

M19-512 Intermediate Biostatistics for Clinical Research

Course Overview

1. Instructors:

- Jingqin Rosy Luo, Ph.D, Professor of Surgery, Biostatistics and Medicine, jingqinluo@wustl.edu
- Feng Gao, Ph.D, Professor of Surgery, Biostatistics, Medicine, feng@wustl.edu

2. Course TA

Yifei Xu, MS: xyifei@wustl.edu

3. Course time arrangements:

Fall 2: 10/23~12/15/2023 Mondays and Wednesdays 9am~12pm

(WashU calendar: Thanksgiving Break – no classes Nov 22-26 (Wednesday-Sunday))

- 9~10:30 Lecture
- 10:30~12:00 Lab

4. PREREQUISITES: M19-511 or consent by instructor

5. TARGET AUDIENCE: medical students, clinicians, clinical and population health researchers

6. Office hours

- Before/after class
- Email scheduling for zoom or in person meeting

7. Course organization

- (1) Lecture notes
- (2) Lab assignment
- (3) HW: usually due 1 week after assignment (submitted by 11:59pm of the due date, see table below)
- (4) Final Project: data analysis project using your own data or public data
 - Project background and objective
 - hypothesis
 - modeling
 - result presentation
 - result interpretation

8. Class schedule, Syllabus and homework schedule:

| # | Date | Handout # / HW # | Topics | HW Due Date |
|---|------|------------------|--------|-------------|
|---|------|------------------|--------|-------------|

| | | | | |
|----|-------|---------------------------------|---|-------|
| 1 | 10/23 | overview, Note 1, Lab 1 / HW 1 | Course overview Recap on R Intro to R markdown Intro to some course datasets EDA and data visualization | 10/30 |
| 2 | 10/25 | Note 2, Lab 2 / HW 2 | Simple linear regression: data visualization, model estimation by least square, interpretation, prediction | 11/1 |
| 3 | 10/30 | Note 3, Lab 3 / HW 3 | Multiple linear regression | 11/06 |
| 4 | 11/1 | Note 4, Lab 4 / HW 4 | Linear Regression model assumption & diagnostic | 11/08 |
| 5 | 11/06 | Note 5, Lab 5 / HW 5 | Binomial Distribution, MLE, and Simple Logistic Regression | 11/13 |
| 6 | 11/08 | Note 6, Lab 6 / HW 6 | Multiple Logistic Regression | 11/15 |
| 7 | 11/13 | Note 7, Lab 7 / HW 7 | Logistic regression : Goodness of fit assessment | 11/20 |
| 8 | 11/15 | Note 8, Lab 8 / HW 8 | Logistic Regression: Case-Control and Matched Designs | 11/21 |
| 9 | 11/20 | Note 9, Lab 9 / HW 9 | Multinomial Logistic Regression | 11/30 |
| | 11/22 | Thanksgiving break, no class | | |
| 10 | 11/27 | Note 10, Lab 10 / HW 10 | Logistic Regression for Ordinal Outcome | 12/4 |
| 11 | 11/29 | Note 11, Lab 11 / HW 11 | Poisson Regression and Zero-Inflated Poisson Model | 12/6 |
| 12 | 12/4 | Note 12, Lab 12 / HW 12 | Intro to Survival Analysis, KM method | 12/11 |
| 13 | 12/6 | Note 13, Lab 13 / Final project | Survival analysis: Cox Proportional Hazards Model | None |
| 14 | 12/11 | Note 14, Lab 14 / Final project | Cox model assessment and handling | None |
| 15 | 12/13 | Note 15 / Final project | Time varying covariates | 12/20 |

9. Course description & objectives: This course is a continuation of the introductory Biostatistics course (M19-511). The topics include basic statistical concepts and methods for various types of data (continuous, categorical, count, and time-to-event outcome data). Through lectures, R labs, homework assignments, and in-class exam, students will learn the concepts and methods commonly used for analyzing various data types, and will implement those methods using R software.

10. Competency: After completing the course, students should

- (1) understand the basic statistical concepts and methods for various types of data
- (2) be able to frame and address research questions using these concepts and methods
- (3) be able to perform data analyses on these types of data using R software, and
- (4) be able to interpret the results in the context of clinical research.

11. Grading: Your grade will be based on:

- Class participation (10%)
- HW: (60% total)
- In class exam or Final project (30%)

12. Grading Scale

A+: 97-100; A: 93-96; A-: 90-92; B+: 87-89; B: 83-86; B-: 80-82; C+: 77-79; C: 73-76; C-: 70-72

13. ATTENDANCE AND PARTICIPATION

Class attendance is required. More than two unexcused absence from class result in 0 point from class participation.

14. POLICY ON LATE ASSIGNMENTS

Late assignments will result in a deduction of 20% of the assignment for each day late (including weekends) unless prior approval is obtained from the instructor or a compelling situation prevents prior approval (i.e. documented health issues or family emergencies).

15. Course textbooks

- (1) Categorical data analysis by Argenti
- (2) Applied survival analysis using R (free online book)
<https://link.springer.com/book/10.1007/978-3-319-31245-3>

16. Software: The course uses R/Rstudio. Please download from the link

@<https://posit.co/download/rstudio-desktop/> to your laptop and bring your laptop to classes for the computing lab assignments.