Instructor: Joanne M. Garrett, PhD

Class Notes: Copies of the class lecture slides are posted on Blackboard under “Course Documents” and “Lectures”; you may want to print them off to bring to class. You will get the most out of the lectures if you read the notes before class – the pace of the class is fast. Also, the lectures have been updated for this year. It’s difficult to know how much material will be covered in each class. The assigned pages for each lecture are a guess. Sometimes we will cover more or less material than indicated.

Several articles are posted on Blackboard. Most of these are suggested readings, but enhance topics in the lectures and/or make good references.

Suggested Texts: (not necessary for class, but you may want to have later)


Class Location: 132 McNider

Class Schedule:

- Thursday 12:00–2:00
- Friday 10:00–12:00

Course Work:

**Problem Sets** – Assignments 1-5 to be completed and a hardcopy turned in at the end of each problem set session. It’s okay to help each other out, particularly if you are stuck on a concept or a Stata command, but you will learn best if you do as much of the work on your own as possible. Final write-ups should be done independently. There will be no final exam.

**Problem Set Sessions**: These sessions will be used to discuss answers to problem sets, as well as any other questions. The smaller class size should allow time for a lot of discussion. There is 1 problem set on exploratory data analysis, 3 on logistic regression, and 1 on survival analysis.
Syllabus – 2011

Course Overview:

**Thurs 1/13** – PUBH 742 overview; Review of PUBH 741 exam – Bring a copy of your final exam and the exam memo and solutions to class (the memo and solutions are posted on Blackboard for PUBH 742 under “Course documents”)

**Part 1: Sampling and Sample Size** (Part 1 has been moved to the end of the course)

**Part 2: Exploratory Data Analysis**

**Fri 1/14**
**Exploratory I**
Topics: Continuous variables; categorical variables
Lecture notes: pp. 2.1.1–2.1.44; 2.2.1–2.2.14 [slides: Part 2\Cont1 & Cat2]

**Thurs 1/20 – Exploratory II**
Topics: Categorical variables (cont.), influential data points, collinearity; missing data
Lecture notes: pp. 2.2.15–2.2.22; 2.3.1–2.3.29 [slides: Part 2\Cat2 & Misc3]

**Fri 1/21 – Exploratory III**
Topics: Missing data (cont.), confounding
Lecture notes: pp. 2.3.30–2.3.50; 2.4.1–2.4.16 [slides: Part 2\Misc3 & Cont4]

**Thurs 1/27 – Exploratory IV**
Topics: Interaction, summary
Lecture notes: pp. 2.5.1–2.5.26; 2.6.1–2.6.7 [slides: Part 2\Interact5 & Sum6]

**Fri 1/28 – Linear Regression Review**
Topics: Review of multiple linear regression
Lecture notes: pp. 3.1.1–3.1.26 [slides: Part 3\Lin1]

**Thurs 2/3 – Problem set #1 on Exploratory: Exp 1 – Exp 4** [Sections 2.1–2.6]
Part 3: Logistic Regression

Fri 2/4 – Logistic I
Topics: Overview of logistic regression: model, odds ratios, confidence intervals
Lecture notes: pp. 3.2.1–3.2.30 [slides: Part 3\Log2]

Thurs 2/10 – Logistic II
Topics: Odds ratios and confidence intervals with interaction; coding the exposure
Lecture notes: pp. 3.2.31–3.2.55 [slides: Part 3\Log2]

Fri 2/11 – Logistic III
Topics: Coding exposure variable – confidence intervals; Maximum likelihood estimation; Likelihood ratio tests; Modeling strategies
Lecture notes: pp. 3.2.56–3.2.64; 3.3.1–3.3.18; 3.4.1–3.4.24
[slides: Part 3\Log2, Log3, & Log4]

Thurs 2/17 – Logistic IV & V
Topics: Examples – traditional Epidemiologic model; Modeling categorical variables with more than 2 categories (intro)
Lecture notes: pp. 3.4.17–3.4.54; 3.5.1–3.5.14 [slides: Part 3\Log4 & Log5]

Fri 2/18 – Logistic V
Topics: Modeling categorical variables with more than 2 categories (cont.); Ordinal and nominal outcomes (more than 2 categories)
Lecture notes: pp. 3.5.15–3.5.56 [slides: Part 3\Log5]


Fri 2/25 – Logistic VI
Topics: Calculating risk ratios from binomial regression models
Lecture notes: pp. 3.6.1–3.6.34 [slides: Part 3\Log6RR]

Thurs 3/3 – Problem set #3 on Logistic Regression: Log V & VI [3.5 & 3.6]

Logistic VII
Topics: Longitudinal models [GEE]
Lecture notes pp. 3.7.1–3.7.13 [slides: Part 3\Log7GEE]

Fri 3/4 – Logistic VII (cont.)
Topics: Continuation of longitudinal models [GEE]
Lecture notes: pp. 3.7.14–3.7.47 [slides: Part 3\Log7GEE]
Part 3: Logistic Regression (cont.)

Thur 3/10 – Spring Break

Fri 3/11 – Spring Break

Thurs 3/17 – Predictive I
Topics: Predicted risk; Strategy: variable selection, descriptive statistics, specifying starting model, variable reduction, quantifying predictive ability, validation
Lecture notes: pp. 3.8.1–3.8.34 [slides: Part 3\Pred1]

Fri 3/18 – Predictive II
Topics: Example study; Quantifying predictive ability
Lecture notes: pp. 3.8.35–3.8.80 [slides: Part 3\Pred2]

Thurs 3/24 – Predictive III [Skip section 3.8.119-3.8.124]
Topics: Comparing ROC curves; Reliability; Model validation
Lecture notes: pp. 3.8.81–3.8.118 [slides: Part 3\Pred3]

Fri 3/25 – Predictive IV
Topics: Validation (cont.); Strategy for a predictive linear regression model

Thurs 3/31 – Problem set #4 on Predictive Models: Pred 1 – Pred 4 [3.8]

Fri 4/1 – Predictive V [may switch 3/31 and 4/1]
Topics: Clinical applications of predictive models (Bill Miller)
Part 4: Survival Analysis

Thurs 4/7 – Survival I
Topics: Review of survival analysis
Lecture notes: pp. 4.1.1–4.1.51  [slides: Part 4\SrvInt1]
Articles:  Tibshirani R. A plain man’s guide to the proportional hazards model.  
Allison PD. Event History Analysis: Regression for Longitudinal Event 

Fri 4/8 – Survival II
Topics: Cox proportional hazards model; Proportional hazards (PH) assumption
Lecture notes: pp. 4.2.1–4.2.33; 4.3.1–4.3.20  [slides: Part 4\SrvCox2 & SrvPH3]

Thurs 4/14 – Survival III
Topics: PH assumption (cont.); Modeling strategy; Exploratory data analysis
Lecture notes: pp. 4.3.21–4.3.37; 4.4.1–4.4.27 [slides: Part 4\SrvPH3 & SrvExmp4]

Fri 4/15 – Survival IV
Topics: Examples (one exposure, predictive model)
Lecture notes: pp. 4.4.28–4.4.58  [slides: Part 4\SrvExmp4]

Thurs 4/21
1) Survival V
   Topics: Time dependent covariates; Repeated events
   Lecture notes: pp. 4.5.1–4.5.20  [slides: Part 4\SrvMisc5]
2) Problem set #5 on Survival Analysis: Surv 1 – Surv 4  [4.1–4.4]

Fri 4/22 – No Class (good Friday)

Part 1: Sampling and Sample Size

Thurs 4/28 – Sampling I
Topics: Sampling and sample size  [Note: only pp. 1.3.1 – 1.3.42 will be covered]
Lecture notes: pp. 1.1.1–1.3.42  [slides: Part 1\Samp1, Samp2, & Samp3]

Fri 4/29 – Sampling II
Topics: Correcting for survey design (sample weights, etc.)
Lecture notes: pp. 1.4.1–1.4.36  [slides: Part 1\Samp4]
correlation from primary care research to inform study design and analysis.  

Thurs 5/5 – Problems on Sample Size calculations: Section 1.1–1.3