
**CLRES 2020:
BIostatISTICS
STATISTICAL APPROACHES IN
CLINICAL RESEARCH**

Dates: 07/06/11 –08/26/11
Meeting days, time: MWF 10:30–noon
Location: VALE 305A/B
Phone contact: 412-692-4868

Course Instructor(s):
(Joyce) Chung-Chou H. Chang, PhD
Charity G. Moore, PhD

Teaching Fellow(s):
Xinxin Dong, MS
Ming Geng, MS
Zhen Jiang, MS
Yi Ren, MS

Email address(es):
changj@pitt.edu
moorecg@upmc.edu

Email address(es):
xid11@pitt.edu
miq2@pitt.edu
zhj4@pitt.edu
yir3@pitt.edu

Overview and Objectives:

This course focuses on basic concepts and methods of statistics and their application to problems in the health and biomedical sciences. Topics include graphical and numerical descriptions and summarization of data, basic probability theory, probability distributions, point and confidence interval estimation, and hypothesis testing with emphasis on one- and two-sample comparisons involving continuous and categorical data. Correlation, simple linear regression, and analysis of variance will be introduced. Scholars will work with clinical datasets and will learn to analyze the data and explain their findings. This rigorous first course in biostatistics will serve as a prerequisite for other biostatistics courses.

At the completion of the course, the trainees should:

- know the basic statistical procedures used to analyze data;
- be able to apply these techniques utilizing a standard statistical package;
- appreciate the concepts of random variation and bias;
- appreciate the wide range of applications of biostatistical methods to problems in medicine and public health;
- know some specific applications in a specialized area of interest; and
- recognize pitfalls in interpreting biomedical and public health data.

Responsibilities:

- Students will be assigned six homework and eight lab assignments that will be graded. All homework and lab assignments will be assigned with a due date. You are encouraged to work together on homework and lab assignments, but you should write up your results individually. Homework assignments are to be turned in at the beginning of class on the due date.
- Attendance and participation in class, recitation and lab are required.
- Evaluation criteria for this class will be based on completion of the homework and lab assignments, participation and attendance, and the midterm and final exams.

Course Requirements:

30% Best 13 of 16 homework assignments
5% Lab assignments
25% Midterm examination (in class)
30% Final examination (in class)

10% Class participation

Attendance Policy:

- Students are expected to sign-in to each class, lab and recitation (computer provided in suite lobby). If a problem is encountered with the sign-in system, please contact the course instructor(s) as well as Lauren Talotta (talottals@upmc.edu) immediately.

NOTE: Homework assignments, course information, and communication will be available at <http://courseweb.pitt.edu>.

Required Textbook(s):

Rosner B (2010). *Fundamentals of Biostatistics, 7th Edition*. Duxbury Press, Pacific Grove, CA.

Supplemental Textbook(s):

- Rosner B (2000). *Study Guide for Fundamentals of Biostatistics*. Duxbury Press, Pacific Grove, CA.
- Glover TJ and Mitchell KJ (2008). *An Introduction to Biostatistics, 2nd Edition*. Waveland Press Inc, Long Grove, IL.
- Van Belle G, Fisher LD, Heagerty PJ, and Lumley TS (2004). *Biostatistics: A Methodology for the Health Sciences, 2nd Edition*. Wiley, New York.
- Daniel WW (2004). *Biostatistics: A Foundation for Analysis in the Health Sciences, 8th Edition*. Wiley, New York.
- Everitt BS (2009). *Modern Medical Statistics: A Practical Guide*. Wiley, New York.

Computer package:

Stata 11, Stata Press, College Station, Texas (www.stata.com)

We recommend that students purchase or have ready access to this program. It can be purchased through www.stata.com. University of Pittsburgh does have a Stata “Grand Plan” available so the cost is reduced to students in this course. Stata is available at the course computer labs (Stata 11), at the Posvar Hall (Stata 11), and at the Falk library in Scaife Hall (Stata 11). For details, please click **Software Info** in the manual box on the left.

Highly recommended books for Stat component of the course

- Hamilton LC. *Statistics with Stata (Updated for Version 10)*. Duxbury Press, Pacific Grove, CA

Software information:

You will need to have access to a computer, a printer and specific statistical software to do the homework in the statistics courses. There are two statistics programs needed to do your homework in the statistics course.

StaTable

StaTable provides interactive way to calculate probability values and/or critical values for the twenty-five most commonly used statistical distributions. The software can be downloaded from the web with no additional charge.

How to download StaTable from the internet

Connect to the Internet. Go to www.cytel.com/. Select **Products** on the menu bar and choose **StaTable**. Follow the instructions. On the right hand side of the new screen, you can choose **StaTable for Windows Free Download** or **StaTable for Palm Free Download**.

Stata 11

Stata 11 can be accessed from our course computing lab, the University computer labs at Posvar Hall or in Falk Library at Scaife Hall. Stata 11 can also be purchased via GradPlans. In the past we have recommended that students purchase Intercooled Stata plus the Getting Started manual (\$155) and the three-volume *Base Reference Manual* plus *Data Management Reference Manual*, *User's Guide*, and *Quick Reference and Index* (\$179). This is the best if you are planning to do multivariable modeling or intensive analyses when you are finished with this course. If your plans are for small studies and the statistics modules in the program, student version of Intercooled Stata 11 and *Getting Started* manual with one-year license (\$95) or student version of Small Stata 11 and *Getting Started* manual with one-year license (\$48) will be sufficient. If you are buying only Small Stata and the Getting Started Manual, you will need to purchase the *Statistics with Stata (Updated for Version 9)* textbook as a reference. If you have decided to purchase Stata it is best to have it in your computer by January 1. (See instructions below).

Important numbers/locations for Stata GradPlans at Pitt

How to order Stata

Go to <http://www.stata.com/info/order/new/edu/gradplans/gp3-order.html>. You can purchase Stata with a purchase order, a university p-card or your credit card. In some cases you may be able to get your department/ grant to pay for software. Ms. Lauren Talotta can also help you order the software. Contact her (talottals@upmc.edu or 412-586-9629) if you need assistance.

Stata is also available for Mac and Linux machines. Make sure you specify the platform you are using when you order the programs.

Academic Integrity:

Students in this course will be expected to comply with the [University of Pittsburgh's Policy on Academic Integrity](#). Any student suspected of violating this obligation for any reason during the semester will be required to participate in the procedural process, initiated at the instructor

level, as outlined in the University Guidelines on Academic Integrity. This may include, but is not limited to, the confiscation of the examination of any individual suspected of violating University Policy. Furthermore, no student may bring any unauthorized materials to an exam, including dictionaries and programmable calculators.

Disabilities:

If you have a disability that requires special testing accommodations or other classroom modifications, you need to notify both the instructors and the [Disability Resources and Services](#) no later than the 2nd week of the term. You may be asked to provide documentation of your disability to determine the appropriateness of accommodations. To notify Disability Resources and Services, call 412-648-7890 (Voice) to schedule an appointment. The Office is located in 140 William Pitt Union.

Copyright notice:

These materials may be protected by copyright. United States copyright law, 17 USC section 101, et seq., in addition to University policy and procedures, prohibit unauthorized duplication or retransmission of course materials. See [Library of Congress Copyright Office](#) and the [University Copyright Policy](#).

Other links:

- CLRES 2005: Computer Based Data Analyses
- [Institute for Clinical Research Education](#)
- [University of Pittsburgh](#)

Course Schedule

Date: July 6, 2011

Session 1: Introduction to Biostatistics

Chang

At the conclusion of this lecture, the student will be able to:

1. Understand basic concepts of statistics
2. Describe data using numerical and graphical summaries

Topics:

1. Course overview
2. Introduction to basic concepts of statistics
3. Descriptive statistics

Required Reading:

- Rosner's Book Chapter 1-2

Homework assignment 1

- Due: Friday, July 8, 2011.

Date: July 8, 2011

Session 2: Probability Concept I

Chang

At the conclusion of this lecture, the student will be able to:

1. Understand basic concepts of probability
2. Use probability rules/laws to solve problems

Topics:

1. Definitions
2. Probability Rules
3. The addition law of probability
4. The multiplication law of probability

Required Reading:

- Rosner's Book Chapter 3

Homework assignment 2

- Due: Wednesday, July 13, 2011.

Due Today: Homework assignment 1

Date: July 11, 2011

At the conclusion of this lecture, the student will be able to:

1. Understand basic concepts of probability and screening test
2. Use probability rules/laws to solve problems

Topics:

1. Conditional Probability
2. Total Probability Rule
3. Bayes' Rule and Screening Tests

Required Reading:

- Rosner's Book Chapter 3

Homework assignment 3

- Due: Wednesday, July 13, 2011.

Date: July 13, 2011**At the conclusion of this lecture, the student will be able to:**

1. Understand concepts of random variable and its distribution
2. Calculate the mean and variance of binomial and poisson distribution
3. Familiar with PMF and CDF of binomial and poisson distribution
4. Distinguish between binomial and poisson distribution

Topics:

1. Definition of a Random Variable
2. The Probability Mass Function (Probability Function; Distribution)
3. The Expected value of a discrete random variable
4. The Variance of a discrete random variable
5. The Cumulative Distribution Function
6. The Binomial Distribution
7. The Poisson Distribution

Required Reading:

- Rosner's Book Chapter 4

Homework assignment 4

- Due: Friday, July 15, 2011.

Due Today: Homework assignment 2 & 3

Date: July 15, 2011

Session 5: Continuous Probability Distributions

Chang

At the conclusion of this lecture, the student will be able to:

1. Calculate the expected value and variance of continuous random variable
2. Familiar with PDF and CDF of normal distribution
3. Use normal tables to solve problems

Topics:

1. Definition of Random Variables (Review)
2. Probability Density Function
3. The Cumulative Distribution Function
4. The Expected Value and the Variance of a Continuous Random Variable
5. The Normal Distribution
6. Using Normal Tables

Required Reading:

1. Rosner's Book Chapter 5

Homework assignment 5

- Due: Wednesday, July 20, 2011.

Due Today: Homework assignment 4

Date: July 18, 2011

Session 6: Estimation I & Sampling Distributions

Chang

At the conclusion of this lecture, the student will be able to:

1. Transfer any normal distribution to standard normal distribution
2. Calculate probability of discrete distribution using normal approximation
3. Understand concepts and theorems of sampling distributions

Topics:

1. Properties of normal distribution
2. Conversion from $N(\mu, \sigma^2)$ to $N(0, 1)$
3. Normal approximations
4. The Central Limit Theorem
5. The Law of Large Numbers

Required Reading:

- Rosner's Book Chapter 6

Homework assignment 6

- Due: Wednesday, July 20, 2011.

At the conclusion of this lecture, the student will be able to:

1. Understand concepts of sampling distribution
2. Estimate the population mean and variance from a sample

Topics:

1. Sampling distribution of an estimator
2. The central limit theorem
3. Sampling distribution of the sample mean
4. Point and interval estimation of population mean
5. Sampling distribution of the sample variance
6. Point and interval estimation of population variance

Required Reading:

- Rosner's Book Chapter 6

Homework assignment 7

- Due: Friday, July 22, 2011.

Due Today: Homework assignment 5 & 6

At the conclusion of this lecture, the student will be able to:

1. Set up null and alternative hypotheses for significance tests
2. Interpret p values
3. Understand type I and type II error
4. Make inference based on one-sample t test

Topics:

1. Significance tests-the general concepts
2. Steps in hypothesis testing
3. Type I and Type II error
4. Test statistics and p value
5. One-sided and two-sided One-sample t tests

Required Reading:

- Rosner's Book Chapter 7

Homework assignment 8

- Due: Wednesday, July 27, 2011.

Due Today: Homework assignment 7

Date: July 25, 2011

Session 9: Hypothesis Testing : Two-Sample Inference

Chang

At the conclusion of this lecture, the student will be able to:

1. Make inference based on two-sample tests
2. Understand the relationship between hypothesis testing and confidence interval

Topics:

1. Review concepts of hypothesis testing
2. Two-sample F tests
3. Two-sample t tests
4. Paired t tests
5. Hypothesis testing vs. confidence interval

Required Reading:

- Rosner's Book Chapter 8

Homework assignment 9

- Due: Wednesday, July 27, 2011.

Date: July 27 & 29, 2011

Session 10 & 11: Midterm review

Chang

At the conclusion of this lecture, the student will be able to:

1. Familiar with descriptive statistics
2. Probability rules and screening test
3. Discrete distribution (binomial and poisson)
4. Continuous distribution (normal distribution)

Topics:

Review lectures 1-6

Required Reading:

- Rosner's Book Chapter 1-6

Date: August 1, 2011

Session 12: Midterm exam

Chang

Topics:

Midterm exam

At the conclusion of this lecture, the student will be able to:

1. Estimate power and sample size for one and two sample Z test

Topics:

1. Review type I, type II error and power
2. Power estimation for one-sample Z test
3. Factors affecting power
4. Sample size estimation for one-sample Z test
5. Power and sample size estimation for two-sample Z test
6. Sample-size estimation for desired margin of error

Required Reading:

- Rosner's Book Chapter 7.5, 7.6, and 8.10

Homework assignment 10

- Due: Friday, August 5, 2011.

At the conclusion of this lecture, the student will be able to:

1. Test equal proportions for one or two groups of binary data
2. Calculate confidence interval for rate with Poisson distribution

Topics:

1. One and two sample tests of proportions and CI for Binomial distribution
2. One sample confidence interval of rate for Poisson distribution

Required Reading:

- Rosner's Book Chapter 10

Homework assignment 11

- Due: Wednesday, August 10, 2011.

Due Today: Homework assignment 10

At the conclusion of this lecture, the student will be able to:

1. Analysis data from contingency tables
2. Analysis matched binary data
3. Measure agreement between two raters

Topics:

1. Power and sample size estimations for proportion tests
2. Chi-square and Fisher's exact test for contingency tables
3. McNemar's test for matched binary data
4. Kappa statistic for rater agreement

Required Reading:

- Rosner's Book Chapter 10

Homework assignment 12

- Due: Wednesday, August 10, 2011.

Date: August 10, 2011

Session 16: One-way Analysis of Variance

Chang

At the conclusion of this lecture, the student will be able to:

1. Test equal means among more than 2 groups using one-way ANOVA
2. Check assumptions for ANOVA
3. Adjust for multiple comparisons

Topics:

1. Basic idea of analysis of variance
2. The hypotheses and assumptions
3. ANOVA table
4. Bartlett's test of homogeneity of variance
5. Post-hoc procedure (Multiple comparisons)

Required Reading:

- Rosner's Book Chapter 12

Homework assignment 13

- Due: Friday, August 12, 2011.

Due Today: Homework assignment 11 & 12

Date: August 12, 2011

Session 17: Correlation and simple linear regression I

Chang

At the conclusion of this lecture, the student will be able to:

1. Measure association between two continuous variables
2. Fit and interpret simple linear regression models

Topics:

1. Correlation coefficients
2. Simple linear regression
3. Parameter estimation
4. Hypothesis testing

Required Reading:

- Rosner's Book Chapter 11

Homework assignment 14

- Due: Wednesday, August 17, 2011.

Due Today: Homework assignment 13

Date: August 15, 2011

Session 18: Correlation and simple linear regression II

Chang

At the conclusion of this lecture, the student will be able to:

1. Calculate interval estimates for simple linear regression models
2. Check model assumptions

Topics:

1. Interval estimations for coefficients
2. Confidence interval for a mean response
3. Prediction interval for a future observation
4. Regression diagnostics

Required Reading:

- Rosner's Book Chapter 11

Homework assignment 15

- Due: Wednesday, August 17, 2011.

Date: August 17, 2011

Session 19: Nonparametric statistics

Chang

At the conclusion of this lecture, the student will be able to:

1. Understand when nonparametric tests were necessary
2. Carry out proper nonparametric tests when necessary

Topics:

1. Overview (parametric vs. nonparametric)
2. The sign test
3. The Wilcoxon sign-rank test
4. The Wilcoxon rank-sum test
5. The Kruskal-Wallis test
6. Spearman rank correlation coefficient

Required Reading:

- Rosner's Book Chapter 9, 11.12, and 12.7

Homework assignment 16

- Due: Friday, August 19, 2011.

Due Today: Homework assignment 14 & 15

Date: August 19, 2011	Session 20: Logistic Regression Analysis Overview	TBA
------------------------------	--	------------

At the conclusion of this lecture, the student will be able to:

1. Understand the basic ideas of logistic regression

Topics:

1. Basic concepts of logistic regression
2. The logit transformation
3. Estimation technique
4. Interpretation
5. Model adequacy checking
6. Multinomial logistic regression

Due Today: Homework assignment 16

Date: August 22, 2011	Session 21: Survival & Correlated Data Analysis Overview	TBA
------------------------------	---	------------

At the conclusion of this lecture, the student will be able to:

1. Understand the basic ideas of survival analysis and correlated data analysis

Topics:

1. Nature of problems for survival analysis
2. Survival or life expectancy
3. Kaplan-Meier curves and actuarial survival
4. Log Rank statistics
5. Hazard of death
6. Parametric survival regression models
7. Cox proportional hazards models
8. Methods for correlated data analysis
9. Incomplete data

Date: August 23 & August 24, 2011
Session 22 & 23: Final exam review

Chang

At the conclusion of this lecture, the student will be able to:

1. One and two sample inference
2. Analysis of variance
3. Nonparametric tests
4. Correlation and simple linear regression
5. Power and sample size calculation
6. Categorical data analysis

Topics:

Review lectures 10-17

Required Reading:

- Rosner's Book Chapter 6-13

Date: August 26, 2011
Session 24: Final exam

Chang

Topics:

Final exam