Unique Number: 14265

Course Objectives:
To provide students with the proper background and experience to use common hypothesis tests, including tests associated with methods such as regression and ANOVA, and to use common computational statistical methods, such as cluster analysis. The emphasis of this course is on practical skills, including the use of MATLAB®. Applications in biology and medicine will be presented.

Instructors:

<table>
<thead>
<tr>
<th>Name</th>
<th>Office Location</th>
<th>Office Phone</th>
<th>Email &amp; AIM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prof. Mia K. Markey</td>
<td>BME 5.202L</td>
<td>+1.512.471.1711</td>
<td><a href="mailto:biostats@utlists.utexas.edu">biostats@utlists.utexas.edu</a></td>
</tr>
</tbody>
</table>

If you have several questions or a long question (e.g., “I tried to do problem 3, but I am stuck at…”), it usually most effective to come to office hours (see below). I will be available in person, by phone, and/or by IM during our office hours.

If you have a short question (e.g., “Will WRS be on the exam?”, “Is there a typo in problem 7?”), post the question on the discussion board on BlackBoard (Bb). Unless a short question specifically pertains to you only (e.g., request for an excused absence), you should post the question on Bb rather than emailing/calling/etc me. This is important because other students may have the same question as you or the answer may impact everyone (e.g., if there is a typo in problem 7).

Please note that the email address is biostats@utlists.utexas.edu and I will not read or respond to course-related emails sent to my personal email addresses rather than this course email address. By using a listserv, course-related messages are easily sorted out from the many other messages that I receive and automatically archived separately from my other messages.

Office Hours:
Office hours will be held each week. Please refer to the calendar on Bb for the times and formats of office hours as they will vary throughout the semester.

Meeting Times: Mondays (M) and Wednesdays (W), 11:00 AM – 12:30 PM.
Location: ETC 5.148 (Austin) and MSB G520c (Houston)

Note: The Houston classroom is not available on 09/27, 10/25, and 11/22. On these dates Houston students should watch streaming video from another location and participate by web>clicker® as usual.

Teaching Approach
Typically, each class session will involve a series of lecture components with active learning activities interspersed. For example, student teams will lead class discussion on scientific papers, perform hands-on exercises using MATLAB®, and practice problem solving. The central theme of my teaching philosophy is the conviction that students require many opportunities to practice and to get feedback in order to learn. Thus, 10% of your grade will be based on participation in class activities (see grading and evaluation section) and we will employ instructional technologies that facilitate immediate feedback (see section on i>clicker®).

Course Assignments and Announcements
Assignments and announcements will be distributed using BlackBoard (Bb, http://courses.utexas.edu/). It is your responsibility to check this site frequently.

Knowledge, Abilities, and Skills Students Should Have Entering this Course
It is assumed that all students (a) have had an undergraduate course in probability or possess equivalent knowledge and (b) are proficient in MATLAB®. Class time will not be spent on review of these topics. You
must pass the placements tests on probability and MATLAB® in order to take BME 380J.5. Details on the placement tests can be obtained from the BME graduate coordinator. If you are concerned about your preparation for this course, please discuss this with the professor immediately.

Knowledge, Abilities, and Skills Students Should Gain from this Course
Detailed lists of learning objectives are listed at the end of the syllabus.

Required Textbooks

You are expected to read the required textbooks. It is highly unlikely that you will be successful in this course if you don't read the textbooks.

“The man who doesn’t read good books has no advantage over the man who can't read them.”
- Mark Twain

i>clicker® and web>clicker®
This course uses the i>clicker® response system, which allows you to respond to questions I pose during class, and you will be graded on your answers and your in-class participation as described in the Grading and Evaluation section. Austin students should purchase a clicker along with their textbooks. In order to receive credit for using i>clicker®, you will need to register your i>clicker® remote online before the third class session. You must have come to class at least once and answered at least one question in order to complete this registration properly. Once you have answered a question in this class, go to http://www.iclicker.com/registration. Complete the fields with your first name, last name, student ID, and remote ID. Your student ID should be UT EID. The remote ID is the series of numbers and sometimes letters found on the bottom of the back of your i>clicker® remote. i>clicker® will be used every day in class, and you are responsible for bringing your remote daily. If you experience any problems with i>clicker®, please contact Morrie Schulman at +1.512.475.6057.

Houston students will use a web-based alternative (web>clicker®). Houston students should contact Morrie Schulman at +1.512.475.6057 in setting up web>clicker®. Notice that Houston students should bring their laptops to class each session and verify as soon as possible that they can access the internet in the classroom. If internet problems are encountered during class time, Amy Jones (right next to entrance door) can obtain assistance for you. Otherwise, contact Stephen Fath at +1 (713) 500-5202 for technical problems in Houston.

All students should be set up to use i>clicker®/web>clicker® by the second class session. If you don’t have you will lose participation and quiz points starting with the third class session if you do not have i>clicker®/web>clicker® set up.

Failure to use i>clicker®/web>clicker® correctly is not grounds for an extension, grading reconsideration, etc. Please refer to the IT Policy.

Laptops
Students are encouraged to use their laptops in class. While bringing a laptop is particularly important for the days that we will do lab exercises in MATLAB®, you may find it helpful to access the course notes and Bb on other days as well. If you do not bring your laptop every day, make sure that you at least bring a hand calculator as we will solve problems in class. Notice that Houston students are required to bring a laptop to each class session in order to use web>clicker®.

BlackBoard
We will use Bb extensively in this course. For example, you are encouraged ask questions and to answer other students’ questions though the discussion board on Bb. You are responsible for knowing how to use Bb. If you
are unsure about how to use any aspect of Bb, please contact ITS Help Desk by calling +1.512.475.9400 or by submitting a help request at http://www.utexas.edu/its/help/blackboard/. Failure to use Bb correctly is not grounds for an extension, grading reconsideration, etc. Please refer to the IT Policy.

**Time Investment**

The amount of time required to be successful in a course varies from person to person depending on his/her definition of success and knowledge/skills going into the course. However, as a “rule of thumb”, you should expect that *at least six hours per week* outside of the scheduled class time will be required in this course.

**Grading & Evaluation**

<table>
<thead>
<tr>
<th>Component</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participation</td>
<td>10%</td>
</tr>
<tr>
<td>Quizzes</td>
<td>20%</td>
</tr>
<tr>
<td>MidTerm Exam I</td>
<td>25%</td>
</tr>
<tr>
<td>MidTerm Assignment</td>
<td>10%</td>
</tr>
<tr>
<td>Final Exam</td>
<td>35%</td>
</tr>
</tbody>
</table>

10% of your grade is based on participation in class activities (problem solving, MATLAB® lab sessions, paper discussions, etc). The class participation grade is assessed via responses collected with i>clicker®/web>clicker®. For each session in which i>clicker®/web>clicker® is used, you can earn at least half of the available points simply by answering, even if your answers are incorrect. In addition, ‘performance’ points can be earned if correct answers are submitted. The number of points that can be earned for correct answers varies, but is capped at maximum of half of the points available in a session. Your overall participation grade excludes any days for which you had an excused absence; in other words, if the maximum number of points that could be earned on the days you were present was 142, then the denominator for calculating your participation percentage is 142. Technical problems such as forgetting to bring an i>clicker® remote or dead batteries are NOT counted as excused absences. Refer to the Attendance, Participation, and Late Work Policy.

20% of your grade is based on quiz scores. A short quiz is administered at the beginning of each class session. Quiz answers are submitted using i>clicker®/web>clicker®. Your overall Quiz grade excludes any days for which you had an excused absence; in other words, if the maximum number of points that could be earned on the days you were present was 142, then the denominator for calculating your Quiz percentage is 142. Technical problems such as forgetting to bring an i>clicker® remote or dead batteries are NOT counted as excused absences. Refer to the Attendance, Participation, and Late Work Policy.

25% of your grade is based on the MidTerm Exam. The MidTerm Exam is closed book, closed notes, no collaboration, no calculator, etc. In other words, just you and your writing utensil. The MidTerm Exam will take place in class and you will only have 75 minutes to complete it. However, there will be a take-home component to be read in advance of the exam time.

10% of your grade is based on the MidTerm Assignment. The details of the MidTerm Assignment are distributed after the MidTerm Exam.

35% of your grade is based on the Final Exam. The Final Exam will take place during the final exam week as scheduled by the registrar. There will not be a take-home portion to read in advance of the Final Exam; instead, you will need to have developed the skills needed to read the related material efficiently during the allotted exam time (3 hours). The Final Exam will be cumulative, but greater emphasis will be placed on the topics covered in the second half of the semester.

Final letter grades will be assigned as:

- 90-100% A
- 80-90% B
- 70-80% C
- <70% F
While the registrar permits the use of +’s and –’s for B, C, and D grades in graduate courses, I do not assign +’s and –’s; likewise, I do not assign D’s in BME 380J.5. The possible grades that can be earned in this course are A, B, C, and F as above.

Is there a curve?
Should the median of the class prove to be less than 85%, the grades will be curved: (M = median, SD = standard deviation, X = your score)

- \( X > M \)  
  - A
- \( M - 1*SD < X < M \)  
  - B
- \( M - 2*SD < X < M - 1*SD \)  
  - C
- \( X < M - 2*SD \)  
  - F

Only the scores of students taking the course for a grade (not “credit / no credit”) will be used in calculating the median and standard deviation. Please note that it is rarely the case that this curve clause is invoked.

Credit / No Credit
Taking this course “credit / no credit” requires prior approval of the instructor. You should be aware that in order to get a grade of "credit", your course score must correspond to a letter grade of at least “C” as determined by the method described under Grading & Evaluation.

Attendance, Participation, & Late Work Policy
Participation is required at all class session and failure to participate in class will impact your grade as described above in the Grading & Evaluation section. Absences will be excused only for a documented emergency, e.g., you or a dependent are ill, or observance of a religious holy day. By UT Austin policy, you must notify me of your pending absence at least fourteen days prior to the date of observance of a religious holy day. Quizzes and in-class questions collected with i>clicker®/web>clicker® will be dropped when an absence is excused; they cannot be “made up” or submitted late. Technical problems such as forgetting to bring a i>clicker®/web>clicker® remote or dead batteries are NOT counted as excused absences.

IT Usage
Difficulties using BlackBoard, i>clicker®/web>clicker®, computer crashes, lost USB flash drives, etc. are not acceptable reasons for requesting an extension or an excused absence. You are responsible for building the possibility of technology problems into your time management, following directions, and learning how to use IT tools relevant to this course. General technical assistance, such as with Bb, can be obtained by the ITS Help Desk by calling +1.512.475.9400. For assistance with i>clicker®, contact Morri Schultan at +1.512.475.6057. For problems with the distance learning technologies, the contacts are Stephen Fath at +1.713.500.5202 for Houston and Jack Prather at +1.512.232.1531 for Austin.

Academic Integrity
Students who violate University rules on scholastic dishonesty are subject to disciplinary penalties, including the possibility of failure in the course and/or dismissal from the University. Since such dishonesty harms the individual, all students, and the integrity of the University, policies on scholastic dishonesty will be strictly enforced. For further information please visit the Student Judicial Services Web site: http://deanofstudents.utexas.edu/sjs/

Collaboration Policy
You are strongly encouraged to study with other students. However, each student is responsible for fully understanding any materials he/she submits to the instructor. Oral examinations will be administered to resolve any questions of potential violations of the collaboration policy.

Notice
The University of Texas at Austin provides, upon request, appropriate academic adjustments for qualified students with disabilities. For more information, please refer to http://deanofstudents.utexas.edu/ssp/
Hyperlinked Files
There are hyperlinks between files distributed for this course. To take advantage of this feature, maintain the
directory structure when copying the files from Bb to your computer. Add updated files of each content type to
the corresponding directory as they are distributed. The “Overview.ppt” file is a graphical depiction of the
relationships among the course topics and between the course topics and related topics outside the scope of
this course. Expect that updated notes or additional readings may be posted to Bb throughout the semester.
Please note that you must be in presentation mode in PowerPoint in order to use the hyperlinks.
Summary of Schedule and Significant Deadlines

This page provides an at-a-glance tentative schedule of the course topics and the corresponding class dates. Please see the following pages for more detail.

While laptop usage is encouraged every day, you are especially encouraged to bring a laptop with MATLAB® for the days marked “Lab/Discussion”.

<table>
<thead>
<tr>
<th>Class Date</th>
<th>Class Topic</th>
<th>Notice</th>
</tr>
</thead>
<tbody>
<tr>
<td>08/25 W</td>
<td>Introductions; One &amp; Two Sample Inference</td>
<td></td>
</tr>
<tr>
<td>08/30 M</td>
<td>One &amp; Two Sample Inference</td>
<td></td>
</tr>
<tr>
<td>09/01 W</td>
<td>One &amp; Two Sample Inference</td>
<td></td>
</tr>
<tr>
<td>09/06 M</td>
<td>No Class</td>
<td></td>
</tr>
<tr>
<td>09/08 W</td>
<td>Lab/Discussion</td>
<td></td>
</tr>
<tr>
<td>09/13 M</td>
<td>Non-Parametric Methods</td>
<td></td>
</tr>
<tr>
<td>09/15 W</td>
<td>Non-Parametric Methods</td>
<td></td>
</tr>
<tr>
<td>09/20 M</td>
<td>Lab/Discussion</td>
<td></td>
</tr>
<tr>
<td>09/22 W</td>
<td>Categorical Data</td>
<td></td>
</tr>
<tr>
<td>09/27 M</td>
<td>Categorical Data</td>
<td>Houston classroom not available</td>
</tr>
<tr>
<td>09/29 W</td>
<td>Lab/Discussion</td>
<td></td>
</tr>
<tr>
<td>10/04 M</td>
<td>MidTerm Exam</td>
<td>MidTerm Exam</td>
</tr>
<tr>
<td>10/06 W</td>
<td>Regression &amp; Correlation; ANOVA</td>
<td></td>
</tr>
<tr>
<td>10/11 M</td>
<td>Regression &amp; Correlation; ANOVA</td>
<td></td>
</tr>
<tr>
<td>10/13 W</td>
<td>Regression &amp; Correlation; ANOVA</td>
<td></td>
</tr>
<tr>
<td>10/18 M</td>
<td>Lab/Discussion</td>
<td></td>
</tr>
<tr>
<td>10/20 W</td>
<td>ROC, Crossvalidation</td>
<td></td>
</tr>
<tr>
<td>10/25 M</td>
<td>Lab/Discussion</td>
<td>Houston classroom not available</td>
</tr>
<tr>
<td>10/27 W</td>
<td>Estimation, Jackknife, Bootstrap</td>
<td>MidTerm Assignment Due</td>
</tr>
<tr>
<td>11/01 M</td>
<td>Estimation, Jackknife, Bootstrap</td>
<td></td>
</tr>
<tr>
<td>11/03 W</td>
<td>PDF Estimation, Bayes Classifier</td>
<td></td>
</tr>
<tr>
<td>11/08 M</td>
<td>Lab/Discussion</td>
<td></td>
</tr>
<tr>
<td>11/10 W</td>
<td>Cluster Analysis</td>
<td></td>
</tr>
<tr>
<td>11/15 M</td>
<td>PCA &amp; LDA</td>
<td></td>
</tr>
<tr>
<td>11/17 W</td>
<td>Lab/Discussion</td>
<td></td>
</tr>
<tr>
<td>11/22 M</td>
<td>Lab/Discussion</td>
<td>Houston classroom not available</td>
</tr>
<tr>
<td>11/24 W</td>
<td>Ethics &amp; Experimental Design</td>
<td></td>
</tr>
<tr>
<td>11/29 M</td>
<td>Lab/Discussion</td>
<td></td>
</tr>
<tr>
<td>12/01 W</td>
<td>Ethics &amp; Experimental Design</td>
<td></td>
</tr>
</tbody>
</table>

The final exam schedule is released one month in advance. Tentative time is Wednesday, December 8, 9:00-12:00 noon. 

Final Exam
Learning Objectives

The following pages contain critical information concerning the relevant sections of the textbook, names of related files in the course notes, and the learning objectives.

The learning objectives are an explicit statement of the knowledge and skills you are expected to gain from this course. They tell you exactly what you need to know to be successful on learning assessments such as exams.

A few notes on the learning objectives….

(a) “Define” and “state” imply that a definition must be committed to memory, which includes memorizing related equations and being able to make related sketches.

(b) “Recognize” implies that you would not be expected to produce an equation from memory but that you would need to be able to recognize it and recall its meaning without assistance. You must also be able to recognize when the concept or method should be applied.

(c) “Calculate” implies that you should be able to work a problem through to a numerical answer and justify your answer. You also need to be able to draw sketches related to your calculations. Calculations will be done with paper and pencil, with a calculator, and with MATLAB®. However, on exams, often you will only be required to set up a calculation rather than work to a numerical answer.

(d) Mathematical notation will be used consistently in this course and you will only receive full credit if you follow the same notation. For example, “X” implies a random variable named “X” whereas “x” implies a specific value that a random variable “X” may take on.
Topics: Hypothesis Testing: One and Two Sample Inference

Related Files: TheBasics.ppt, OneSampleT.ppt, QQplot.ppt, PairedT.ppt, TwoEqualT.ppt, VarsF.ppt, TwoUnequalT.ppt, SampleSize.ppt

Related Sections in Textbooks: Rosner 6.5, 7.1-7.6, 8.1-8.2, 8.4, 8.6-8.7, 8.10

Comments:
- We will follow the “p value” approach in this course. You will not be assessed on anything about confidence intervals.

Learning Objectives:
- Compare and contrast type I and type II errors
- Compare and contrast significance and power
- Define p-value and interpret calculated values
- Identify the appropriate test, calculate, and interpret the results:
  - for comparing the mean of a normal distribution to a given value
  - for comparing means of two paired normal samples
  - for the equality of variances of two independent normal samples
  - for comparing means of two independent normal samples with equal variances
  - for comparing the means of two independent normal samples with unequal variances
- Use qqplot to assess whether data are sampled from a normal distribution
- Calculate the sample sizes needed for one and two sample tests on the means of normal distributions
- Summarize the relationships between significance, power, sample size, differences in the means, and the variance(s)
Topics: Non-Parametric Methods
Related Files: NonParam.ppt, Sign.ppt, WSR.ppt, WRS.ppt
Related Sections in Textbooks: Rosner 9.1-9.4

Learning Objectives:
- Compare and contrast parametric and non-parametric statistical methods
- Calculate and interpret the sign test
- Calculate and interpret the Wilcoxon signed-rank test
- Calculate and interpret the Wilcoxon rank-sum test
Learning Objectives:
- Calculate and interpret the Chi-square test for contingency tables of categorical data
- Calculate and interpret Fisher’s exact test for contingency tables of categorical data
- Calculate and interpret McNemar’s test for binomial proportions for matched-pair data
- Calculate and interpret the kappa statistic to quantify the degree of association of categorical variables
Learning Objectives:

- Define independent variable, dependent variable, regression line, regression component, and residual and interpret regression coefficients
- Fit simple linear regression and multiple linear regression models using least-squares
- Calculate and interpret F tests for assessing regression coefficients
- State the assumptions made and assess the goodness-of-fit of linear regression models
- Define, calculate, and interpret the sample correlation coefficient
- Calculate one-sample z test for comparing the correlation coefficient of a normal to a given value
- Calculate two-sample z test for comparing the correlation coefficients of two normal samples
Learning Objectives:

- Compare and contrast ANOVA, two-sample t tests, and multiple regression
- Calculate F test for one-way, fixed-effects ANOVA
- Calculate t test for comparisons of groups in one-way, fixed-effects ANOVA, using the Bonferroni adjustment if needed
- Recognize the most common variants of ANOVA
Learning Objectives:
- Compare and contrast the concepts of accuracy, prevalence, sensitivity, specificity, true positive fraction, true negative fraction, false negative fraction, false positive fraction, positive predictive value, negative predictive value, and receiver operating characteristic curve
- Calculate non-parametric ROC curve and the area under the non-parametric ROC curve
- Compare and contrast non-parametric and semi-parametric approaches to ROC analysis
- Compare and contrast cross-validation, bootstrap, and jackknife methods
- Use cross-validation to evaluate a model or classification method
Topics: Estimation, Resampling: Jackknife, Bootstrap
Related Files: Estimation.ppt, Jackknife.ppt, Bootstrap.ppt, ROCboot.ppt,
Related Sections in Textbooks: Coursepacket sections C, D, & E

Learning Objectives:
- Define estimator
- Compare and contrast bias, variance, mean squared error
- Use jackknife sampling to estimate the standard error and the bias of a statistic
- Use bootstrap sampling to estimate the standard error and the bias of a statistic
- Use bootstrap sampling to conduct hypothesis tests on the ROC AUC
Learning Objectives:

- Compare and contrast parametric and non/semi-parametric approaches to PDF estimation
- Define Monte Carlo simulation
- Use histogram methods to estimate probability density functions
- Use kernel methods to estimate probability density functions
- Use the Expectation-Maximization algorithm to estimate the parameters of multivariate mixture models
- Develop and interpret a simple Bayesian classifier
Topics: Cluster Analysis
Related Files: Clustering.ppt, AgglomHierarch.ppt, Kmeans.ppt
Related Sections in Textbooks: Coursepacket sections K, L, & M

Learning Objectives:
- Compare and contrast unsupervised and supervised learning
- Define three properties of distance metrics
- Recognize common distance metrics used to measure similarity
- Perform and interpret agglomerative hierarchical cluster analysis
- Perform and interpret k-means cluster analysis
Topics: PCA & LDA
Related Files: PCA.ppt, LDA.ppt
Related Sections in Textbooks: Coursepacket section N

Learning Objectives:
- Perform and interpret Principal Components Analysis
- Perform and interpret Linear Discriminant Analysis
Topics: Ethics and Experimental Design
Related Files: ProfEthics.ppt, HumanSubjects.ppt, StudyDesign.ppt, ClinicalTrials.ppt, FDA.ppt, AnimalSubjects.ppt
Related Sections in Textbooks: Coursepacket sections O & P

Learning Objectives:
- Compare and contrast common study designs:
  - Completely random factorial
  - Randomized block
  - Crossover
  - Split plot
- Discuss uses of different kinds of positive and negative controls
- Describe the 5 types of clinical trials
- Describe the four phases of clinical trials
- Describe examples of the types of medical products regulated by the FDA
- Discuss statistical analysis and experimental design in the context of practical approaches to avoiding scientific misconduct
- Discuss statistical analysis and experimental design in the context of responsible conduct of research with human participants
- Discuss statistical analysis and experimental design in the context of responsible conduct of research with animal subjects