Fall 2010 MATH 300-A: Probability & Statistics I

Instructor information:

Dr. Brad Thiessen
Office: 414 Ambrose Hall
Phone: 563.333.6160
Site: http://www.bradthiessen.com
Hours: MW 3:30-4:30; F 12:00-1:00
Email: ThiessenBradleyA@sau.edu
twitter.com/Thiessen

Catalog description:

Descriptive statistics, probability concepts, random variables, discrete, continuous, and sampling distributions. Two sample inferences. Sampling considerations. Point and interval estimation of distribution parameters and single sample tests of hypotheses. Prerequisite: MATH 191 or Departmental approval.

Prerequisites:

Calculus I is listed as a prerequisite primarily to ensure students have facility with algebra and a comfort with mathematical concepts and notation. While we will use concepts of Calculus to investigate and derive statistical theory, this course will focus more on statistical reasoning than on specific mathematical techniques. No previous knowledge of probability or statistics is assumed.

By the end of this course, students will...

- Communicate statistical analyses and results using appropriate terminology and notation
- Test hypotheses through simulation and randomization methods
- Develop and use appropriate probability models to solve complex problems
- Apply their understanding of Calculus to derive formulas and calculate probabilities
- Describe data using appropriate and meaningful visual displays and numerical summaries
- Evaluate interpretations and analyses for a given dataset, study, or situation
- Describe common pitfalls and sources of error in statistical analyses
- Accurately interpret output from several statistical applications
- Apply appropriate probability distributions and statistical analyses to make valid decisions
- Evaluate the fit between statistical models and reality

Course materials:

Required: Print out activities prior to each class (posted on Blackboard or bradthiessen.com)
Calculator (TI-8x recommended); Binder for class activities and notes

Recommended: Devore, J.L. *Probability & Statistics for Engineering & the Sciences (7th edition)* (ISBN: 978-0-495038217-1). While this book will also be recommended for MATH 301, you can choose to research topics online rather than buying the book.

Optional Apps: *Stata* (powerful, easy to use; $49/year; $179 lifetime): http://www.stata.com/coursegp.html
SPSS (easy to use; $80/year): http://www.spss.com/
R (powerful, free; takes some time to learn): http://cran.r-project.org/
R-Cmdr (makes R easier to use): http://socserv.mcmaster.ca/jfox/Misc/Rcmdr/
JGR (graphical interface for R): http://jgr.markushelbig.org/Download.html

Grading methods & criteria:

Each unit score is calculated with the following weights:

	No homework/Project	Homework/Project
Exam score	85%	60%
Assignments/Quizzes/Participation	15%	15%
Homework or Project	----	15%

The grading scale for your final course score (an average of your unit scores) will be no higher than:

A (100-90%) B (90-80%) C (80-70%) D (70-60%) F (60-0%)

Extra credit policy:

Students scoring below 70% on unit exams will be required to complete additional projects to demonstrate their mastery of the unit objectives. Successfully completing the project will boost unit grades to a maximum of 70%.

Additionally, one extra credit opportunity per unit may be announced in class.

Attendance policy:

Many concepts and methods will be presented in class through activities that cannot be fully reproduced outside of class. Because of this, it is important that you attend class and arrive on-time. While I will not deduct points from your grade for absences, I have found that students with poor attendance generally do poorly in the class.

If you must miss class, I'd appreciate it if you let me know in advance. You **must** get my approval before missing any scheduled exams or turning in any assignments late. An assignment turned-in past the due date will earn a score no higher than 70%.

Accommodations policy:

Students with disabilities who believe they may need accommodations in this class are encouraged to contact the Office of Services for Students with Disabilities at 333-6275 as soon as possible to better ensure that such accommodations are implemented in a timely fashion

Academic integrity policy:

I encourage you to collaborate in studying or completing in-class activities, homework, or projects. You must work alone on tests. Review the SAU policy at http://web.sau.edu/Registration/documents/AcademicIntegrityPolicy.pdf

Policy on the use of electronic equipment:

Calculators and computers will be used extensively in this course. I recommend you save yourself the embarrassment of having your ringtone play during class.

Course procedures and expectations:

This course will introduce students to modern conceptualizations, applications, and methods of probability (the mathematical study of uncertainty or randomness) and statistics (the science of gaining insight from data).

> *I keep saying that **the sexy job in the next 10 years will be statisticians**. The ability to take data – to be able to understand it, to process it, to extract value from it, to visualize it, to communicate it – that's going to be a **hugely important skill in the next decades**, not only at the professional level but even at the educational level for elementary school kids, for high school kids, for college kids.* -- Hal Varian, Chief Economist at Google (The McKinsey Quarterly, 1/09)
>
> **For Today's Graduate, Just One Word: Statistics**. Lohr, S. (2009, August 6). *New York Times*. http://is.gd/2gZJa
> **The Future of Work: It's Data, Baby**. Ali, I. (2009, August 12). *New York Times.* http://is.gd/2gZEq

Each topic will be introduced through guided classroom activities (which you should download from the website and print before each class). Working together, we will learn important concepts, methods, and statistical reasoning by analyzing realistic applications. During these activities, I will ask questions – lots of questions. If you attend class, read about the topic ahead of time, and participate in these activities, you will learn the material. In addition, you will be able to use these activities (and any notes you write on them) during the unit exams.

The calendar on the next page shows assigned readings and homework problems. I recommend you read the material and complete the homework problems before the topic is discussed in class. That way, you'll get more out of the in-class activities. The vast majority of assigned homework problems are odds, so you'll be able to immediately check your answers in the back of the textbook. You can also ask questions about the homework at the beginning of each class.

Throughout each unit, I will assign some take-home activities or projects. These activities will be opportunities for you to conduct statistical analyses and write short summaries of your findings. You will always have at least one week to complete these take-home activities and I encourage you to collaborate with other students. Just make sure you write, understand, and turn-in your own answers.

Exams will be open-note but, due to time constraints, not open-book. You can expect true/false, multiple-choice, short-answer, and take-home items designed to check your mastery of the objectives listed at the end of this syllabus. The test questions will only sample content we cover in class.

Immediately following the unit exams, I will collect homework/project portfolios. Students completing all the assigned homework problems will receive credit for their efforts. Students who do not complete homework will not be penalized. These students can choose to either (a) have the unit exam count for a greater percentage of their unit score, or (b) complete a project related to the unit content.

I am not terribly interested in your ability to memorize and use formulas and "statistical recipes." I want you to learn the concepts, logic, applications, applicability, and pitfalls of statistics. To do this, you will need to come prepared to class everyday. This means you should read about the assigned topics, complete all assignments, actively participate in class discussions, and review your notes periodically.

Do not fall behind in this class! If you have any questions or need assistance, feel free to work with other students, send me questions (via email, twitter, Blackboard post, or voicemail), or visit my office during my posted office hours.

Fall 2010 **MATH 300-A: Probability & Statistics I**

Week 1 8/25 – 8/27	Course overview and introductions Activity #1: Statistical Thinking Activity #2: Probability Theory	 Read 2.1: 1, 9 Read 2.2: 11, 12, 13, 15, 17, 21, 25
Week 2 8/30– 9/3	Activity #3: Probability Applications Activities #3abc: Randomization Methods	Read 2.3: 29, 30, 31, 33, 39, 41 **Assigned: Permutation Tests**
Week 3 9/6– 9/10 No class 9/6	Activity #4: Discrete random variables; Conditional Probability Activity #5: Conditional probabilities; Independence	Read 2.4: 45, 47, 49, 50, 53, 55, 59 **Assigned: Randomizations** Read 2.5: 71, 80, 83, 108
Week 4 9/13– 9/17	Activity #6: Expected Value and Variance Activity #7: Binomial Distribution (chance for 1 million extra credit pts)	Read 3.1: 1, 7 Read 3.2: 11, 13, 17 **Assigned: Presidential Election** Read 3.2: 11, 13, 17 Read 3.3: 29, 31, 44
Week 5 9/20 – 9/24	Activity #7a: Sign Test Review for exam (Dog Resemblance activity, 7b)	Read 3.4: 46, 47, 49, 50, 55, 57 **Exam #1: Probability (Unit 1 work due)**
Week 6 9/27 – 10/1	Activity #8: Useful Discrete Distributions Activities #8b: Practice Exercises Activity #9: Continuous Random Variables Activity #9b: Battery Problem	Read 3.5: 69, 71, 73, 75 Read 3.6: 79, 81, 83, 85 **Assigned: 3rd base coach, practice exercises** Read 4.1: 1, 3, 5 Read 4.2: 11, 15 **Assigned: Industrial Robot problem**
Week 7 10/4 – 10/8	Activity #10: Exponential distribution Activity #11: Normal Distribution	Read 4.4: 59, 61 Read 4.3: 28, 31, 33, 35, 37, 47 *Extra credit opportunity: Sections 4.5, 4.6* **Assigned: Normal Probability Practice**
Week 8 10/11 - 10/15 No class 10/15	Activity #12a: Central Tendency Activity #13: Exploratory Data Analysis (Introduction to Tableau) Activity #14: Statistical Inference – Point Estimation	 Read 1.3: 33, 37, 39 Read 1.4: 47, 49, 51
Week 9 10/18 - 10/22	*Activity #15: Maximum Likelihood (optional)* Activity #16: Sampling Distributions Activity #17, 17a: Central Limit Theorem	Read 6.1: 1, 3, 5 *Read 6.2: 20 (optional)* Read 5.3: 37 Read 5.4: 46, 47, 49, 53
Week 10 10/25 – 10/29	Activity #17b: Central Limit Theorem Stata Bootstrap Demonstration (if time allows)	Read 5.5: 59, 65 **Assigned: Practice exam**
Week 11 11/1 - 11/5	Activity #18: Confidence Intervals	**Exam #2: Applied Probability (Unit 2 work due)** Read 7.1: 1, 3, 5 Read 7.2: 13, 17, 19, 23
Week 12 11/8 - 11/12	Activity #19: Student's t-distribution Activity #20: Hypothesis Testing	Read 7.3: 29, 30, 33, 35, 37 Read 8.1: 1, 3, 5, 7
Week 13 11/15 – 11/19	Activities #21: T-tests Activity #21bc: Hypothesis test & Confidence Intervals	Read 8.2: 17, 19, 21, 23, 25 Read 8.4: 45, 47, 51, 53
Thanksgiving 11/22 – 11/26	No class	
Week 14 11/29 - 12/3	Review hypothesis testing Activity #22b: Test for proportions Activity #23: Independent Samples t-test theory Activity #24: Independent Samples t-test	Read 8.3: 35, 37, 39 Read 9.1: 1 **Assigned: 2 Group Randomization** Read 9.2: 21, 23c, 25, 27
Week 15 12/6 - 12/10	Extra Credit Activity #24a Activity #25: Dependent Samples t-test Review for final exam	**Assigned: 2 Group Randomization** Read 9.3: 39, 41
FINAL EXAM	Scheduled for:	

6/7/2010

Fall 2010 MATH 300-A: Probability & Statistics I

Unit 1:

Overview (Activity #1)

1. Define the terms *probability, statistics (descriptive statistics),* and *Statistics (statistical inference)*
2. Make decisions based on reported probabilities
3. Explain why random assignment is important in designing studies
4. Calculate simple proportions, odds, relative rates, and odds ratios from a contingency table
5. Given the general definition of a null hypothesis, write the null hypothesis for a specific study
6. Use a simulation method (randomization) to estimate the p-value of a statistical test
7. Explain the difference between a probability and a likelihood

Probability Basics (Activities 2-3)

8. Write out the sample space for simple and compound experiments
9. Apply the slots method (multiplication, factorial, permutation rules) or combinations to calculate the potential outcomes of an experiment
10. Explain the difference between a combination and a permutation using an example of each
11. Recognize when outcomes from an experiment are equally likely (or are not equally likely) to occur
12. Use the complement rule to calculate probabilities of disjoint (mutually exclusive) events
13. Give an example of two or more events that are disjoint; give an example of two events that are not disjoint
14. Draw Venn Diagrams to represent probability rules (such as the general addition rule or the complement rule)
15. Use Venn Diagrams or the general addition rule to calculate probabilities
16. Write out probability statements using correct set notation
17. Given a frequency table, determine probabilities; given a list of probabilities, fill-in a frequency table
18. Use combinations and the definition of probability to solve simple probability problems (beginning hypergeometric probabilities)
19. Use a simulation to estimate probabilities from an experiment
20. Explain solutions to the birthday problem (probability of at least 2 people sharing a birthday) and the Let's Make a Deal problem
21. Explain the following problem (sample space comprehension): Amy has 2 children; the older child is a female. Barb has 2 children; one child is female. The probability that the youngest child is female is 1/2 for Amy and 1/3 for Barb.

Randomization Methods (Activity 3a)

22. Explain the difference between an experimental study and an observational study
23. Write out the null and alternate hypotheses for a given study
24. Write out all possible randomizations from a study
25. Use combinations to determine the number of ways of randomly assigning X people into G groups
26. Use randomization methods to calculate a p-value for a given hypothesis
27. Write the potential consequences of Type I and Type II errors in a given study
28. Write out an interpretation of a p-value in a given study
29. Given a simple data set, complete an analysis using simulation or randomization methods

Discrete Random Variables & Conditional Probability (Activities 4, 5, 5a)

30. Given a list of probabilities, sketch a probability mass function and cumulative distribution function
31. Given a cumulative distribution function, calculate specific probabilities: What is $P(X<3)$?
32. Explain why a cumulative distribution function always starts at zero and ends at 1.0
33. Explain what is meant by a *conditional* probability
34. Write conditional probabilities using correct set notation: $P(A \mid B)$
35. Use Venn Diagrams or the conditional probability rule to calculate conditional probabilities
36. Use the general multiplication rule to calculate probabilities: $P(A \cap B)$
37. Given a frequency table, calculate conditional and joint probabilities (like the Alan & Beth movie example)
38. Explain what it means for two events to be independent
39. Use the conditional probability rule or the general multiplication rule to show that two events are independent
40. Use the law of total probability to calculate the probability of an event
41. Use Bayes' Theorem to calculate posterior probabilities
42. Verify de Morgan's Laws by sketching and shading Venn Diagrams

Discrete Random Variables, Expectation, and Variance (Activity #6)

43. Use the formula to calculate the expected value and variance of a random variable
44. Explain what the expected value and variance of a random variable represent
45. Explain the difference between an expected value and our expectation (of an outcome of an experiment)
46. Use the properties of expected values to determine what will happen if each value of X undergoes a linear transformation
47. Use the formula to calculate the standard deviation of a random variable
48. Use the properties of variances to determine what will happen if each value of X undergoes a linear transformation

Binomial Distribution (Activity #7 & Dog Experiment)

49. Define a discrete random variable
50. Explain the properties of a Bernoulli random variable
51. Derive the expected value and variance of a Bernoulli random variable
52. Evaluate a situation to see if the Binomial distribution applies (independent trials, constant probability of success)
53. Derive the pmf of a binomial distribution

54. Use the pmf to calculate binomial probabilities: P(X<a) P(X=a) P(X>a)
55. Derive the expected value of a binomial distribution
56. Calculate the expected value and variance of a binomial random variable
57. Use a calculator to calculate binomial probabilities
58. Use a binomial table to find binomial probabilities
59. Use the complement rule and a calculator to calculate P(X>a) under a binomial distribution
60. Verify the results of a published research article

Binomial Test & Sign Test (Activity #7a)

61. Write null and alternate hypotheses (for the binomial or sign tests) using correct probability notation
62. Assuming the null hypothesis is true, calculate the p-value of an experiment (probability of observing something as or more extreme)
63. Draw an appropriate conclusion from a p-value
64. Conduct a complete analysis using the binomial test
65. Conduct a complete analysis using the sign test

If time permits...

66. Calculate the desired probability using information typically collected under the method of randomized responses (in surveys)
67. Calculate false positive, false negatives (sensitivity, specificity) rates for a given situation (drug testing, polygraph)

Unit 2:

Discrete Distributions (Activities 8, 8a, 8b, 8c)

68. Define the term *random*
69. Derive pmf's for variables following geometric, negative binomial, hypergeometric, or poisson distributions
70. Derive formulas for the expected value and variance under geometric, negative binomial, hypergeometric, and poisson distributions
71. Use the formulas to calculate probabilities, expected values, and variances under these distributions
72. Sketch the pmf and cdf for a specific case under each distribution
73. Interpret the probabilities, expected values, and variances in an application of each distribution
74. Given an application, determine the appropriate discrete probability distribution to use
75. Explain how hypergeometric random variables differ from binomial random variables (sampling w/o replacement)
76. Explain the conditions under which each distribution can be used to model probabilities
77. Calculate probabilities for each distribution using a calculator (or computer)
78. Apply discrete distributions to solve realistic problems (hypergeometric probabilities, odds ratios, relative rates)
79. Use the multinomial distribution to calculate probabilities

Continuous Distributions (Activities 9, 9a, 9b)

80. Define a continuous random variable
81. Classify variables as either discrete or continuous
82. Given a histogram, calculate probabilities
83. Given a histogram, explain how probability is equivalent to the area under the histogram
84. Explain how to calculate probabilities (area under curves) or continuous random variables
85. Explain the conditions needed to demonstrate that a pdf is legitimate (area sums to 1.0; always positive)
86. Using either geometry or integration, calculate the area under a variety of curves
87. Explain why P(X=a) is always zero with continuous random variables
88. Derive the formulas for the expected value and variance of continuous random variables
89. Given a valid pdf, calculate expected values, variances, and percentiles of continuous distributions
90. Define the pth percentile of a distribution
91. Apply pdfs, area calculations, expected values, and percentiles to solve a realistic problem (industrial robot problem)

Special Continuous Distributions (Activity 10)

92. Calculate probabilities and expected values for variables following a uniform distribution
93. Explain the conditions under which exponential distributions may be appropriate
94. Given the pdf of an exponential distribution, derive the cdf
95. Model a situation using an exponential distribution and calculate probabilities
96. Provide an example of the memoryless property of exponential distributions
97. Use the gamma, beta, and weibull distributions to calculate probabilities
98. Determine the appropriate continuous distribution to model a given situation

Normal Distribution (Activity 11)

99. Describe the visual characteristics of a normal curve
100. Compare and contrast normal distributions for a variety of variables (location, spread)
101. Determine the appropriateness of using a normal distribution to model given random variables
102. Sketch symmetric, positively skewed, and negatively skewed distributions
103. Given the pdf for a standard normal distribution, prove the inflection points are at +/- 1
104. Sketch normal distributions given an expected value and variance
105. Explain why we cannot integrate to calculate normal curve probabilities
106. Use the empirical rule to state what percentage of observations lie within 1, 2, and 3 standard deviations of the mean

107. Interpret a z-score from a normal distribution
108. Calculate z-scores from a normal distribution
109. Explain what happens to the shape of a normal distribution when we transform it to a standard normal distribution
110. Calculate probabilities from a normal distribution
111. Use a calculator to calculate normal distribution probabilities
112. Calculate the pth percentile of a normal distribution
113. Standardize scales using z-scores and interpret the results (identify limitations)

Data Collection (Activity 12)

114. Classify variables into quantitative, qualitative, nominal, ordinal, interval, and ratio
115. Explain simple random, stratified, cluster, systematic, and convenience sampling methods
116. Determine whether a sampling method introduced bias (selection, response, measurement, nonresponse biases)
117. Define *independent* and *dependent* variables
118. Identify the sampling method, independent variable(s), and dependent variable(s) from a given research article

Exploratory Data Analysis (Activities 12a, 13)

119. Derive various methods to determine the "center" of a distribution
120. Demonstrate how the median minimizes the sum of absolute deviations
121. Prove (using calculus or graphical methods) how the sample mean minimizes the sum of squared deviations
122. Given a dataset, create a visual display (stemplot, histogram, and boxplot)
123. Given a visual display, draw conclusions about a dataset
124. Calculate the mean, median, and mode of a dataset
125. Determine which measure of central tendency is most appropriate for a given dataset
126. Determine the impact of an outlier (or linear transformation) on measures of central location
127. Calculate the pth percentile of a sample dataset
128. Calculate the IQR, range, variance, and standard deviation of a sample dataset
129. Calculate the population variance and unbiased estimate of the population variance
130. Determine the impact an outlier (or linear transformation) on measures of spread
131. Derive formulas to show the impact of a linear transformation on the mean and variance

Estimates and Estimators (Activities 14, 15)

132. Use correct notation to define population parameters (Greek letters) and point estimators
133. Use intuition to select the most appropriate point estimator of a parameter
134. Explain what is meant by a biased point estimate
135. Prove that the sample mean and sample proportion are unbiased estimators for the population mean and proportion
136. Explain why we want the variance of a point estimate to be minimized
137. Derive the variance of the sample mean and variance
138. Interpret the components of the mean square error of a point estimate
139. Explain why the unbiased estimate of the population variance has (n-1) in the denominator
140. Explain the term *degrees of freedom*
141. Explain the concept of *maximum likelihood*
142. Maximize the likelihood function in a simple (binomial) case
143. Use maximum likelihood to estimate item parameters in an Item Response Theory example

Sampling Distributions (Activity 16)

144. Given a small population, simulate the sampling distribution of the sample mean through repeated sampling
145. Given a population with a normal distribution, explain the shape and center of the sampling distribution for the sample mean
146. Define the term *standard error*
147. Explain what happens to the standard error as our sample size increases
148. Prove that the expected value of the sample mean is equal to the population mean
149. Prove that the variance of the sample means is equal to the population variance divided by the square root of the sample size
150. Given a population mean and standard deviation, sketch the sampling distribution of the sample mean for various sample sizes

Sampling Distributions (Activities 17, 17a, 17b)

151. Write out the Central Limit Theorem (explain it in your own words)
152. $P(\bar{X} < a)$ Explain the conditions under which the CLT "works"
153. Calculate probabilities from the sampling distribution:
154. Run a computer simulation to verify the CLT

Unit 3:

Confidence Intervals & t-distribution (Activities 18, 19)

155. Derive the formula for a confidence interval for the population mean
156. Interpret a confidence interval
157. Explain why we cannot say, "We are xx% confident that the population mean falls in our interval."
158. Explain how our chosen level of confidence impacts the width of a confidence interval

159. Explain how the sample size impacts the width of a confidence interval
160. Given a desired confidence interval width, determine the necessary sample size
161. Calculate confidence intervals for given situations
162. Derive the formula for a confidence interval for the population proportion
163. Explain the formula for the standard error of the sampling distribution of the proportion
164. Calculate confidence intervals for the proportion in a given situation
165. Explain why a t-distribution must be used when the population standard deviation is unknown
166. Determine the degrees of freedom for a confidence interval for the population mean
167. Calculate confidence intervals using the t-distribution
168. Evaluate different interpretations of confidence intervals (determine which are correct/incorrect)
169. Explain why a confidence interval may be meaningless with a nonrepresentative sample

Hypothesis Testing (general concepts) (Activity 20)

170. Given a study, identify the research goal, target population, sample, sampling procedure, parameter of interest, observed estimator, dependent variable, and independent variable
171. Determine if a given study is observational or experimental
172. Write out the null and alternate hypotheses
173. Explain why hypotheses are written with respect to parameters (and not statistics)
174. Determine if a 1-tailed or 2-tailed test should be used in a given study
175. Define Type I (alpha) error and Type II (beta) error
176. Define power
177. Explain the potential consequences of alpha and beta errors in a given study
178. Set an appropriate level for alpha
179. Explain the logic of statistical inference (hypothesis testing)
180. Explain why we must assume the null hypothesis is true in order to conduct a hypothesis test
181. For a given study, sketch the sampling distribution and locate the critical value (z-score)
182. Convert a critical value in the z-score metric to the sample mean metric
183. Convert an observed sample mean into a z-score
184. Determine whether to retain or reject the null hypothesis by comparing observed and critical values
185. Explain why we never accept the null hypothesis
186. Calculate and interpret the p-value from a hypothesis test
187. Given distributions for both the null and alternate hypotheses, shade the areas corresponding to alpha, beta, and power
188. Calculate the probability of a Type II error and power for a given study

One-sample hypothesis tests for the mean and proportion (Activities 21, 21b, 21c, 22, 22a)

189. Explain the conditions under which we use a t-test rather than a z-test
190. Estimate the p-value under a t-test
191. Given a specified Type I error rate and a p-value, determine whether to retain or reject the null hypothesis
192. Explain the difference between *statistical significant* and *practically significant* results
193. Determine the impact of alpha error rate on power
194. Determine the impact of sample size on power
195. Given a set of data, complete a hypothesis test and write the conclusions
196. Explain the relationship between hypothesis testing and confidence intervals
197. Conduct hypothesis tests on a calculator or computer
198. Interpret results from computer output of a hypothesis test
199. Derive formulas in order to test hypotheses about a population proportion

Independent samples hypothesis tests for the mean and proportion (Activities 23, 24, 24a, 24b)

200. Write appropriate hypotheses regarding the means from two independent samples
201. Derive and sketch the sampling distribution of the difference in means
202. Derive the formula for s_{pooled} (weighted average standard deviation)
203. Determine when to pool variances (equal variance assumption) or to use the Welch-Satterthwaite Method
204. Derive the formula for the confidence interval of the difference in means
205. Conduct an independent samples hypothesis test (z-test and t-test)
206. Determine the appropriate degrees of freedom from a given study
207. Write appropriate conclusions from an independent samples t-test
208. Explain the assumptions needed to conduct an independent samples t-test (independence, equal variances, normal distributions)
209. Explain how to test if two populations have equal variances
210. Explain how to test the normality of a distribution (p-p plots, histograms, chi-square tests)

Dependent samples (matched-pairs) hypothesis tests for the mean (Activity 25)

211. Given a study, determine if the groups are independent or dependent (matched)
212. Derive the formulas and sampling distribution for a dependent samples t-test
213. Conduct a complete dependent samples t-test
214. Explain why dependent samples tests have higher power than independent samples tests

Online Statistics Textbooks – Look up something we've covered in class (or use as your primary textbook):

* Online Stat Book with interactive applets: http://www.onlinestatbook.com/
* UCLA AP Statistics Textbook: http://wiki.stat.ucla.edu/socr/index.php/EBook
* Wikipedia (search for a concept): http://www.wikipedia.org/
* Wikipedia Online Statistics Textbook: http://en.wikibooks.org/wiki/Statistics
* Wikipedia Online Probability Textbook: http://en.wikibooks.org/wiki/Probability
* Stat Primer: http://www.sjsu.edu/faculty/gerstman/StatPrimer/

StatSoft Statistics Encyclopedia: http://www.statsoft.com/textbook/stathome.html
HyperStat: http://davidmlane.com/hyperstat/
Introductory Statistics: http://www.psychstat.missouristate.edu/sbk00.htm
Multivariate Statistics: http://www.psychstat.missouristate.edu/multibook2/mlt.htm
Engineering Statistics: http://www.itl.nist.gov/div898/handbook/
A New View of Statistics: http://www.sportsci.org/resource/stats/index.html
Vassar Concepts and Applications of Inferential Statistics: http://faculty.vassar.edu/lowry/webtext.html
Seeing Statistics: http://www.seeingstatistics.com/
Statistics Notes: http://www2.chass.ncsu.edu/garson/pa765/statnote.htm
Maximum Likelihood Estimation (advanced): http://statgen.iop.kcl.ac.uk/bgim/mle/sslike_1.html
Multivariate Notes (advanced): http://www.gseis.ucla.edu/courses/ed231a1/lect.html

Online Probability Textbooks – Look up something we've covered in class (or use as your primary textbook):

* Elementary Probability: http://www.openmathtext.org/lecture_notes/probability_book2.pdf

Introduction to Probability: http://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/amsbook.mac.pdf
Compendium of Common Probability Distributions: http://www.causascientia.org/math_stat/Dists/Compendium.pdf

Interactive Applets and Statistical Calculators:

* Web Pages That Perform Statistical Calculations (600+ links): http://statpages.org/javastat.html

UCLA Statistics Online Computational Resources: http://www.socr.ucla.edu/
SticiGUI: http://www.stat.berkeley.edu/~stark/SticiGui/index.htm
QuickCalcs: http://www.graphpad.com/quickcalcs/index.cfm
Free Statistics Calculators: http://www.danielsoper.com/statcalc/

Resources for Extra Credit Projects:

* Chance News: http://www.dartmouth.edu/~chance/

Exploring Data: http://exploringdata.cqu.edu.au/
Online Statistics Journals: http://www.math.ucla.edu/~tom/ProbStatJ.html
Electronic Encyclopedia of Statistical Examples and Exercises: http://www.whfreeman.com/eesee/eesee.html
Correlation or Causation: http://jonathan.mueller.faculty.noctrl.edu/100/correlation_or_causation.htm

Statistics Software:

* R Project (free, but somewhat difficult to learn): http://cran.r-project.org/
* Stat Crunch (cheap, web based app): http://www.statcrunch.com/
* Stata (more expensive, but my favorite): http://www.stata.com/

SPSS (even more expensive, but easier to use): http://www.spss.com/

* Statistical Computing: http://www.ats.ucla.edu/stat/

Free Statistics Software: http://statpages.org/javasta2.html

Statistics Education:

Center for Teaching Statistics: http://cts.stat.ucla.edu/
Journal of Statistics Education: http://www.amstat.org/publications/jse/
Statistics and mathematics: tension and cooperation: http://www.stat.purdue.edu/~dsmoore/articles/Statmath.pdf
Undergraduate programs and the future of academic stats: http://www.stat.purdue.edu/~dsmoore/articles/Undergrad.pdf
Statistics Among the Liberal Arts: http://www.stat.purdue.edu/~dsmoore/articles/LibArts.pdf
New pedagogy and new content: the case of stats: http://www.stat.purdue.edu/~dsmoore/articles/PedagogyContent.pdf
Data-based, active learning, post-calculus intro to stats: http://www.rossmanchance.com/iscat/RossmanChanceIASEpaper.pdf