SCHOOL OF PUBLIC HEALTH
UNIVERSITY OF MINNESOTA

PUBH 7450, SECTION 001

Survival Data Analysis
Fall 2019

COURSE & CONTACT INFORMATION

Credits: 3
Meeting Day(s): Tuesday and Thursday
Meeting Time: 2:30 -- 3:45 pm
Meeting Place: Nils Hasselmo Hall 2-101

Instructor: Haitao Chu
Email: chux0051@umn.edu
Office Phone: 612-625-2138
Fax: 612-626-0660
Office Hours: 1:00pm -- 2:00pm TH with appointments
Office Location: A426 Mayo Building

Teaching Assistants: Roland Brown (brow4288@umn.edu) and Ziyu Ji (jixxx311@umn.edu)
Office Hours: 10:00am -- 11:00am Monday with Roland and 1:00 – 2:00pm Wednesday with Ziyu
TA Office Address: Mayo Building A446

COURSE DESCRIPTION

This course deals with statistical methods analyzing data-to-event data. This is an applied course, though both theoretical and applied materials will be covered. In particular, this is not a course only on how to run SAS/R programs and interpret the outputs. In addition, you are expected to understand underlying statistical theories, modeling assumptions and possibly derivations; you are expected to understand some general statistical principles for survival data analysis including how to deal with censoring, truncation and competing risks. Both SAS and R will be used.

COURSE PREREQUISITES

- Theory of Statistics at the level of Stat 5101-5102 or Stat 8101-8102;
- Regression courses: PUBH 7405 (Linear Regression) and PUBH 7406 (Advanced Regression);
- Some programming experience in both SAS and R;

COURSE GOALS & OBJECTIVES

This course deals with methods analyzing survival times or time-to-event data, which may be censored and/or truncated. The main topics are: 1) estimating a survival curve; 2) comparing two (or more) survival curves; and 3) regression analysis including competing risks analysis.

METHODS OF INSTRUCTION AND WORK EXPECTATIONS

PUBH7450 is a **3** credit course. The University expects that for each credit, you will spend a minimum of three hours per week attending class or comparable online activity, reading, studying, completing assignments, etc. over the course of a 15-week term. Thus, this course requires approximately **135** hours of effort spread over the course of the term in order to earn an **average grade**.

In-class lectures are the main method of instruction. There is a brief break during the 1 hour and 15 minutes' class session. Students are expected to come to each class and encouraged to ask questions and participate in discussions in class, and

read textbooks and finish assignments after class. **Late assignment is not accepted unless with in advance permission from the instructor (see below) or with other legitimate reasons (such as illness).**

Learning Community

School of Public Health courses ask students to discuss frameworks, theory, policy, and more, often in the context of past and current events and policy debates. Many of our courses also ask students to work in teams or discussion groups. We do not come to our courses with identical backgrounds and experiences and building on what we already know about collaborating, listening, and engaging is critical to successful professional, academic, and scientific engagement with topics.

In this course, students are expected to engage with each other in respectful and thoughtful ways.

In group discussion, this can mean:

- Respecting the identities and experiences of your classmates.
- Avoid broad statements and generalizations. Group discussions are another form of academic communication, and responses to instructor questions in a group discussion are evaluated. Apply the same rigor to crafting discussion posts as you would for a paper.
- Consider your tone and language, especially when communicating in text format, as the lack of other cues can lead to misinterpretation.

Like other work in the course, all student to student communication is covered by the Student Conduct Code (https://z.umn.edu/studentconduct).

COURSE TEXT & READINGS

RESERVED IN THE BIO-MEDICAL LIBRARY (DIEHL HALL) AND PROBABLY IN THE BIOSTATISTICS READING ROOM (A460)

- **REQUIRED**: J.P. Klein and M.L. Moeschberger, **Survival Analysis**, 2nd edition. Springer, 2003.
- Optional:
  - P.D. Allison, **Survival Analysis using the SAS System, A Practical Guide**, 2nd edition. SAS Institute Inc., 2010.
  - T.M. Therneau, P.M. Grambsch, **Modeling Survival Data**. Springer, 2000.

COURSE OUTLINE/WEEKLY SCHEDULE

KM for J.P. Klein and M.L. Moeschberger, **Survival Analysis** (2003)

Week	Topic	Readings	Activities/Assignments
Week 1: 9/3, 9/5	**Introduction to survival analysis**	**Lecture notes, KM Chapter 1, 3**	
Week 2: 9/10, 9/12	**Estimating a survival curve**	**Lecture notes, KM Chapter 4-6**	
Week 3: 9/17, 9/19			**Homework 1, 9/17**
Week 4: 9/24, 9/26	**Comparing survival curves**	**Lecture notes, KM Chapter 7**	**Homework 2, 9/26**
Week 5: 10/1, 10/3			**10/3, Midterm Review;**
Week 6: 10/8, 10/10	**Proportional hazards regression**	**Lecture notes, KM Chapter 8, 9, 11**	**Homework 3, 10/8**
Week 7: 10/15, 10/17			**Homework 4, 10/15** **Midterm Exam, 2:30 to 3:45 p.m., 10/17**
Week 8: 10/22, 10/24			
Week 9: 10/29, 10/31			**Homework 5, 10/31**
Week 10: 11/5, 11/7	**Parametric regression models**	**Lecture notes, KM Chapter 12**	
Week 11: 11/12, 11/14	**Other semi-parametric models**	**Lecture notes**	**Homework 6, 11/14**
Week 12: 11/19, 11/21	**Multivariate survival analysis**	**Lecture notes, KM Chapter 13**	
Week 13: 11/26	**Competing risk analysis**	**Lecture notes**	**Homework 7, 11/26**
Week 14: 12/3, 12/5	**Joint survival and longitudinal models**	**Lecture notes**	
Week 15: 12/10			**Homework 8, 12/10** **12/10 Final Review**
Week 16: 12/19	**Final Exam on 10:30 a.m. to 12:30 p.m., Thursday, December 19**		

EVALUATION & GRADING

- Homework assignments will involve some theoretical problems and running SAS or R programs to analyze data.
- I assume that everyone has working knowledge about using SAS and R. I will distribute and put relevant SAS or R programs of discussed examples in the course Moodle website. You are strongly encouraged to try these programs. Larger data sets for homework will be accessible from the course Moodle site, or from http://www.mcw.edu/biostatistics/statisticalresources/Survival-Analysis-Book.htm or in R package KMsurv.
- Please note that it is required to include in your write-up your SAS or R programs, only relevant parts of output, major steps of hand calculations, and necessary interpretations/conclusions.
- Students are allowed to discuss homework problems, however, each one is expected to program and write up independently; **Copying other's work, including computer programs, is cheating or plagiarism, which will lead to an automatic ``F" and possible reporting to the University office**.
- **Each assignment is due at the beginning of class, typically one week after assigned. Hard copy is preferred. Email attachment is allowed if confirmed with instructor beforehand and sent to the instructor before the start of the class. Late homework is counted down 20% for each day of lateness, with the first 20% accruing to homework handed in after class on the due date.**
- With some legitimate reasons (e.g. illness with appropriate documents), you need to notify the instructor in advance or as early as possible to obtain my approval and thus receive no credit deduction.
- **There will be an in-class midterm exam in Week 7 (scheduled on Tuesday October 15, 2:30pm-3:45pm), and a final exam (scheduled on Thursday December 19 10:30 a.m.-12:30 p.m.). An in-class exam will be closed book, but you can bring one page (for mid-term) or two pages (for final exam) of notes (8x11 in size, single sided). For either exam, you need to understand your course notes/textbooks (and homework problems). In particular, exam problems will not be only a subset of homework problems. If you cannot take the midterm exam for some legitimate reasons either unforeseen or with my approval in advance, there will be no make-up exam and your grade will be based on the final exam, suitably pro-rated.**
- **Course evaluation will be based on homework assignments, midterm exam and final exam with weights 40%, 20% and 40% respectively. The final grade is based on a weighted average score of a student's performance in the above three items.**
- S = Achievement that is satisfactory will be expected to complete all assignments and receive a minimum of 65% to receive a passing score.

Please refer to the University's Uniform Grading Policy and Grading Rubric Resource at https://z.umn.edu/gradingpolicy]

Grading Scale

The University uses plus and minus grading on a 4.000 cumulative grade point scale in accordance with the following, and you can expect the grade lines to be drawn **approximately (the actual grade lines are set by some cut points close to some percentiles of the grade distribution)** as follows:

% In Class	Grade	GPA
93 - 100%	A	4.000
90 - 92%	A-	3.667
87 - 89%	B+	3.333
83 - 86%	B	3.000
80 - 82%	B-	2.667
77 - 79%	C+	2.333
73 - 76%	C	2.000
70 - 72%	C-	1.667
67 - 69%	D+	1.333
63 - 66%	D	1.000
< 62%	F	

- A = achievement that is outstanding relative to the level necessary to meet course requirements.
- B = achievement that is significantly above the level necessary to meet course requirements.
- C = achievement that meets the course requirements in every respect.
- D = achievement that is worthy of credit even though it fails to meet fully the course requirements.
- F = failure because work was either (1) completed but at a level of achievement that is not worthy of credit or (2) was not completed and there was no agreement between the instructor and the student that the student would be awarded an I (Incomplete).
- S = achievement that is satisfactory, which is equivalent to a C- or better
- N = achievement that is not satisfactory and signifies that the work was either 1) completed but at a level that is not worthy of credit, or 2) not completed and there was no agreement between the instructor and student that the student would receive an I (Incomplete).

Evaluation/Grading Policy	Evaluation/Grading Policy Description
Scholastic Dishonesty, Plagiarism, Cheating, etc.	You are expected to do your own academic work and cite sources as necessary. Failing to do so is scholastic dishonesty. Scholastic dishonesty means plagiarizing; cheating on assignments or examinations; engaging in unauthorized collaboration on academic work; taking, acquiring, or using test materials without faculty permission; submitting false or incomplete records of academic achievement; acting alone or in cooperation with another to falsify records or to obtain dishonestly grades, honors, awards, or professional endorsement; altering, forging, or misusing a University academic record; or fabricating or falsifying data, research procedures, or data analysis (As defined in the Student Conduct Code). For additional information, please see https://z.umn.edu/dishonesty The Office for Student Conduct and Academic Integrity has compiled a useful list of Frequently Asked Questions pertaining to scholastic dishonesty: https://z.umn.edu/integrity. If you have additional questions, please clarify with your instructor. Your instructor can respond to your specific questions regarding what would constitute scholastic dishonesty in the context of a particular class-e.g., whether collaboration on assignments is permitted, requirements and methods for citing sources, if electronic aids are permitted or prohibited during an exam. Indiana University offers a clear description of plagiarism and an online quiz to check your understanding (http://z.umn.edu/iuplagiarism).
Late Assignments	**Late homework is counted down 20% for each day of lateness, with the first 20% accruing to homework handed in after class on the due date,** unless you have a documented reason for missing a homework assignment which has been pre-arranged with the instructor.
Attendance Requirements	Students are expected to attend lectures, submit the homeworks and finish exams on time.
Extra Credit	Some homeworks contain bonus questions that students can choose to answer and earn extra credit.

CEPH COMPETENCIES

Competency	**Learning Objectives**	**Assessment Strategies*** (*see Assessment Descriptions below this table)
Analyze time-to-event data using appropriate methods and software.	**Time-to-event outcomes** • Appropriately handle censoring and truncation • Estimate survival, hazard and cumulative functions by hand and software, interpret results, and write summary report • Calculate an appropriate confidence interval for a population parameter (e.g., 1-year survival, median survival) from data via their chosen software. • Compare survival functions by hand and software • Identify proper tests and statistics to conduct hypothesis tests to compare survival, hazard or cumulative hazard functions, and carry out the tests via their chosen software. **Regression** • Create exploratory plots to visualize the relations among outcomes and covariates. • Appropriately handle censoring and truncation • Carry out appropriate regression analyses (e.g., PH, AFT, competing risk) via their chosen software. • Create diagnostic plots via their chosen software to assess how well the model fits the data.	• Homework • Exams
Interpret results of data analysis for public health research, policy or practice.	**Descriptive and Graphical Summaries** • Interpret survival, hazard, cumulative hazard functions and graphs • Interpret analysis results from time-to-event regression models	• Homework • Exams