

General Information

Instructor: Robert Weiss
Office: CHS 51-269
Mail box: CHS 51-254
Office Hours: Monday 1:00
Thursday 2:00 and by appointment
June 6, 7 11:00
E-mail: robweiss at ucla.edu
Book Web Site: <http://faculty.biostat.ucla.edu/robweiss/book>
Course Web Site: <http://faculty.biostat.ucla.edu/robweiss/biostat236>

TA Daniel Conn
Office Hours: Monday 10:00am in CHS A1-279
Office Hours: Thursday 10:00am in CHS A1-279
E-mail: DJCONN17@gmail.com

Class: Monday 11:00–12:50 Room 61-235
Class: Wednesday 11:00–11:50 Room 61-235
Lab Tuesday 11:00–11:50 Room A1-241

Capsule Summary

Longitudinal Data occurs when we repeatedly take the same response measurement (height, weight, blood pressure, test scores, viral particles per cc, pain tolerance, ozone, stock price) over *time* on experimental units (ie people, subjects, rats). The multiple responses within a person are *correlated*, and this correlation must be accommodated in the analysis.

This course will cover the analysis of continuous responses for which a multivariate normal model may be assumed. We will learn how to think about longitudinal data; plot the data and how to specify the mean and variance of the longitudinal response. Advanced topics include introductions to clustered, multivariate and discrete longitudinal data.

Prerequisites

- **Biostat 200a** or other good regression course,
- **Two years** undergraduate mathematics (calculus, linear algebra, multivariable calculus),
- At least one of
 - **Statistical sophistication** as provided by taking Stat 100ABC or Biostat 202A/B or Biostat 215 or other 200-level Stat or Biostat course OR
 - Substantial previous dissertation (or sometimes work) experience with longitudinal data and a strong work ethic.
- **Access** to a computer with SAS software after hours to do homework and projects and for finishing labs.
- Some students should take another course. For example, students (a) without 200-level Biostatistics training (not 201AB, not 406) or (b) with weak mathematical skills should take a different course or take more statistics courses before tackling this course. Students whose prior statistical training consists only of Biostat 100a, 100b, 406 and 201a should generally take Biostat 411 which is designed specifically for those students and I look forward to seeing you in 411.

Text, Notes and Web Resources

Weiss, R. E. (2005). Modeling Longitudinal Data. Springer. Required.

Weiss, R. E. (2013). Modeling Continuous Longitudinal Data. Course lecture notes. Copies of the slides from lecture. Download from the web. *Please bring to class.*

General SAS (version 9.3) information is at <http://support.sas.com/documentation/cdl/en/statug/63962/HTML/default/viewer.htm#titlepage.htm>. Under the Contents tab on the left hand side, click on the + beside procedures. That will open up a menu of all SAS procedures. We will be using MIXED, GLIMMIX, NLMIXED.

The MIXED Procedure documentation is at http://support.sas.com/documentation/cdl/en/statug/63962/HTML/default/viewer.htm#mixed_toc.htm

The GLIMMIX Procedure documentation can be found at http://support.sas.com/documentation/cdl/en/statug/63962/HTML/default/viewer.htm#glimmix_toc.htm

NLMIXED documentation is at http://support.sas.com/documentation/cdl/en/statug/63962/HTML/default/viewer.htm#nlmixed_toc.htm

HPMIXED is a new procedure with interface similar to MIXED but designed for large scale problems. HP stands for “High Performance”.

Examples on the course web page are available. See at the bottom of the course web page. These date back to SAS version 8. We’re currently up to SAS version 9.4.

Garrett M. Fitzmaurice, Nan M. Laird, James H. Ware (2011). Applied Longitudinal Analysis, 2nd edition. Wiley. Great book, extremely well written. Covers the material at a lower mathematical level and lower depth than our text, but covers many more topics. Highly recommended as a second text.

Computing & Stat Packages

1. Lab will meet 9 times, including the first week.
2. We will be using SAS Proc Mixed, Glimmix and NLMixed.
3. You will need regular access to SAS to complete labs, homeworks and projects.

Lab Schedule

Lab 1	1-10	Data Entry & Exploration
Lab 2	1-17	Plots
Lab 3	1-24	Simple Analyses
Lab 4	1-31	Fitting Longitudinal Models
Lab 5	2-7	Covariance Models
Lab 6	2-14	Clustered Data
Lab 7	2-21	Discrete Data
Lab 8	2-28	Bivariate Longitudinal
Lab 9	3-6	Residuals

Grading

Grade will be based upon homework, data analysis project (DAP) and the final data analysis project. No tests. You need to find your own data for the final data analysis project (FDAP). As part of the FDAP, you must turn in an abstract due Feb 13, and you must meet with me by week 9 to discuss your FDAP. The abstract and meeting count as part of the grade for the FDAP. The DAP and FDAP should be written up as reports/paper.

Homework	35%
Assigned Data Analysis Project	25%
Final Data Analysis Project	40%
Attendance is required	

Grading will be on an extremely stringent curve. 95% will be required to get an A. Everyone is able *in principle* to get an A. 90% and above will be a B. The curve may be revised downwards, but won't be revised up.

Homeworks General Homework Guidelines

- Homework that is turned in on time: If there are errors in the homework, you have 7 days after the homework is returned (before 11:59pm) to fix the errors and turn the homework back in for potentially full credit. **Please turn in the previous homework plus the full new homework.**
- Because of this, grading will be rather stringent.
- You may work on homeworks singly or in groups of 2 people. Turn in a single homework, and put both names on the top of the front page.
- Data analysis projects may *not* be worked on as part of a group.
- Turn in neat homework.
- Homework and projects must be typed.
- Do NOT turn in pages of computer output with the numbers circled.
- Convert important results neatly into tables, plots or text.
- Delete all output not needed.
- Completely document the information presented in tables and plots. Define all columns, rows, the model used, and important conclusions. Round results appropriately.
- Use full sentences to state your conclusions, to justify your conclusions, to indicate what is important and why.

Data Analysis Project and Final DAP Grading Guidelines

- Data Analysis Project (DAP) will be graded by the TA. Revisions will be graded by the professor.
- Final data analysis project abstract and meeting with the professor: These may be repeated as necessary until the abstract and meeting are sufficient.
- Final data analysis projects will be graded by the professor. Final DAP turned in on time: if the project is deficient, you will receive feedback and an incomplete. After turning in a revised project you will receive credit appropriate to the revision.
- Why? We are here for me to teach and you to learn. A distant secondary purpose of the course is to grade students appropriately via the grading system. It seems vastly preferable that students learn hence, if students demonstrate that they have not learned material, they will revise the homework, DAP or final DAP and turn in the revised homework/DAP.

Homework Due Dates

Homework is due in class on the day assigned. Class attendance is required.

HW1	Apr 18	Intro
HW2	Apr 25	Simple Analyses
HW3	May 2	Dental Data exploration
Abstract	May 4	For your final data analysis project
HW4	May 9	Dental Data Mean Models
HW5	May 16	Dental Data Covariance Models
DAP	May 25	DAP: Study Report
Meeting	Before or by week 9	with Weiss about your FDAP
FDAP	Jun 8 3pm	FDAP report due in my mailbox in 51-254

Lecture Topics and Course Readings (Chapter in book)

- Introduction to Longitudinal Data (1)
- Plotting (2)
- Simple Analyses (3)
- Critiques of Simple Analyses (4)
- Multivariate Normal Theory (5)
- Tools (6.1–6.5)
- Specifying Covariates (7)

Modeling the Covariance Matrix (8)

Random Effects Models (9)

Topics (time permitting)

Discrete Longitudinal (11)

Bivariate Longitudinal (13)

Missing Data (12)

Residuals (10)

Competencies

item	Upon completion of this course, the student should be able to:	Degree Competencies		
		MS	DrPH	PhD
1	Understand the need for correlated data analyses for longitudinal and hierarchical data and why simplified analyses are inappropriate.	2b 2f 3e 2c	2b 2f 3e 4b	1b 1e 3d 4f
2	Understand, specify and apply longitudinal and hierarchical models as a means of addressing scientific and public health issues, including: <ul style="list-style-type: none"> a Specifying the mean model in longitudinal models. b Specifying the covariance model in longitudinal models. c Specify additional hierarchical levels as needed. d Plot data and residuals from longitudinal models. e Interpreting the assumptions in these models f Interpreting the results from the analyses. 	1b 1f 2b 2c	2b 2f 3e 4b	1b 1e 4f
3	Fit these models in SAS.	1f 1h	2f	1e
4	Present the results from a hierarchical/longitudinal data analysis to biostatisticians, public health professionals and scientists.	1f 1g 1h 2b 2c	2f 2g 3e 4b	1e 1f 4f

Relevant Competencies from the Biostat MS, DrPH, and PhD degrees Competencies Lists.

UCLA Biostatistics MS Competencies

- 1b Formulate a public health question in statistical terms. (1,2)
- 1f Conduct appropriate statistical analyses of study data and interpret the results. (1-4)
- 1g Effectively communicate the assumptions and results of analyses through oral and written communications to the collaborative team. (3,4)
- 1h Use statistical software to answer research questions and communicate the results to other research professionals. (3,4)
- 2b Effectively communicate statistical concepts and reasoning to public health collaborators. (1,2,4)
- 2c Learn to write and disseminate substantive field publications and communicate the statistical portion of the methodology to a substantive field audience. (1,2,4)

UCLA Biostatistics DrPH Competencies

- 2b Formulate a public health question in statistical terms. (1,2)
- 2f Conduct appropriate statistical analyses of study data and interpret the results. (1-4)
- 2g Effectively communicate the assumptions and results of analyses through oral and written communications to the collaborative team. (3,4)
- 3e Learn to write and disseminate substantive field publications and communicate the statistical portion of the methodology to a substantive field audience. (1,2,4)
- 4b Write and present effective and clear reports or publications about the application of statistical methods to health problems. (1,2,4)

UCLA Biostatistics PhD Competencies

- 1b Formulate a public health or scientific question in statistical terms. (1-2)
- 1e Conduct appropriate statistical analyses of study data and interpret the results. (1-4)
- 1f Effectively communicate the assumptions and results of analyses through oral and written communications to the collaborative team. (4)
- 3d Evaluate the statistical properties of new methodologies using numerical, mathematical and philosophical approaches. (1,2)
- 4f Learn to write and publish substantive field publications and communicate the statistical portion of the methodology to a substantive field audience. (1,2,4)