==========   Unit #1   ==========

Activity #1: Introduction to Probability & Statistical Inference

1. Explain the logic behind statistical inference
2. Use simulation methods (randomization), as directed, to estimate the p-value from simple studies
3. Make decisions based on reported probabilities
4. Explain the importance of random assignment
5. Given the general definition of a null hypothesis, write out the null hypotheses for simple studies

Activity #2: Introduction to Probability & Counting

6. Define the terms *random*, *probability*, and *likelihood*.
7. Explain the difference between probability and likelihood
8. Write out probability models (sample space and associated probabilities) for simple and compound experiments
9. Explain and contrast the relative frequency, classical, and subjective approaches to estimating probabilities
10. Use the relative frequency approach to estimate (or to design a method to estimate) simple probabilities
11. Apply the slot method (multiplication, factorial, and permutation rules) to count the number of outcomes in an experiment
12. Choose the correct counting rule by determining whether order matters and whether sampling is done with or without replacement
13. Recognize when outcomes from an experiment are equally likely (or are not equally likely) to occur
14. Use combinations to calculate the number of ways to divide n objects into x groups
15. Use the complement rule to calculate probabilities of disjoint (mutually exclusive) events
16. Draw Venn Diagrams to represent probability rules (such as the general addition rule or the complement rule)
17. Write out probability statements using correct set notation
18. Given a frequency table, calculate simple probabilities
19. Given a list of probabilities, fill-in a frequency table
20. Use the general addition rule to calculate probabilities
21. Use combinations and the general definition of probability to solve simple probability problems (beginning hypergeometric probabilities)
22. Explain solutions to the birthday problem (probability of at least 2 people sharing a birthday) and the Let's Make a Deal problem

Activity #3: Applications of Counting Rules; Permutation Tests

23. Use enumeration methods or exact calculations (hypergeometric probabilities) to confirm probabilities obtained from simulation
24. Explain the importance of random sampling
25. Explain the difference between an experimental study and an observational study
26. Write out the null and alternate hypotheses for a given study
27. Write out all possible randomizations from a small study
28. Write out the potential consequences of Type I and II errors in a given study
29. Given a simple data set, design and complete an analysis using simulation or randomization methods
30. Write out an interpretation of a p-value in a given study

Activity #4: Discrete Random Variables; Conditional Probability; Bayes Theorem

31. Define *discrete random variable* and *continuous random variable*
32. Sketch a probability mass function and cumulative function given a set of probabilities
33. Explain why a cumulative distribution always starts at 0 and ends at 1
34. Classify variables as discrete or continuous
35. Explain what is meant by *conditional* probability
36. Use Venn Diagrams to derive the formula for P(A|B)
37. Use the formula and generate frequency tables to calculate conditional probabilities
38. Derive the general multiplication rule
39. Use the general multiplication rule and generate frequency tables to calculate joint probabilities
40. Given Bayes' Theorem and P(A|B), calculate P(B|A)

Activity #5: Conditional Probability; Independence

41. Explain what it means for two events to be independent
42. Use the conditional probability rule or the general multiplication rule to show that two events are independent
43. Derive the law of total probability and use it to calculate the probability of an event
44. Use Bayes' Theorem to calculate posterior probabilities
45. Verify de Morgan's Laws by sketching and shading Venn Diagrams

Activity #6: Discrete Random Variables - Expected Value & Variance

46. Use the formula to calculate the expected value and variance of a discrete random variable
47. Explain what the expected value, variance, and standard deviation of a random variable represent
48. Calculate the expected value and variance of a variable under a linear transformation

Activity #7: Binomial Distribution

49. List the characteristics of a Bernoulli random variable

50. Derive the expected value and variance of a Bernoulli random variable
51. Evaluate a situation to see if the Binomial distribution applies (independent trials, constant probability of success)
52. Derive the pmf of a binomial distribution
53. Use the pmf to calculate binomial probabilities (by hand or using technology): P(X<a) P(X=a) P(X>a)
54. Derive the expected value of a binomial distribution
55. Calculate the expected value and variance of a binomial random variable
56. Conduct a binomial test, estimate a p-value, and make a valid conclusions
57. Conduct a sign test, estimate a p-value, and make valid conclusions

==========   Unit #2   ==========

Activity #8: Discrete Distributions

58. Derive pmf's for the geometric, negative binomial, hypergeometric, and poisson distributions
59. Derive formulas for the expected value and variance of each discrete distribution
60. Use formulas and technology to calculate probabilities, expected values, and variances for each of these distributions
61. Choose the most appropriate discrete distribution to model a given situation
62. Explain the conditions under which each discrete distribution can be used to model probabilities

Activity #9: Continuous Random Variables

63. Calculate probabilities of continuous random variables via integration techniques
64. Determine whether a formula represents a valid probability density function
65. Explain why P(X=x) = 0 for continuous random variables
66. Derive formulas for the expected value and variance of continuous random variables
67. Given a valid pdf, calculate the expected value, variance, and specific percentiles
68. Define the *pth percentile* of a continuous distribution

Activity #10: Special Continuous Distributions

69. Calculate probabilities and expected values for the uniform and exponential distributions
70. Explain how a Q-Q plot can help determine whether a set of values are adequately modeled by a given distribution
71. Derive the pdf and cdc for an exponential distribution
72. Model a situation using an exponential distribution and calculate probabilities
73. Provide an example of the memoryless property of exponential distributions
74. Determine the most appropriate continuous distribution to model a given situation

Activity #11: Normal Distribution

75. Describe the visual characteristics of a normal distribution; compare and contrast different normal distributions
76. Determine the appropriateness of using a normal distribution to model given random variables
77. Sketch examples of symmetric, positively skewed, negatively skewed, unimodal, and bimodal distributions
78. Given the pdf for a standard normal distribution, prove the inflection points are at +/- 1
79. Sketch a normal distribution with a given expected value and standard deviation
80. Explain why we do not integrate to find probabilities under a normal distribution
81. Use the empirical rule to state what percentage of observations fall within 1, 2, and 3 standard deviations of the mean
82. Interpret a z-score from a normal distribution
83. Calculate a z-score from a normal distribution
84. Calculate probabilities under a normal distribution using a table or technology
85. Calculate percentiles of a normal distribution
86. Standardize variables using z-scores and interpret the results (also identify limitations of this approach)

Activity #12: Visual and statistical summaries

87. Derive and evaluate various methods to determine the "center" of a distribution
88. Demonstrate how the median minimizes the sum of absolute deviations
89. Prove (using calculus or graphical methods) how the sample mean minimizes the sum of squared deviations
90. Given a dataset, create a visual display (stemplot, histogram, and boxplot) using technology
91. Given a visual display, draw conclusions about a dataset
92. Given a dataset, calculate numerical summaries of central tendency, location, and spread
93. Determine the impact of an outlier (or linear transformation) on measures of central location and variation
94. Calculate the pth percentile of a sample dataset
95. Derive formulas to show the impact of a linear transformation on the mean and variance
96. Explain other (robust) numerical summaries, such as trimmed and windsorized means, and graphical displays, such as kernal density plots
97. Draw conclusions from bivariate scatterplots, least squares lines, and lowess regression curves

Activity #13: Point Estimates & Maximum Likelihood Estimation

98. Explain the concept behind maximum likelihood estimation
99. Derive the maximum likelihood estimate of a proportion
100. Use maximum likelihood to estimate examinee ability in a simple Item Response Theory example
101. Derive the maximum likelihood estimate for lambda in an exponential distribution
102. Use correct notation to define population parameters (Greek letters) and point estimators
103. Explain what is meant by a biased point estimate
104. Prove that the sample mean and sample proportion are unbiased estimators for the population mean and proportion
105. Explain why we want the variance of a point estimate to be minimized
106. Derive the variance of the sample mean
107. Explain why the unbiased estimate of the population variance has (n-1) in the denominator
108. Explain the term degrees of freedom

Activity #14: Sampling Distributions

109. Given a small finite population, simulate the sampling distribution of the sample mean through repeated sampling
110. Given a population with a normal distribution, explain the shape and center of the sampling distribution for the sample mean
111. Define the term *standard error*
112. Explain what happens to the standard error as our sample size increases
113. Derive key features of the Central Limit Theorem
114. Given a population mean and standard deviation, sketch the sampling distribution of the sample mean for various sample sizes
115. Use the CLT to calculate probabilities regarding sample means
116. Write out the CLT in your own words, describing the conditions under which the CLT applies
117. Run a computer simulation to verify results from the CLT
118. Calculate probabilities from the sampling distribution

========== Unit #3 ==========

Activity #15: Interval estimation

119. Derive the formula for a confidence interval for the population mean and population proportion
120. Correctly interpret a confidence interval for a population mean or proportion
121. Explain why we cannot say, "We are xx% confident that the population mean falls in our interval."
122. Explain how our chosen level of confidence impacts the width of a confidence interval
123. Explain how the sample size impacts the width of a confidence interval
124. Given a desired confidence interval width, determine the necessary sample size
125. Use technology to calculate confidence intervals
126. Explain why a t-distribution must be used when the population standard deviation is unknown
127. Determine the degrees of freedom for a confidence interval for the population mean
128. Use technology to calculate confidence intervals using the t-distribution
129. Provide an example in which confidence interval is meaningless due to a non-representative sample
130. Use randomization methods to generate a confidence interval for a mean or proportion
131. Explain the bootstrap method and how it can be used to generate a confidence interval

Activity #16: Hypothesis Testing

132. Given a study, identify the goal, target population, sample, sampling procedure, parameter of interest, and observed estimator
133. Given a study, identify the dependent variable(s) and independent variable(s)
134. Determine if a given study is observational or experimental
135. Write out the null and alternate hypotheses
136. Explain why hypotheses are written with respect to parameters (and not statistics)
137. Explain the difference between a1-tailed or 2-tailed test
138. Define - and interpret in the context of a study - Type I (alpha) error and Type II (beta) error
139. Define - and interpret in the context of a study - power
140. Explain why we must assume the null hypothesis is true in order to conduct a hypothesis test
141. For a given study, sketch the sampling distribution and locate the critical value (z-score)
142. Convert a critical value in the z-score metric to the sample mean metric
143. Convert an observed sample mean into a z-score
144. Determine whether to retain or reject the null hypothesis by comparing observed and critical values
145. Explain why we never accept the null hypothesis
146. Calculate and interpret the p-value from a hypothesis test
147. Given distributions for both the null and alternate hypotheses, shade the areas corresponding to alpha, beta, and power
148. Calculate the probability of a Type II error and power for a given study
149. Explain the conditions under which we use a t-test rather than a z-test
150. Use technology to estimate the p-value under a t-test

151. Explain the difference between statistical significant and practically significant results
152. Determine the impact of alpha error rate on power
153. Determine the impact of sample size on power
154. Given a set of data, complete a hypothesis test and write the conclusions
155. Explain the relationship between hypothesis testing and confidence intervals
156. Conduct hypothesis tests on a calculator or computer
157. Interpret results from computer output of a hypothesis test
158. Use technology to compare a single mean to a hypothesized value using randomization methods
159. Explain the process behind the randomization test for a single mean

Activity #17: Independent Samples t-test

160. Write appropriate hypotheses regarding the means from two independent samples
161. Derive and sketch the sampling distribution of the difference in means
162. Derive the formula for s_{pooled} (weighted average standard deviation)
163. Determine when to pool variances or to use other methods (Welch-Satterthwaite, randomization, nonparametric)
164. Derive the formula for the confidence interval of the difference in means

Activity #18: Testing Differences in Means Between Two Independent Groups

165. Use technology to conduct an independent samples t-test
166. Determine the appropriate degrees of freedom from a given study
167. Write appropriate conclusions from an independent samples t-test
168. Explain the assumptions underlying an independent samples t-test
169. Given output from a computer program, draw valid conclusions from an independent samples t-test
170. Determine how many t-tests would be needed to make all pairwise comparisons among g group means
171. Calculate the overall (family-wise) alpha error rate when conducting n t-tests from the same data
172. Explain why we need to be careful when running multiple t-tests
173. Use randomization methods to compare two group means
174. Interpret the concept behind, and results from, a Bayesian model used to compare two group means (BEST)
175. Explain the concept of analysis of variance to compare 2+ group means

Activity #19: Matched-pairs, ANOVA, effect sizes

176. Given a study, determine if the groups are independent or dependent (matched or repeated)
177. Derive the formulas and sampling distribution for a dependent samples t-test
178. Conduct a complete dependent samples t-test
179. Conduct a sign test and interpret the p-value
180. Explain why dependent samples tests have higher power than independent samples tests
181. Randomize SAD statistics to compare multiple group means
182. Calculate and interpret effect sizes (such as Cohen's d)
183. Explain the process of Bayesian estimation