weeks): Generalized linear mixed models.
- (2 weeks): Nonlinear mixed-effect models.
- (2 weeks): Model diagnostics.

# From classical regression to mixed-effect models

**Classical regression model**:

$$y_k = \alpha + \beta x_k + \epsilon_k, \quad k = 1, \dots, K,$$

where the $\{\epsilon_k\}$ are independent and identically distributed random variables with zero mean and constant variance $\sigma^2$.

**Mixed-effect model**:

$$y_{ij} = \alpha_i + \beta x_{ij} + \epsilon_{ij}, \quad i = 1, \dots, N, \; j = 1, \dots, n_j,$$

where the double index $(ij)$ means the $j$-th unit in $i$-th cluster. Furthermore, we assume

$$\alpha_i = \alpha + b_i, \quad i = 1, \dots, N,$$

where $\{b_i\}$ are i.i.d. $N(0, \sigma_b^2)$.

# From classical regression to mixed-effect models

**Classical regression model**:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

where $\boldsymbol{\epsilon} \sim N(0, \sigma^2 \boldsymbol{I})$.

**Mixed-effect model**

$$\boldsymbol{y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \epsilon_i, \quad i = 1, \ldots, N,$$

where $\epsilon_i \sim N(0, \sigma^2 \boldsymbol{I})$ for $i = 1, \ldots, N$, and $\boldsymbol{b}_i \sim N(0, \sigma^2 \boldsymbol{D})$ for $i = 1, \ldots, N$.

# Mixed-effect Models

Model hierarchy:

$$\boldsymbol{b}_i \sim N(0, \sigma^2 \boldsymbol{D})$$
$$\boldsymbol{y}_i \mid \boldsymbol{\beta}, \boldsymbol{b}_i \sim N(\boldsymbol{X}_i \boldsymbol{\beta} + \boldsymbol{Z}_i \boldsymbol{b}_i, \sigma^2 \boldsymbol{I})$$

**Mixed-effect Model**

- Known: $\boldsymbol{X}_i$, $\boldsymbol{Z}_i$
- Estimate: $\boldsymbol{\beta}$, $\boldsymbol{b}_i$, $\sigma^2$, $\boldsymbol{D}$.

**Hierarchical Bayesian Model**

- Known: $\boldsymbol{X}_i$, $\boldsymbol{Z}_i$, $\sigma^2 \boldsymbol{D}$ and prior for $\sigma^2$.
- Estimate: $\boldsymbol{\beta}$, $\boldsymbol{b}_i$, $\sigma^2$

# Why mixed-effect models?

## Heterogeneity

No two leaves are alike!

Personalized/individualized inference.

# Extensions

- Nonlinear:
$$\boldsymbol{y}_i = \boldsymbol{f}_i(\boldsymbol{\beta}) + \boldsymbol{Z}_i(\boldsymbol{\beta})\boldsymbol{b}_i + \epsilon_i, \quad i = 1, \ldots, N$$
- Generalized Linear:
$$g(\boldsymbol{y}_i) = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \epsilon_i, \quad i = 1, \ldots, N$$

# From Data to Conclusion

1. Data
2. Model
3. Estimation
   - Close-form
   - Algorithmic
4. Properties
   - Feasibility
   - Bias, Variance
   - Asymptotics
5. Implementation
   - R/Python
   - algorithms
6. Diagnostics
7. Conclusion

# To do a good job, an artisan needs the best tools

- Linear algebra.
- Matrix calculus.
- Regression theory.