Analyzing Correlated Data: Basics of the Linear Mixed Effects Model

John Rice, PhD

What is clustered/correlated data?

- **Correlated observations:** arise when pairs or clusters of observations are related and thus are more similar to each other than to observations outside of that pair/cluster
  - Multiple (longitudinal) measurements on the same subject
  - Observations of multiple items on the same subject (e.g., left and right eye data)
  - Sibling, twins, members of the same household
  - Patients from the same practice or provider
- **Repeated measures** are another kind of correlated data
  - Could also be viewed as multiple measurements on a unit (cluster)
  - For example, standardized test scores from students in the same classroom in same school

Advantages and disadvantages of correlated data:

- Advantages
  - Only longitudinal data gives information on individual patterns of change
  - Longitudinal studies economize on subjects (fewer patients needed for similar power, e.g.)
- Disadvantages
  - More complicated analyses are often necessary
  - Have to deal with missing data
  - Interpretation of results may be more difficult

Simple approaches to analyzing correlated data:

(Not recommended in general)

- Change score as outcome
- Baseline as covariate
- Hybrid

Linear mixed models are the preferred approach to analyzing correlated data:

- Similar framework to classical linear regression methods
  - Assumes normality of residuals (but extensions to other distributions also exist)
  - Regression coefficients have similar interpretations
  - Estimation methods are similar
  - Testing/inference is based on the same statistics (e.g., ratios of coefficient estimates to estimated standard errors)
- Additional specification is required (Need to describe the correlation structure of the data within each cluster/unit)
- **Covariance pattern models** specify a correlation structure for the outcome (e.g., independence, compound symmetry, etc.)
- **Random effects models** specify a set of random coefficients

Analyzing Correlated Data: Basics of the Linear Mixed Effects Model

John Rice, PhD

[Notes]