BIOST/EPI 537
Winter Quarter 2009
Dr. Heagerty
January 6, 2009

**BIOST/EPI 537**
**SURVIVAL DATA ANALYSIS IN EPIDEMIOLOGY**
**WINTER 2009**

| | |
|---|---|
| PREREQUISITES: | BIOST/EPI 536; or permission of instructor |
| HOURS: | Lecture: Tuesday 1:30pm-3:20pm, HSB T-625<br>Lecture: Thursday 1:30pm-3:20pm, HSB T-625<br>Discussion: Tuesday 12:30pm-1:20pm, HSB K-069 |
| INSTRUCTOR: | Patrick Heagerty, PhD<br>Professor<br>Department of Biostatistics<br>F-667 HSB; 206-616-2720<br>e-mail: `heagerty@u.washington.edu`<br>Office Hours: Tuesday 3:30-5pm, or by appointment. |
| TEACHING ASSISTANTS | Sean Devlin<br>Office:<br>e-mail: `devlin@u.washington.edu`<br>Office Hours:<br><br>Tanya Granston<br>Office:<br>e-mail: `granston@u.washington.edu`<br>Office Hours:<br><br>Brandon Guthrie<br>Office:<br>e-mail: `brguth@u.washington.edu`<br>Office Hours: |

| | |
|---|---|
| REQUIRED TEXT: | Hosmer, Lemeshow & May: *Applied Survival Analysis: Regression Modeling of Time to Event Data*, Wiley, 2008. |
| | Lecture notes handed out in class. |
| RECOMMENDED BOOKS: | Breslow & Day: *Statistical Methods in Cancer Research, Volume II: The Design and Analysis of Cohort Studies*, IARC Sci. Pub. #82, Oxford, 1988. |
| | Collett: *Modeling Survival Data in Medical Research*, Chapman & Hall, 2003. |
| | Clayton & Hills: *Statistical Methods in Epidemiology*, Oxford, 1993. |
| EXTRA HANDOUTS: | Extra handouts from class sessions will usually be available at the next lecture. In addition, PDF files of the lecture notes will be posted on the course web page. |
| VIDEOTAPES: | Audio with screen capture will be recorded for each lecture and posted on the course web site. Not all aspects of class will be captured (reliably) with this technology, so students are expected to attend class sessions. |
| COMPUTER SOFTWARE: | We will be using STATA 10 which is available in the Health Sciences Microcomputer Laboratory. Personal copies of STATA are available for UW Health Sciences faculty, students, and staff via the STATA web site. |
| CLASS WEBSITE: | Homework exercises and course information will be available on the website: `http://faculty.washington.edu/heagerty/Courses/b537` |
| DISABILITY: | If you would like to request academic accomodations due to a disability, please contact Disabled Student Services, 448 Schmitz, 543-8924 (V/TDD). If you have a letter from Disabled Student Services indicating you have a disability that requires academic accomodations, please present the letter to me so that we can discuss the accomodations you might need for class. |
| DISCUSSION SECTION: | The discussion section will be used to discuss STATA examples, homework exercises, and outstanding questions that arising from the lectures. |

COURSEWORK: Weekly Exercises
Two Quizes (February 3, 2009 and March 3, 2009)
Midterm Exam (take-home), due February 12, 2006 in class.
Final Exam (exam week).

GRADING: Numerical class grades will be based on the final exam (30%), the midterm exam (40%), and two quizes (30%). In addition, weekly homework questions will be marked "check" or "minus" for whether or not it represents a good faith effort to answer the question. The course grade will be the minimum of the grade based on the average for the two exams and the exercise effort according to the chart below:

| Course Grade | Minimum percent of exercises completed |
|---|---|
| 3.7 | 85 |
| 3.3 | 75 |
| 2.8 | 65 |

BIOST/EPI 537
Winter Quarter 2009
Dr. Heagerty

**LEARNING OBJECTIVES**
**BIOST/EPI 537**
**SURVIVAL DATA ANALYSIS IN EPIDEMIOLOGY**
**WINTER 2009**

Upon entering this course you are expected to have completed courses in introductory statistics or biostatistics, multiple regression, elementary categorical and censored survival data analysis, and multiple logistic regression analysis of data from epidemiologic case-control studies. You should know how to fit multiple linear and logistic regression models, and how to perform hypothesis tests about regression coefficients. You should be familiar with case-control, cross-sectional, and cohort study designs. You should have been exposed to simple statistical methods for analyzing censored survival data, including the Kaplan-Meier estimator and the log-rank test.

After completing this course the student can ordinarily expect to:

1. Estimate survival curves using the Kaplan-Meier estimator.
2. Estimate the (grouped or smoothed) instantaneous mortality or hazard rate.
3. Compare two or more survival curves using a log-rank or related test.
4. Fit appropriate Cox regression models to continuous time data from epidemiologic cohort studies using STATA, and evaluate the fit of these models.
5. Interpret regression coefficients from Cox regression models fit to continuous-time data, and test hypotheses about them.
6. Use time-dependent covariates in the Cox model and interpret the coefficients.
7. Fit appropriate Poisson regression models to grouped data from epidemiologic cohort studies using STATA, and evaluate the fit of these models.
8. Interpret regression coefficients from Poisson regression models fit to grouped data, and test hypotheses about them.

BIOST/EPI 537
Winter Quarter 2009
Dr. Heagerty

**BIOST/EPI 537**
**SURVIVAL DATA ANALYSIS IN EPIDEMIOLOGY**

# Course Outline

| TOPIC (estimated number of lectures), ∗=optional | READING |
|---|---|
| 1. **Introduction to longitudinal studies, survival distributions and event-time data.** (1) | H&L: Ch. 1<br>B&D: Ch. 1 |
| 2. **One, Two, and K-sample Nonparametric methods for censored survival data.** (2) | |
| Product-limit estimator | H&L: 2.2, 2.3 |
| Nelson cumulative hazard estimator | H&L: 2.5; B&D 5.3 |
| Kernal hazard estimator | B&D: 5.3(a) |
| Two and K-sample log-rank test | H&L: 2.4; Collett: 2.5, 2.6 |
| Trend test; Stratified test | H&L: 2.4; Collett: 2.7, 2.8 |
| 3. **Introduction to the Cox regression model** (3) | |
| Forms of the model | H&L: 3.1, 3.2; B&D: 5.1(c) |
| Interpretation of coefficients | H&L: 4.1-4.4; Collett: 3.7 |
| Relationship to K-sample methods | Collett: 3.9 |
| Survival curve and hazard function estimation | H&L: 4.5, 4.6; B&D: 5.3(b)<br>Collett: 3.8 |
| Stratification | H&L: 7.2; B&D: 5.1 |
| 4. **Advanced topics in Cox regression and its application to cohort studies** (6) | |
| Left truncation due to staggered entry | H&L: 7.4; B&D: 5.6(a) |
| Time-dependent covariates | H&L: 7.3; B&D: 5.1(b)<br>B&D: 5.5(g), Ch. 6 |
| Time-dependent strata | B&D: 5.1 |
| Different time scales | B&D: 5.1, 5.6(a),(b) |
| Comparison with population rates (external) | B&D: 5.1(d), 5.5(a),(c),(d) |
| Goodness-of-fit and regression diagnostics | B&D: 4.3, 5.2(c), 5.6(c)<br>Collett: Ch. 5 |
| Competing Risks | |
| Post-stratification | |

# Course Outline
**(continued)**

| | TOPIC (estimated number of lectures) | READING |
|---|---|---|
| 5. | **Sampling strategies / designs** (2) | |
| | Sampling from the risk set | H&L: 9.4; B&D: 5.4 |
| | | B&D: 5.4, 5.5(e) |
| | Case-cohort design | |
| 6. | **Additive hazards regression model** (1*) | |
| | Form of the model | H&L: 9.5 |
| | Interpretation of regression coefficients | B&D: 5.1(c), 5.2(d) |
| | Inference/testing of hypotheses | |
| 7. | **Parametric hazard models / accelerated failure model** (1) | |
| 8. | **Analysis of grouped survival data**(4) | |
| | Introduction to Poisson regression | B&D: 4.2 |
| | Poisson regression | B&D: 4.3-4.5, 4.8, 5.2(a) |
| | Calculation of person years | B&D: 3.1, 3.2 |
| | Comparison with population rates (external standard) and the SMR | B&D: 4.6, 5.5(b) |
| 9. | **Regression models for prediction** (2) | |
| | Defining error of prediction | H&L: 5.4, 6.5 |
| | Bias versus variance | |
| | AIC and BIC as criteria | |
| | Accuracy ideas for survival data | |
| | Validation | |
| 10. | **Regression models for multiple outcomes / recurrent events** (1*) | |
| | Recurrent events | H&L: 9.2 |