CLASS SESSIONS
Thursdays 1:00-3:50 PM, ARB 8th Floor Auditorium

INSTRUCTOR
Qixuan Chen, PhD
Associate Professor of Biostatistics
Email: qc2138@cumc.columbia.edu
Office: 722 West 168th Street, R644

TEACHING ASSISTANTS
Fanyu Cui (fc2754@cumc.columbia.edu)
Yanran Li (yl5465@cumc.columbia.edu)

OFFICE HOURS
Tuesdays 4:00 – 5:00 PM (Room: ARB R638)
Wednesdays 12:00 – 1:00 PM (Room: ARB R638)

COURSE DESCRIPTION

This course will introduce the statistical methods for analyzing censored data, non-normally distributed response data, and repeated measurements data that are commonly encountered in medical and public health research. Topics include estimation and comparison of survival curves, Cox models for survival data, multinomial logistic regression models for nominal outcomes, cumulative logistic regression models for ordinal outcomes, Poisson and negative binomial models for count data, generalized estimating equations, and mixed models. Examples are drawn from the health sciences.

PREREQUISITES

This course is designed for students who have already completed P8100 (Applied regression analysis I) and have some programming experiences of statistical software package SAS.

COURSE LEARNING OBJECTIVES

Students who successfully complete this course will be able to:
- Identify types of censoring in survival data analysis, create Kaplan-Meier survival curves, generate descriptive statistics for survival data, compare survival functions among two or more groups, and use Cox models to analyze survival data.
- Use logistic regression models to analyze binary, nominal, or ordinal response data, and use Poisson and negative binomial regression models to analyze count response data.
- Use generalized estimating equations and mixed models to analyze repeated measures data.
- Write simple SAS programs and read SAS outputs.
- Write data analysis section, create tables and figures, and interpret data analysis results for scientific papers.

TEXTBOOKS

Required textbook: None

Recommended textbooks:

- Allison PD (2010). *Survival Analysis Using SAS: A Practical Guide Second Edition*. SAS Institute. [eBook]

- Hosmer DW, Lemeshow S and May S (2008). *Applied Survival Analysis: Regression Modeling of Time-to-Event Data*. John Wiley & Sons, Inc.: New York. [eBook]

- Dupont WD (2009). *Statistical Modeling for Biomedical Researchers: A Simple Introduction to the Analysis of Complex Data*. Cambridge University Press. [eBook]

- Dobson AJ (2001). *An Introduction to Generalized Linear Models, 2nd Edition*. Chapman & Hall. [eBook]

Recommended lecture notes:

- Davidian M. Applied Longitudinal Data Analysis, Lecture Notes (available in the Canvas).

ASSESSMENT AND GRADING POLICY

The course will consist of lectures, homework assignments, quizzes, midterm project, and final exam. Grading will be based on homework (30%), quizzes (20%), midterm project (20%), final exam (25%), and attendance (5%).

- The homework should be submitted via the Canvas (allow multiple times of submission before the deadline). **No late homework will be accepted!!!** Students are encouraged to work together on homework, although only in the spirit of learning. **There should be no copying of work. The homework with the lowest score will not be counted toward the final grade.**

- Quizzes will be **closed books closed notes** and conducted at the beginning of classes.

- Students will be divided into groups with **3-4 students per group**. Students in each group will work together on the midterm project.

- The final exam will be **closed book closed notes** but allow bringing a one-page **two-sided** cheat sheet.

GRADING

A+   Reserved for highly exceptional achievement.
A   Excellent. Outstanding achievement.
A-   Excellent work, close to outstanding.
B+   Very good. Solid achievement expected of most graduate students.
B   Good. Acceptable achievement.
B-   Acceptable achievement, but below what is generally expected of graduate students.
C+   Fair achievement, above minimally acceptable level.
C   Fair achievement, but only minimally acceptable.
F   Failure. Course usually may not be repeated unless it is a required course.

ADDITIONAL COURSE INFORMATION

We will use SAS. SAS is available for use on PCs in the library and in the Student Learning Center. All course notes and handouts will be posted on the Canvas web site. It is students' responsibility to print out all materials before coming to class. Hard copies of handouts will NOT be provided in class. All students are required to arrive on time. Cell phones must be turned off, set to vibrate/silent. No phone calls are allowed in the class.

MAILMAN SCHOOL POLICIES AND EXPECTATIONS

Academic Integrity
Students are required to adhere to the Mailman School Conduct and Community Standards, which includes the Code of Academic Integrity. Columbia Mailman and Columbia University take academic integrity very seriously. This instructor and course are no different. Should any student be suspected of an academic integrity violation, there will be a report submitted to the Center for Student Success & intervention/Student Conduct. After these offices conduct their process, if a student is found responsible for violating an academic integrity policy (see Standards & Discipline/Academic Violations and Student Honor Code & Professional Guidelines), they will be assigned a grade penalty, with a possible outcome being a 0% on the assignment. *[note to faculty: please update this according to your course specific policy]* Please review the university, school, and course policies, as you are responsible for behaving according to the outlined expectations.

Personal Support
Students sometimes experience life challenges that require additional support and connection to resources. If you are experiencing difficult circumstances, please reach out for help and support. Student Support Services in the Office of Student Affairs is poised to connect with students, provide resource referrals, and provide ongoing, non-clinical support. They are a good place to start if you do not know where to turn.

Disability Access
In order to receive disability-related academic accommodations, students must first be registered with the Office of Disability Services (ODS). Students who have or think they may have a disability are invited to contact ODS for a confidential discussion at 212.854.2388 (V) 212.854.2378 (TTY), or by email at disability@columbia.edu. If you have already registered with ODS, please speak to your instructor to ensure that they have been notified of your recommended accommodations by Meredith Ryer (mr4075@cumc.columbia.edu), Assistant Director of Student Support and Mailman's liaison to the Office of Disability Services.

Bias Incidents
Our community at Columbia University's Mailman School of Public Health is committed to creating an inclusive working, learning, and living environment where all are respected. The occurrence of bias related incidents, involving conduct, speech, or expressions reflecting prejudice are an opportunity for learning and growing as a community.

As part of our efforts to create as inclusive a community as possible, when bias incidents occur at Columbia, we provide an opportunity for those involved to engage in education, advocacy and conversation. In this way, we work to address the incident and minimize the potential for future

occurrences. Our community's tools to address bias include a reporting process and the Bias Incident Resource Team, plus resources within schools and various offices. You can access information about the Bias Reporting Process and FAQs here.

Why Reporting Matters and How to File a Report
Our priority is ensuring that Columbia University is a safe community and workplace where we can learn, live, work and express ourselves. As members of the community, we have a shared responsibility to uphold these standards and report behavior that violates these standards. The reporting options provide Columbia University community members an opportunity to share important information directly with appropriate offices. If you or a member of the community needs support please take the time to complete a report so we may provide support, care, and accountability: https://universitylife.columbia.edu/report

COURSE SCHEDULE

Week 1 – Syllabus & Introduction to Survival Analysis

Sept. 4   Learning Objectives:
    ▪ Syllabus
    ▪ Introduction to survival analysis
    ▪ Definition of time-to-event
    ▪ Censoring vs. truncation
    ▪ Survival function estimator without censoring
    ▪ Survival functions with censoring using Kaplan-Meier method
    ▪ Kaplan-Meier survival curves
    ▪ Greenwood's variance estimator for survival functions
    ▪ Confidence interval estimates for survival functions
    Assignments:
    ▪ HW # 1 is assigned

Week 2 – Quantile and Mean Estimation

Sept. 11   Learning Objectives:
    ▪ Quantile estimation
    ▪ Confidence intervals for quantiles
    ▪ Mean survival time estimation
    ▪ SAS syntax and outputs for survival functions, quantiles, and means
    Assignments:
    ▪ HW # 2 is assigned
    ▪ HW # 1 is due

Week 3 – Comparison of Survival Functions & Introduction to Hazard Functions

Sept. 18   Learning Objectives:
    ▪ Quiz #1
    ▪ Comparison of survival functions using log-rank and Wilcoxon test
    ▪ SAS syntax and outputs for survival functions comparison
    Assignments:
    ▪ HW # 3 is assigned
    ▪ HW # 2 is due

Week 4 – Cox Models

Sept. 25   Learning Objectives:
    ▪ Quiz #2
    ▪ Hazard functions
    ▪ Introduction to the PH models
    ▪ Cox model estimation and interpretation
    ▪ Nominal scale covariates and continuous scale covariates
    Assignments:
    ▪ HW # 4 is assigned
    ▪ HW # 3 is due

Week 5 – Survival Data Analysis Case Study	
Oct. 2	<u>Learning Objectives:</u> ▪ Quiz #3 ▪ Confounders and interaction effects ▪ Cox models comparison ▪ Survival data analysis case study ▪ *Distribute midterm project* <u>Assignments:</u> ▪ HW # 5 is assigned ▪ HW # 4 is due

Week 6 – NO CLASS	
Oct. 9	<u>Assignments:</u> ▪ *Work on* the *midterm project* ▪ HW # 5 is due

Week 7 – Cox Models Case Study & Non-Proportional Hazards	
Oct. 16	<u>Learning Objectives:</u> ▪ Quiz #4 ▪ Survival data analysis case study ▪ Model checking for the PH models ▪ Interactions with time ▪ Stratification <u>Assignments:</u> ▪ *Work on the midterm project*

Week 8 – Time-Dependent Covariate	
Oct. 23	<u>Learning Objectives:</u> ▪ Time-dependent covariates ▪ Review writing sample <u>Assignments:</u> ▪ *Work on the midterm project*

Week 9 – Logistic Regression & Multinomial Logistic Regression	
Oct. 30	<u>Learning Objectives:</u> ▪ Review of logistic regression for binary outcomes ▪ Logistic regression for nominal outcomes <u>Assignments:</u> ▪ *Midterm project report 1 is due* ▪ HW # 6 is assigned

Week 10 – Ordinal Logistic Regression & Poisson Regression	
Nov. 6	<u>Learning Objectives:</u> ▪ Ordinal logistic regression ▪ Poisson regression <u>Assignments:</u> ▪ HW # 6 is due

Week 11 – Poisson Regression & Negative Binomial Regression

Nov. 13   <u>Learning Objectives</u>:
    ▪ Poisson regression
    ▪ Negative Binomial regression
<u>Assignments</u>:
    ▪ HW # 7 is assigned
    ▪ *Midterm project report 2 is due*

Week 12 – GEE

Nov. 20   <u>Learning Objectives</u>:
    ▪ Quiz #5
    ▪ GEE & case studies
<u>Assignments</u>:
    ▪ HW # 8 is assigned
    ▪ HW # 7 is due

Week 13 – Random intercept and slope models

Dec. 4   <u>Learning Objectives</u>:
    ▪ Quiz #6
    ▪ Random intercept model
    ▪ Random slope model
<u>Assignments</u>:
    ▪ HW # 9 is assigned
    ▪ HW # 8 is due

Week 14 – Random effects models

Dec. 11   <u>Learning Objectives</u>:
    ▪ Review HW #9
    ▪ Random effects models

Week 15 – Final Exam
Dec. 18   Final exam