

Biostatistics 524
Biostatistics for Clinical Researchers

Clinical Research Design and Statistical Analysis (CRDSA)
OJOC Cohort 17
October 2015 – July 2016
Course Information & Syllabus

INSTRUCTOR: Dr. Lili Zhao
Department of Biostatistics
Office: M2541 SPH II
Tel: (734) 763-6898
Email: zhaolili@umich.edu

Dr. Mousumi Banerjee
Department of Biostatistics
Office: M4150 SPH II
Tel: (734) 764-5451
Email: mousumib@umich.edu

GSI: Wenjia Wag
Department of Biostatistics
Email: icywang@umich.edu

CLASS TIME: OJOC weekends

LOCATION: 1655 SPH1

COURSE DESCRIPTION: Biostat 524 is an introductory course in biostatistics. Topics to be covered include: (1) descriptive statistics; (2) basic probability theory and statistical distributions; (3) overview of study designs; (4) fundamental concepts of statistical estimation, and hypothesis testing; (5) simple and multiple linear regression; (6) analysis of variance (ANOVA) and covariance (ANCOVA); and (7) nonparametric approaches. Statistical concepts, and the specification and interpretation of statistical models will be emphasized and illustrated with real data sets. The SAS computer program will be used for performing analyses of real data sets.

REQUIRED TEXTBOOK:

David S. Moore, George P. McCabe, and Bruce A. Craig. *Introduction to the Practice of Statistics, 7th Edition*. New York: W.H. Freeman & Company.

This new edition of a popular and well-respected text covers the early course materials in

great detail, but will need to be supplemented for some of the later materials in the course. Some of the course material and examples may be collected from other books and peer-reviewed articles in order to provide you with a comprehensive lecture note. For some of the later topics, supplementary materials from other books and articles will be assigned as required readings. In such instances, both the lecture notes and required reading materials will be required to complete the homework assignments and exams. The sequencing of the materials covered in the course might differ from that in the textbook.

REFERENCE BOOKS (not required):

1. Rosner, B. Fundamentals of Biostatistics; Duxbury Thomson Learning
2. Motulsky H. Intuitive Biostatistics; Oxford University Press.
3. Colton T. Statistics in Medicine; Little, Brown and Company, Boston.

COMPUTING: Statistical software used in the course is SAS for Windows.

CLASS WEBPAGE: Available on <https://ctools.umich.edu/> All lecture notes, handouts, required reading materials, homework assignments, their solutions, and various other help materials for the course will be posted on the course website on a regular basis throughout the semester. Although I will send e-mail notifications for lecture note and homework uploads, you will be responsible for checking the ctools site periodically so that you do not miss important materials/announcements. Please bring your own downloaded copy of the lecture notes to class.

HOMEWORKS: There will be monthly homework assignments covering theoretical material, applications, data analysis using statistical software, and interpretation of results. Homeworks will be typically due one month after they are assigned unless otherwise noted. Your answers should be submitted on standard size paper, with the sheets stapled together provided there is more than one page. Computer outputs should be attached where necessary. An un-annotated computer output will not constitute an acceptable answer unless it is entirely self-explanatory, which is seldom the case. It is advisable to present a judicious integration of written answers and computer printouts by a “cut-and-paste” job if necessary.

Please make arrangements to submit your assignment if you are unable to attend a class on the date of submission.

COURSE ORGANIZATION: Monthly sessions will include formal lectures, informal dialogue/discussion, homeworks, computer exercises for data analysis using SAS. The objective of homework is to point to areas that need more discussion. Discussion is particularly encouraged since CRDSA students often have many practical experiences and may contribute towards better understanding of issues and concerns.

COURSE EVALUATION:

Homeworks: 30%

Exam 1 (January, in class): 20%

Exam 2 (April, in class): 20%

Final (Take home, assigned in June, due in July): 30%

Your course grade will be based on the weighting scheme presented above. There is no fixed grading scale for this course; conversion from your cumulative score to letter grades will be carried out at the end of the course.

OUTLINE OF TOPICS:

Descriptive statistics, Basic probability theory and statistical distributions, Fundamental concepts of statistical inference

The nature of statistics, scales of measurement (*MMC* 1.1); descriptive statistics, measures of location, spread, shape (*MMC* 1.1, 1.2); population distributions, normal (*MMC* 1.3); random variables, rules of probability, probability models (*MMC* 4.1-4.5); sampling distribution of means and Central Limit Theorem (*MMC* 5.1), sampling distribution of proportions (*MMC* 5.2); confidence interval for a population mean based on normal distribution (*MMC* 6.1)

Study Designs, Hypothesis Testing, Comparisons of Groups

Randomized vs observational studies, designs for comparing groups (*MMC* 3.1); hypothesis testing (*MMC* 6.2-6.4); Small-sample tests and confidence intervals for a mean based on the t distribution (*MMC* 7.1), confidence intervals for differences in means: paired design (*MMC* 7.1), and independent groups design (*MMC* 7.2); binomial distribution, inference for a population proportion (*MMC* 8.1), comparing two proportions (*MMC* 8.2), chi-square test for two-way tables (*MMC* 9.1)

Nonparametric methods

Nonparametric tests: sign test, rank sum, signed rank (*MMC* 15.1, 15.2)

Simple Linear Regression

Describing relationships: association, correlation and regression (*MMC* 2); Simple linear regression, inference and model checks (*MMC* 10)

Multiple Linear Regression

Multiple regression (*MMC* 11), confounding, interaction, interpretation of observational data, polynomial regression, model assessment, dummy variable regression (*MMC* 11)

Analysis of Variance and Covariance

Multiple Linear Regression (contd.), Analysis of variance and covariance (*MMC* 12, 13).

Introduction to Bootstrap

Bootstrap and resampling methods (*MMC* 16)

CORE COMPETENCIES (BIOSTATISTICS)

1. Describe the roles biostatistics serves in the discipline of public health.
2. Describe basic concepts of probability, random variation, and commonly used statistical probability distributions.
3. Distinguish among the different measurement scales and the implications for selection of statistical methods to be used based on these distinctions.
4. Apply descriptive techniques commonly used to summarize public health data.
5. Learn concepts underlying statistical estimation and hypotheses testing.
6. Apply common statistical methods for inference.
7. Learn concept, use, and assumptions underlying simple and multiple linear regression; and apply to public health data.
8. Learn concept, use, and assumptions underlying ANOVA and ANCOVA; and apply to public health data.
9. Learn concept, use, and assumptions underlying nonparametric methods; and apply to public health data.
10. Apply descriptive and inferential methodologies according to the type of study design for answering a particular research question.
11. Interpret results of statistical analyses found in public health studies.
12. Achieve proficiency in using the SAS statistical package to implement the statistical methods learnt in the course.