EEES 6400/8400 Biostatistics Spring 2015

1. General Information

Instructor:	Song S. Qian	Term	Spring 2015
Email	song.qian@utoledo.edu	Class Day/Time:	Mon/Wed 2:00 - 3:50
Office Hours:	Mon/Wed 9:30-12:00	Class Location:	BO 1005
Office Phone:	419 530 4230	Credit Hours:	4
Office Location:	BO 3001E		

2. Course/Catalog Description

   Application of statistical tools (estimation, hypothesis testing, and statistical modeling) to data analysis in environmental and ecological studies.

3. Course Overview:

   The course covers exploratory analysis of data using graphical methods, statistical inference using hypothesis testing and estimation, and linear regression models. The class will emphasize investigating assumptions and interpreting results. A major focus is on writing literate data analysis reports and critiquing statistical analysis. Examples in the class are from the environmental and ecological sciences. The open-source software package R will be used in class.

   Materials covered in the class are grouped into the following topics

   (a) Basics of statistical inference
      i. Assumptions and graphical presentation of data
      ii. estimation – standard error and confidence interval, bootstrapping
      iii. Hypothesis testing – t-test and nonparametric alternatives

   (b) Statistical modeling
      i. Linear models – ANOVA, multiple comparisons and linear combinations of means, linear regression, nonlinear regression
      ii. Graphical analysis of residuals

4. Student Learning Outcomes

   Upon completion of this course, the student will be able to:

   (a) Propose appropriate models/assumptions based on relevant scientific and statistical knowledge,

   (b) Conduct basic statistical estimation and hypothesis testing – t-test, ANOVA, linear regression, and nonparametric statistical methods for hypothesis testing,

   (c) Use graphical tools for assessing statistical assumptions, and

   (d) Utilize critical thinking in applying statistics.

5. Teaching Strategy

   This course is designed to stimulate student learning through analysis of real world data in environmental and ecological studies, both as in-class examples and in homework assignments. Students are expected to read all assigned readings before each class and actively participate in-class discussions. Lectures include both statistical methods and statistical computing. A unique feature of the class is the data analysis project. Students will identify problems (and associated data) and carry out the analysis during the semester. By the end of the semester, students are to present their results in class.

6. Prerequisite

   Data management or working experience with R

7. Required Texts and Ancillary Materials

   (a) Main text – Qian, S. (2010), *Environmental and Ecological Statistics with R*, Chapman & Hall/CRC Press. (required) Online materials are here: `http://songqiansblog.blogspot.com/2013/04/environmental-and-ecological-statistics.html`

   (b) R reference text – Petra Kuhnert and Bill Venables, (2006). *An Introduction to R: Software for Statistical Modelling & Computing*, available at the CRAN site, `http://cran.r-project.org/doc/contrib/Kuhnert+Venables-R_Course_Notes.zip`)

   (c) Recommended reading –

      i. Salsburg, 2002. *The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century*, Owl Books.

      ii. Levitt & Dubner, 2005. *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything*, Harpe Perennial.

   (d) Software: R – an open source implementation of the S Language, available from CRAN at `http://cran.r-project.org`

   (e) Data sets – R data sets are from several R packages. Additional R data sets will be posted on the class folder.

8. Course Expectations

   Class rules

   (a) Students are expected to attend weekly lectures and actively participate in class activities.

   (b) Late work will not be accepted without prior consent from the instructor.

   (c) Students requesting re-grade must make these requests within one week of receiving the graded material. Attach a note explaining the re-grade issue to your homework or exam and submit to instructor. The instructor has the option to re-grade the entire homework or exam.

   (d) Take-home exams must be completed independently.

   Course work expectations

   (a) Data analysis project is an important part of the class. Students are encouraged to select their own project topics in the first two weeks of the semester. The class will host a public presentation session at the end of the semester, open to the public.

   (b) Examples are an important part of the class. But any given example is unlikely to be intuitive to all. Students are encouraged to submit/recommend examples.

   (c) Statistics uses mathematics, but is not mathematics. The class will not teach mathematical details of statistical methods. Instead, students are expected to learn the statistical modeling process and the rationale behind the process. Specifically, students are to learn the three types of statistical models commonly used in environmental and ecological studies.

   (d) Students are encouraged to study together in order to better understand the materials. Students must complete individual assignments independently. Exams (including take-home exams) must be completed independently.

9. Grading Scheme:

- Homework (30%)
- R assignment (15%)
- Midterm and final exams (30%)
- Project report and project presentation (25%)

Each category will be graded in percentage and the final score is the weighted average of the 4 categories (also in percentage). Project will be graded based on both the final report and poster presentation.

Midterm grades will be entered within one week of the midterm exam. They are used to assit students with determining where they stand academically in the course.

Final grades will be curved based on the final score. But in general, A (>95%), A- (90-95%), B (75-90%), C (50-75%), F (<50%).

Students who do not attend class or stop attending at some point throughout the semester will be given a final grade of F which will impact your overall grade point average. To formally withdraw from this or any other course you need to contact the Registrars Office.

10. Communication Guidelines

    As your instructor, I am here to help, and will do my best to respond to mail within 24 to 48 hours. Students are expected to check their UT email account frequently for important course information. In addition, if you are having difficulty in the course or trouble understanding any aspect of it, please let me know.

    A question-answer session will be held during Mondays lecture. Questions submitted via Email will be addressed on the following Monday.

11. University Policies

    - Policy Statement on Non-Discrimination on the basis of Disability (ADA)

      The University is an equal opportunity educational institution. Please read The Universitys Policy Statement on Nondiscrimination on the Basis of Disability Americans with Disability Act Compliance.

    - Academic Accommodations

      The University of Toledo is committed to providing equal access to education for all students. If you have a documented disability or you believe you have a disability and would like information regarding academic accommodations/adjustments in this course please contact the Student Disability Services Office.

12. Academic Policies

    As a student in my course and enrolled at The University of Toledo you should be familiar with the policies that govern the institutions academic processes, for example, Academic Dishonesty, Enrollment Status, and Grades and Grading. Please read Graduate Academic Policies.

    Students are expected to attend every class meeting of courses in which they are registered. Please read the Missed Class Policy.

13. Course Schedule

Date	Topic	Readings	Contents/Activities	Assignment due
Jan. 12	Data analysis in R	Q. ch 1-2, Box (1976)	Discussing Box (1976)	
Jan. 14	R essentials	Q. ch 1-2 Rstudio	Import data, explore with plots, summary stats	
Jan. 19	MLK Day	No class		
Jan. 21	Assumptions	Q. ch 3	Summaries, plots, and charts in R	HW 1 due
Jan. 26	Statistical processes	Q. ch 3	Find data, produce a one-plot summary	Presentation on Feb. 2
Jan. 28	Graphical Methods	Q. ch 3	Display/compare distributions	HW 2 due
Feb. 2	5 Min presentations of graphed data			Graphing presentation
Feb. 4	Probability distributions			
Feb. 9	Normal distribution Central limit theorem	Q. 4.1		HW 3 due
Feb. 11	Estimation confidence interval	Q. 4.1		
Feb. 16	Hypothesis testing	Q. 4.2		HW 4 due
Feb. 18	Hypothesis testing	Q. 4.3-4		
Feb. 23	Hypothesis testing	Q. 4.5		
Feb. 25	Statistical power	Handout		HW 5 due
Mar. 2	Examples	Q. 4.7.1-3		
Mar. 4	Midterm Exam			HW 6 due
Mar. 9-11	Spring Break			HW 7 due
Mar. 16	ANOVA	Q. 4.6		
Mar. 18	Multiple comparisons	Hurlbert (1984)		HW 8 due
Mar. 23	More ANOVA	Q. 4.6		
Mar. 25	Linear regression	Q. 5.1		
Mar. 30	Linear regression	Q. 5.2		
Apr. 1	Linear Regression	Q. 5.3		HW 9 due
Apr. 6	Class project		HW 10:	
Apr. 8	Regression	Q. 5.4	Find a dataset	HW 10 due
Apr. 13	ANOVA	Q. 5.5		
Apr. 15	Simulation	Q. 5.4, 9.1-3		
Apr. 20	Nonlinear model	Q. 6.1		
Apr. 22	Statistical thinking		Strong inference discussion	
Apr. 27	Parametric v. nonparametric	Seaman & Jaeger (1990)	Class discussion	
Apr. 29	Stat. & science	Box (1976)	Project report	report due
May 7	Final Exam		12:30-2:30	

Last modified: Wednesday 7$^{\text{th}}$ January, 2015, 09:47