BTRY 494/694: Applied Survival Analysis
2008 Spring Semester

Professor: Rob Strawderman
1172 Comstock Hall
rls54@cornell.edu
255-1642

Class Times: T,Th 10:10-11:25, 1187 Comstock

Office Hours: TBA

Grading: Homework and exam(s); exact format TBA and depends on class size

Course website:

`http://www.bscb.cornell.edu/~strawderman/B694.html`

Prerequisites:

BTRY 301/302 or 601/602, BTRY 408/409, or equivalent coursework. Co-registration in BTRY 602 and 409 acceptable, but is not necessarily encouraged.

The classroom will comfortably hold 18-20 students. Because of this and the difficulty of finding another classroom during this time slot, course enrollment will be limited to approximately 20 students. Enrollment preference will therefore be given to undergraduates in the Biometry & Statistics major, MPS students in applied statistics, and graduate students in fields of Statistics and Biometry taking the class for credit.

The homework will involve some computer work and data analysis, as well as some theory and hand calculations. I will focus primarily on *R*, a free package that is very similar to Splus (a package that can also be used). However, SAS code and analyses are also fine.

Important Dates:

- 3/18 & 3/20: No class (Spring Break)
- 5/1: Last class

As per university policy, all syllabi should contain some reference to the **Cornell Code of Academic Integrity**. *Violations are dealt with seriously, so please read it:*

http://cuinfo.cornell.edu/Academic/AIC.html

It goes without saying that this policy is in effect at all times.

References

Required: *Modelling Survival Data in Medical Research, 2nd ed, 2003, David Collett*

Other useful references:

- *Applied Survival Analysis*, DW Hosmer and S Lemeshow (basic)
- *Survival Analysis*, J Klein and M Moeschberger (intermediate)
- *Modeling Survival Data: Extending the Cox Model*, T Therneau and P Grambsch (intermediate-advanced)
- *Statistical Models and Methods for Lifetime Data*, J Lawless (advanced)
- *Statistical Analysis of Failure Time Data*, J Kalbfleisch and R Prentice (advanced)

Format

Didactic. Class lectures will be seminar-style: computer-based presentation with screen projection, the lecture notes being handed out at the start of each class. The upside: we will be able to cover more material in greater depth, and you will also be able to focus on what is being said / presented rather than frantically copying things down. The downside: the pace may be relatively quick at times.

Outline of Topics Covered

Survival analysis deals with the analysis of censored (i.e., incomplete) time-to-event response data. Censored response data arise in many application areas, including actuarial science (e.g., developing insurance and pension schemes); biomedical, epidemiological and public health studies (e.g., clinical trials, disease history and longevity studies); economics (e.g., durations of strikes or unemployment); engineering and industrial applications (e.g., life testing and reliability); finance (e.g., bankruptcy, credit scoring); political science (e.g., duration of regimes, timing of legislation); and sociology (e.g., rates of recidivism).

In addition to the presence of censoring, a distinguishing feature of survival analysis is the focus on estimating the so-called "hazard" and "survivor" functions. We will define these quantities and discuss their utility in analyzing censored time-to-event data. We will develop and discuss several nonparametric methods of estimation (e.g., life table, actuarial, and Kaplan-Meier estimators). Also developed and discussed are (i) methods appropriate for comparing the survival experience between 2 or more groups of subjects and (ii) several important hazard regression models, including the Weibull and Cox regression models. Finally, time permitting, attention may turn to more advanced topics in the field (e.g., grouped time-to-event data, time-dependent covariates, recurrent events, and multivariate survival data). Analyses of real data – primarily drawn from biomedical, epidemiology, and public health applications – will be presented and used for illustration. These analyses will include computer code, mainly in *R* but also frequently in *SAS*.

Rough outline of topics covered (subect to change):

1. Introduction (Ch 1)
2. Nonparametric Methods
    - Estimating survivor functions (§2.1-2.2)
    - Estimating cumulative hazard & hazard functions (§2.3)
    - Estimating percentiles (§2.5-2.6)
    - Comparing 2 or more groups (§2.6, 2.7)
3. Semiparametric Methods
    - Proportional hazards (Cox) regression model (§3.1-3.2)
    - Estimating the regression coefficients (§3.3)
    - Testing and CIs (§3.4)
    - Interpreting output (§3.7)
    - Basic model selection (§3.5-3.6)
    - Estimating the survivor/hazard functions (§3.8)
    - Two sample problem revisited (§3.9)
    - Other semiparametric methods*
4. Parametric Methods
    - Single sample estimation (§5.1-5.3)
    - Comparing two groups (§5.4)
    - Weibull proportional hazards regression (§5.5, 5.6)
    - Accelerated life regression models* (§6.1-6.6)
5. Possibly other topics as time permits

* if time