2021
Biostatistics 140.641: Survival analysis

General information

- Instructor: Yuxin Zhu <yzhu15@jhu.edu>
- Teaching Assistants:
  Chunnan Liu <cliu173@jhmi.edu>; Jiyang Wen <jwen22@jhu.edu>
  Chunnan Liu Laboratory: TBA Jiyang Wen Laboratory: TBA
- Lectures (Zhu): Tu/Thur 3:30 -5:00pm, Room TBD
  Office hour by zoom: US EST Friday 9-10 am
  https://jhjhm.zoom.us/j/91481312079?pwd=NFVIdDFnOFFxU1hXbC9oMzhzblkzUT09
- Evaluation based on 3 homework sets, one computing homework and one closed-book inclass exam. Grade is based on homework (50%) and one final exam (50%).
- Course material at intermediate level; proper math/stat background required
- Course material available at CoursePlus website http://courseplus.jhsph.edu

Course Description
The course will introduce fundamental concepts, theory and methods in survival analysis. The emphasis is on statistical tools and model interpretations which are useful in medical follow-up studies and in general time-to-event studies. The content of the course includes hazard functions, survival functions, types of censoring and truncation, Kaplan-Meier estimates, log-rank tests and their generalization. Parametric inference includes likelihood estimation and the exponential, Weibull, log-logistic and other relevant distributions. Statistical methods and theory for the proportional hazard models (Cox model) are discussed in detail, with extensions to time-dependent covariates. Clinical and epidemiological examples included in class presentation and homework illustrate various statistical procedures.

Prerequisite Biostatistics 140.621-4 or 140.651 or equivalent. Calculus I and II. Knowledge of fundamental probability and statistical theory is required.

Requirements There will be 3 regular problems sets (30%), one computing homework (20%) and one online exam (50%) which will be held on 10/21/2021 (Thursday, last class).

Course material is available at CoursePlus website http://courseplus.jhsph.edu

References

*Lee, E. "Statistical Methods for Survival Data Analysis", Wadowrth, 2nd-edition.

*Scheike and Martinussen. Dynamic Regression Models for Survival Data. Springer - Statistics for biology and health.

*Collett, D. "Modelling Survival Data in Medical Research", Chapman and Hall, 1994.

**Kleinbaum, DG (1996) Survival Analysis: A Self Learning Text. Springer.

**Cox, R. and Oakes, D. "Analysis of Suvival Data", Chapman and Hall.

**Kalbfleisch J. D. and Prentice, R. L. "The Statistical Analysis of Failure Time Data", Wiley.

*Hosmer D.W., Lemeshow, S. and May S. Applied Survival Analysis: Regression Modeling of Time to Event Data. Wiley Series in Probability and Statistics.

*: Introductory; **:Advanced

SYLLABUS

Biostatistics 140.641 Survival Analysis

1. Introduction
   - Survival, hazard and cumulative hazard functions
   - Definitions of right-censoring, interval-censoring, left-truncation and right-censoring
   - Fundamental concepts and properties
2. Parametric models
   - Exponential, Weibull, log-normal, log-logistic and Gamma distributions
   - Maximum likelihood estimation
   - Likelihood formulation for censored data
   - Examples
3. One sample estimation
   - Empirical distribution estimation for complete data
   - Right censored data
     a. Likelihood functions
     b. Independent censoring
     c. Kaplan-Meier estimator and its properties
     d. Greenwood's formula
4. Proportional hazards model
   - Assumptions and interpretations

- Partial likelihood and estimation
- Tied survival data
- Models for discrete survival data
- Models for time-dependent covariates

5. PHM and beyond

    - Baseline hazard estimation
    - Model fitting
    - Risk estimation and prediction

6. Hypothesis testing

    - Complete follow-up
      a. Chi-square test
      b. Wilcoxen test
    - Incomplete follow-up
      a. Consideration of so-called nonparametric testing
      b. Log-rank test
      c. Generalized Log-rank test

7. Competing risks model (extra)

    - Cause-specific hazard function
    - Cumulative incidence function (CIF)
    - Cause-specific PHM and CIF regression