Longitudinal Data Analysis

Introduction with examples

Chad Pickering

BIOS 244
Department of Biostatistics

April 30, 2019

Presentation overview

1. Introduction
   - Goals of longitudinal analysis
   - Examples of repeated measures and study designs
2. Hypothesis testing
   - 1 group, 2 time points
   - 1 group, 3+ time points
   - 2 groups
3. Examples & models
   - t-test and mixed effects model for 2 group, 2 measures
   - Mixed effects model example for 2 group, 3+ measures
4. Formatting your data

UCLA FIELDING SCHOOL OF PUBLIC HEALTH

What is longitudinal analysis?

- Measuring the **same variable(s)** from the **same subject** over time
- Reasons for performing repeated measurement studies:
  - **Limiting number of subjects** reduces the between-subject variation of measurements
  - **Efficiency.** Decreases the number of subjects needed to be recruited for the study
  - Allows one to **study changes over time**

Goals of longitudinal analysis

- Compare **patterns of change** between groups over time
- **Assess associations** between response and other factors that also change over time
- Learn about meaningful features underlying observed patterns and how they vary across time

Examples of repeated measurements

- **Clinical treatment trial**
  - Measure baseline status (measurement 1)
  - Perform treatment
  - Measure post-treatment status (measurement 2)
- **Oral glucose tolerance test** (screening for type II diabetes)
  - Baseline blood sample is drawn (measurement 1)
  - Subject consumes glucose solution
  - Blood samples are drawn after 30, 60, and 120 minutes (measurements 2, 3, 4)

More examples of repeated measurements

The following represent **time series data** (large number of measurements are performed on a high-frequency set interval):

- **Weather** - measure and record temperature every minute (measurements 1 ... n)
- **Stock price** - measure and record stock price every millisecond (measurements 1 ... n)

Instead, we will focus on the analysis of a **low number** of repeated measurements (e.g. 2-10 measurements).

UCLA FIELDING SCHOOL OF PUBLIC HEALTH

Common study designs

- One group, 2 measurements per subject
- One group, 3+ measurements per subject
- Two groups, 2 measurements per subject (Example 1)
  - Typically used in **randomized placebo controlled (clinical) trials**
- Two groups, 3+ measurements per subject (Example 2)
  - Typically used in **longitudinal cohort studies**

UCLA FIELDING SCHOOL OF PUBLIC HEALTH

- Fortunately, common statistical tests work well in a longitudinal context
- Always use methods that utilize the fact that **there are multiple measurements from the same subject**
- Useful to **calculate differences** to reduce dimensions:
  - $\Delta =$ Later measurement - Baseline measurement

and conduct a hypothesis test like $H_o : \Delta = 0$ vs. $H_a : \Delta \neq 0$

SUBJECT	GROUP	WEIGHT BASELINE	WEIGHT END
1	A	50	56
2	A	61	63
3	A	70	65
4	B	84	80
5	B	46	47

WEIGHT DELTA
6
2
-5
-4
1

UCLA FIELDING SCHOOL OF PUBLIC HEALTH

1 group, 2 time points

- Consider measurements from first time point as block 1, and second time point as block 2
- Use to test differences between blocks:
  - Parametric: **paired t-test** (calculate Δs)
  - Nonparametric: **Wilcoxon signed-rank test** for dependent groups

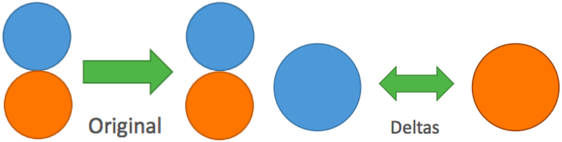
UCLA FIELDING SCHOOL OF PUBLIC HEALTH

1 group, 3+ time points

- Consider measurements from the time points as blocks 1 ... n
- Use to test differences between blocks:
  - Parametric: repeated measurements analysis of variance **(rANOVA/mixed effects model)**
  - Nonparametric: **Friedman test** (rank-based ANOVA analog)
- If there is a significant difference between the time points, continue with pairwise comparison of the baseline time point to every other time point

2 groups, 2 time points

- **Calculate deltas** between the time points:
    - Deltas from the first group is block 1
    - Deltas from the second group is block 2
- Use to test differences between blocks:
    - Parametric: **t-test** (calculate Δs) or **mixed effects model**
    - Nonparametric: **Mann-Whitney U test** (a.k.a. Wilcoxon rank-sum test) for independent groups



UCLA FIELDING SCHOOL OF PUBLIC HEALTH

Mixed effects longitudinal models

Mixed methods are especially useful when:

- There are **3+ time points**
- There are **missing values** or an **unbalanced design**
- Need to adjust for **confounding variables** in multiple time points (instead of adjusting for only baseline value)

Fixed vs. random effects

- **Fixed effects**
  - Group under study and common covariates
  - Usually of interest for interpretation of results
- **Random effects**
  - Used to describe the different "blocks" in the data
  - Usually not of interest for interpretation of the results
  - Study center, family, location, and even subject

2 group, 2 points example

- **Premise: Two diets** are being compared. **50 female subjects** are recruited and randomized to the two diets such that the design is **balanced.** One **body weight** measure is taken **before the diet intervention** and one is taken **one year later.** Is one nutritional intervention better than the other?
- The interest is in **how the diets differ from each other,** not how the individual subjects differ.
  - Diet = fixed effect
  - Subject = random effect

	subj	diet	time	wt
1	1	A	0	190.6076
2	1	A	1	152.2019
3	2	A	0	176.6450
4	2	A	1	144.3953
5	3	A	0	181.4530
6	3	A	1	169.1339

Diet example results

In this case, the random effect makes no difference because the data is **balanced** (same number of data points per subject). Each subject has one Δ, the within-subject differences of weights.

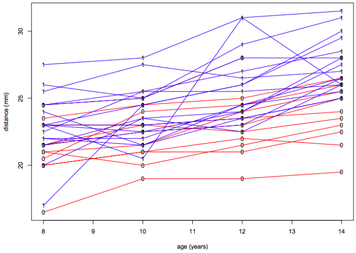
- Use **t-test:**
  - $\bar{Y}_{\Delta_A} = 16.950$, $\bar{Y}_{\Delta_B} = 30.343$, $\bar{Y}_{\Delta_A} - \bar{Y}_{\Delta_B} =$ -13.394
  - t = -2.633, p-value = 0.011
- Or use **mixed effects model** (lowest AIC: random int. & slope):
  - **Model:** $Y_{i,j}^{wt} = \beta_{0,j} + \beta_{diet} X_{diet} + \beta_{time} X_{time} + \beta_{diet \times time} X_{diet} X_{time} + \epsilon_{i,j}$
  - $\hat{\beta}_{diet \times time} =$ -13.394, t = -2.633, p-value = 0.011

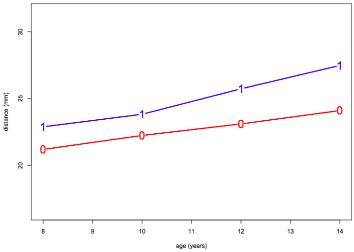
Conclusion: There is a significant difference in the weights of subjects between diets after one year of intervention.

UCLA FIELDING SCHOOL OF PUBLIC HEALTH

2 groups, 3+ points example

Premise: We have a sample of **27 children (16 boys, 11 girls).** On each child, the **distance (mm)** from the center of the pituitary to the pterygomaxillary fissure is measured at **ages 8, 10, 12, and 14.** (Pothoff and Roy (1964)) Does the distance change over time? Is the pattern different for boys and girls?





We will work with the **subject-specific** parameterization, which models individual behavior (with **random intercept** and **random slope**):

$$Y_{ij}(t) = \beta_{0i} + \beta_{1i} t_{ij} + \epsilon_{ij}(t),$$

subject i, time point j.

- $\beta_{0i} + \beta_{1i} t_{ij}$ is the "inherent trajectory" (linear line) of subject i
- $e_{ij}(t)$ is due to
  - fluctuation at t
  - measurement error
- $e_{ij}(t)$ and $e_{ij}(t')$ tend to be in the **same direction** if times are close together $\rightarrow$ **within-subject autocorrelation**

UCLA FIELDING SCHOOL OF PUBLIC HEALTH

- **Random intercept:** $\beta_{0i} = \gamma_{0G}(1 - G_i) + \gamma_{0B} G_i + b_{0i}$ for subject i
- **Random slope:** $\beta_{1i} = \gamma_{1G}(1 - G_i) + \gamma_{1B} G_i + b_{1i}$ for subject i
  - $G_i = 1$ for boys, $G_i = 0$ for girls
  - $(\beta_{0i}, \beta_{1i}) \perp (\beta_{0k}, \beta_{1k}) \quad \forall i \neq k$ pairs
- b's are **mean-zero random effects** that show how much subject i **deviates** from the average intercept/slope γ for a particular sex

Dental example results

$$Y_{ij} = \gamma_{0G}(1 - G_i) + \gamma_{0B}G_i + \gamma_{1G}(1 - G_i)t_{ij} + \gamma_{1B}G_i t_{ij} + b_{0i} + b_{1i}t_{ij} + \epsilon_{ij}$$

- Correlation between ages is relatively constant for all age pairs $\rightarrow$ choose **exchangeable** correlation structure
- Test $H_o : \gamma_{1B} - \gamma_{1G} = 0$
- **Results:** $\hat{\gamma}_{1B} - \hat{\gamma}_{1G} = 2.145$, p-value=0.009
- When **interaction term** is added: $\hat{\beta}_{sex \times age} = 0.305$, p-value=0.026
- **Conclusion:** Boys' distances are greater than girls' on average, and boys' distances also grow faster over time than do girls'.

UCLA FIELDING SCHOOL OF PUBLIC HEALTH

Wide-format

- **Most commonly used** format for storing research data
- **Each subject has a single row**, with multiple columns for each variable
  - Repeated measurement from each time point are in their own column/variables

The preferred format

Long-format is usually preferred for repeated measurements.

- Each subject has a **row for each time point**
- It keeps the measurements in a **single variable** (easier to manipulate, transform, and validate)
- Forces the user to explicitly **store all important information for each time point**, even variables only recorded at baseline)
- Transformation from long-format to wide-format is **much easier!**

UCLA FIELDING SCHOOL OF PUBLIC HEALTH

Resources to get started

- Text: "Modeling Longitudinal Data", Robert E. Weiss
- **Recommended:** lme4/lmerTest or nlme packages in R for mixed effects modeling
- Alternatively, use PROC MIXED in SAS

Why understand the theory?

The syntax of the software is directly tied to the statistical model!

UCLA FIELDING SCHOOL OF PUBLIC HEALTH

Questions?