BMI/STAT 741. Survival Analysis: Theory and Methods (Spring 2023)

Goals: This course provides a survey of modern statistical methodology for analysis of censored time-to-event data arising from clinical, epidemiological, sociological, and engineering studies. It focuses on intuitive explanations of statistical theory, such as counting-process martingale, and its application to real-world problems. In doing so we combine methodological discussions with extensive case studies, mostly drawing from health sciences research (sample R/SAS code for case studies will be provided). *The overall emphasis of this course is on the practical aspects of censored-data analysis.*

The course is divided into three parts. The first part focuses on methods for univariate event times, e.g., Kaplan—Meier curve, log-rank test, and Cox proportional hazards model. Building on this foundation, the second part expands the scope to complex outcomes such as recurrent or multivariate events, (semi-)competing risks, joint analysis of survival and longitudinal data, multistate data, composite endpoints, and so forth. Special topics for censored data from cutting-edge research areas such as causal inference and machine learning are discussed in the third part.

After taking the course, students will be able to:

- Understand the features of censored data and their implications in statistical inference
- Choose proper non- and semi-parametric methods for analysis of various types of data
- Understand and check the assumptions needed for estimation and inference
- Implement the inference procedures to solve real-world problems using statistical packages such as R (or SAS)
- Interpret and present the analytic results in a clear and coherent way to answer substantive questions

Prerequisites: Students are expected to have basic knowledge in statistical concepts such as random variables, expectation, variance, and maximum likelihood estimation, and to have taken first courses in statistical hypothesis testing (e.g., *t*-test, ANOVA, etc.) and (generalized) linear regression models. Prior experience with R or SAS will be helpful but is not required.

Time and Location: MW 2:30—3:45pm; HSLC 1220—1222.

Instructor: Lu Mao, PhD
WARF 207A, 610 Walnut St,
Madison, WI 53726
Email:lmao@biostat.wisc.edu
Phone: 608-263-5674

Office Time and Location: T&Th 2--3pm, or by appointment.
Zoom link:
https://uwmadison.zoom.us/j/92228225217?pwd=WkdkSHlGdWtlVk1hdVk1eDU0UGZIUT09
(If you intend to join the meeting, send me an email in advance.)

Teaching Assistant:

Tuo Wang (twang437@wisc.edu); Office hours: W 9—11am, or by appointment.
Zoom link:
https://uwmadison.zoom.us/j/93775833743?pwd=ZGZManorRGRobXdIVENERE1GTWxXUT09
(If you intend to join the meeting, send an email to Tuo in advance.)

Readings:

- [Required] Coursebook
  Applied Survival Analysis: From Univariate to Complex Time-to-Event Outcomes
  To be posted periodically by chapter on Canvas
- [Recommended for methodological reader] Kalbfleisch, J. D. & Prentice, R. L. (2002). *The statistical analysis of failure time data* (2nd Ed). John Wiley & Sons.
- [Recommended for applied reader] Klein, J. P. & Moeschberger, M. L. (2003). *Survival analysis: techniques for censored and truncated data* (2nd Ed). Springer Science & Business Media.
- Fleming, T. R. & Harrington, D. P. (1991). *Counting processes and survival analysis*. John Wiley & Sons.

Course Schedule

1/25	Lecture:	Overview
	Reading:	Syllabus; Chapter 1

Part I: Univariate Time-to-Event Data

1/30	Lecture:	Introduction
	Reading:	Chapter 1
2/1	Lecture:	Mathematical Foundations
	Reading:	Chapter 2
2/6	Lecture:	Nonparametric Estimation of the Survival Curve
	Reading:	Chapter 3
2/8	Lecture:	Comparing Survival Rates between Groups
	Reading:	Chapter 3
2/13	Lecture:	The Cox Proportional Hazards Model – Assumptions and Inference
	Reading:	Chapter 4
2/15	Lecture:	The Cox Proportional Hazards Model – Extensions and Diagnostics
	Reading:	Chapter 4
2/20	Lecture:	Non-Hazard-Based Approaches
	Reading:	Chapter 5
2/22	Lecture:	Study Design and Sample Size Calculation
	Reading:	Chapter 6
2/27	Lecture:	Left Truncation

	Reading:	Chapters 7
3/1	Lecture:	Interval Censoring
	Reading:	Chapter 7

Part II: Complex Time-to-Event Data

3/6	Lecture:	Multivariate Event Times – Conditional (Frailty) Models
	Reading:	Chapter 9
3/8	Lecture:	Multivariate Event Times – Marginal Models
	Reading:	Chapter 9

------ Assignment of midterm project ----

3/20	Lecture*:	Recurrent Event
	Reading:	Chapter 10
3/22	Lecture:	Competing and Semi-competing Risks
	Reading:	Chapter 11

------ Midterm project due ----

3/27	Lecture:	Joint Analysis of Longitudinal and Survival Data
	Reading:	Chapter 12
3/29	Lecture:	Multi-State Models – Introduction
	Reading:	Chapter 13
4/3	Lecture:	Multi-State Models – Cox-Type Markov and Semi-Markov Models
	Reading:	Chapter 13
4/5	Lecture:	Composite Endpoints – Nonparametric Estimation
	Reading:	Chapter 14
4/10	Lecture:	Composite Endpoints – Semiparametric Regression
	Reading:	Chapter 14

Part III: Special Topics

4/12	Lecture:	Causal Inference with Censored Data – The Basics
	Reading:	Chapter 15
4/17	Lecture*:	Causal Inference with Censored Data – IPTW and Standardization
	Reading:	Chapter 15
4/19	Lecture:	Causal Inference with Censored Data – Marginal Structural Models
	Reading:	Chapter 15

4/24	Lecture:	Machine Learning with Censored Data – Variable Selection
	Reading:	Chapter 16
4/26	Lecture:	Machine Learning with Censored Data – Nonlinear Regression
	Reading:	Chapter 16
5/1	Lecture:	Guest Lecture (Dave DeMets or others)
5/3	Discussion:	Recap

*Probably online because of the ENAR meeting.
*Probably online due to the instructor's scheduled visit to UPenn

Homework and Exams: One homework every other week; one mid-term, and one data analysis final project.

Grading: 15% attendance; 35% homework; 20% mid-term; 30% final project.