STAT 5353: Probability and Statistics for Data Science and Bioinformatics

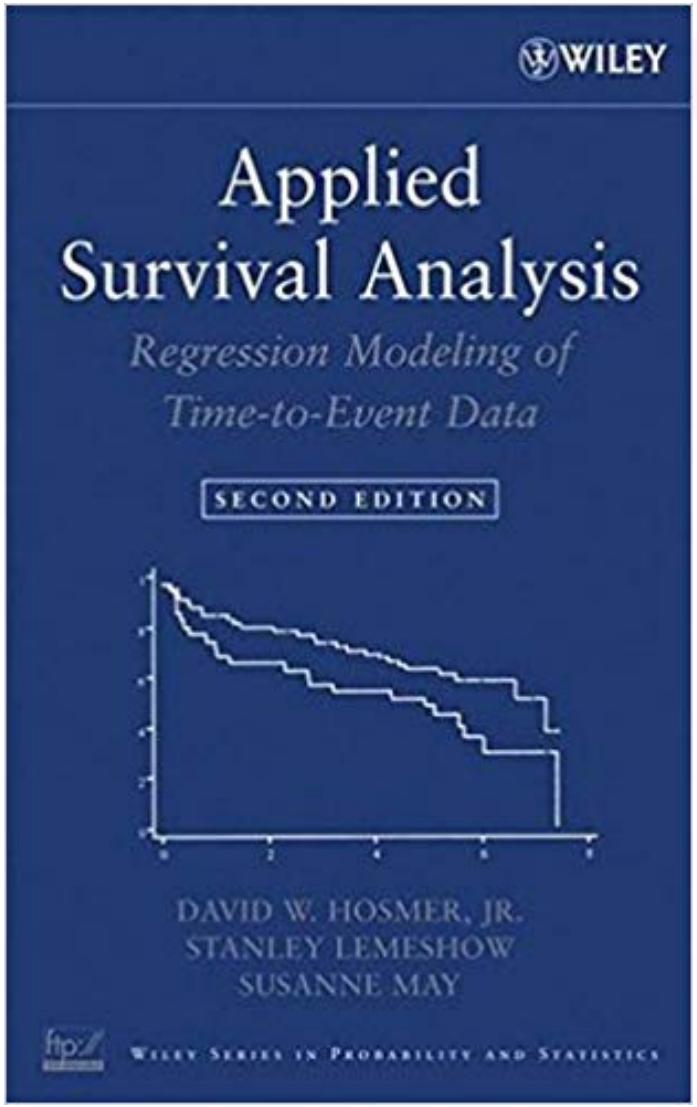
Steven Chiou

Department of Mathematical Sciences,
University of Texas at Dallas

Course information

Instructor	Sy Han (Steven) Chiou
Email address	schiou@utdallas.edu
Office Location	FO 2.410A
Office Hours	Tuesday, Thursday, 12:30 pm - 1:30 pm, or by appointment.
Syllabus	`www.sychiou.com/pdf/fall2018stat6390.pdf`

Required text



- David W. Hosmer, Stanley Lemeshow, and Susanne May
- 2nd edition
- ISBN: 978-0-471-75499-2

Prerequisite

- Calculus through multivariate calculus
- Basic knowledge of regression methods
- Statistical methods for estimation and inferences
- Basic knowledge about `R`
- Basic knowledge about `RMarkdown` and LaTeX
- Basic knowledge about `GitHub`

Course website

- `http://elearning.utdallas.edu/`
- The website will contain
  - Syllabus
  - Lecture notes, both the original and annotated versions
  - `R` scripts
  - Homework and exams
  - Project topics

Grading criteria

Two options

- Traditional option
- Project option

Traditional option

- Homework (50%)
  - Biweekly; 5 $\sim$ 6 assignments
  - Requires `R`
  - Be prepared using *RMarkdown* or *knitr*
  - Due in class or submit via `GitHub`.
- 1 midterm & 1 Final (25% each)
  - Take-home portion that requires `R`
  - One week to complete.

Project option

- General policy
  - A list of potential topics will be posted on both `GitHub` and eLearning.
  - Must inform me of your project topic by November 1.
- Project presentation (50%)
  - 30 minutes
  - last week of the class (week of 12/4)
- Report
  - A *one-page summary* (20%) about the project is due one week before the presentation.
  - *Final report* (30%) is due on the final exam date.

Grading criteria

Grade assignment:

A+: 96 - 100	**A**: 93 – 95.99	**A–**: 90 – 92.99	
B+: 86 - 90	**B**: 83 – 85.99	**B–**: 80 – 82.99	
C+: 76 - 80	**C**: 73 – 75.99	**C–**: 60 – 72.99	**F**: 0 – 59.99.

Tentative course schedule

	Coverage	Topics
Week 1 (8/21)	Chapter 1	Typical censoring and truncation mechanisms
Week 2 (8/28)	Chapter 2	Estimating the survival function (Kaplan-Meier estimator)
Week 3 (9/4)	Chapter 2	Comparison of survival functions (log-rank test)
Week 4 (9/11)	Chapter 2	Other functions of survival time and their estimators (Nelson-Aalen estimator)
Week 5 (9/18)	Chapter 2	Other functions of survival time and their estimators (Nelson-Aalen estimator)
Week 6 (9/25)	Chapter 8	Parametric regression models
Week 7 (10/2)	Chapter 8	Parametric regression models
Week 8 (10/9)	Exam	
Week 9 (10/16)	Chapter 3	Proportional hazards regression model
Week 10 (10/23)	Chapter 3	Proportional hazards regression model
Week 11 (10/30)	Chapter 4	Interpretation of a fitted proportional hazards regression model
Week 12 (11/6)	Chapter 4	Interpretation of a fitted proportional hazards regression model
Week 13 (11/13)	Chapter 5	Model development
Week 14 (11/27)	Chapter 5	Model development
Week 15 (11/4)	Chapter 9	Other models (accelerated failure time model and others)