Stat 426 Assignment 7 (due Thur, Mar 13)
Topics: Confidence Intervals, Hypothesis Testing, Types of Errors, Power Function, Size, *p*-Value

1. Some student researchers at Cal Poly took a random sample of 100 students and calculated the ratio of backpack weight to body weight for each student. They found a sample mean ratio of .07713, and a sample standard deviation of .03664.
a) Determine a 90%, 95%, and 99% confidence interval for the population mean μ.
b) Do these intervals suggest that the average Cal Poly student carries less than 10% of his/her body weight in his/her backpack? Explain.
c) Determine a 95% confidence interval for the population standard deviation σ.

2. Suppose $X_1, X_2, \ldots, X_n$ are i.i.d. Normal random variables with mean μ and variance σ^2. Consider the *t*-interval for μ and chi-square interval for σ^2.
a) Determine the width of the 95% *t*-interval, as a function of the sample standard deviation S, when $n = 5$, when $n = 10$, and when $n = 20$.
b) Determine the width of the 95% chi-square-interval, as a function of the sample standard deviation S, when $n = 5$, when $n = 10$, and when $n = 20$.

3. In class we used simulation to investigate the robustness of the *t*-interval against departures from normality of the underlying population. In this exercise you will used simulation to investigate the robustness of the chi-square interval for σ^2.
a) Generate 10,000 simulated samples of size $n = 10$ from a normal distribution. Specify whatever mean μ and variance σ^2 you use. Then for each of your 10,000 simulated samples, produce a 95% confidence interval for σ^2. Calculate the length of each interval, produce a histogram of the lengths, and determine the mean and standard deviation of those lengths. Also determine how many and what proportion of the 10,000 intervals succeed in capturing the actual value of σ^2. Is this close to 95%?
b) Repeat a), but generating the simulated samples from a continuous uniform distribution on the interval (0,10). In addition to everything asked for in a), be sure to report the value of σ^2 here. At the conclusion of your analysis, also comment on how well the chi-square procedure works here.
c) Repeat b), generating the simulated samples from an exponential distribution with parameter equal to 1.

4. Suppose that a basketball player has had a .5 probability of making a free throw throughout his career. Now he claims that he has improved to where his probability of making a free throw is .75. You ask the player to take a random sample of 20 free throws and see how many he makes. Let π represent the success probability, so you want to test H_0: $\pi = .5$ vs. H_1: $\pi = .75$, based on a binomial random variable X with parameters $n = 20$ and π. You decide to reject H_0 whenever $X \geq c$, for some appropriate integer c.
a) Determine the value of c for a significance level of .05.
b) Determine the size of this test.
c) Determine the probability of type II error with this test.
Now increase the sample size from $n = 20$ to $n = 50$.

d) Re-answer the previous three questions for this larger sample size. Comment on how the values change.

5. Reconsider the previous problem, with the sample size of $n = 20$.
a) For each possible value of c (integers from 0 to 20, inclusive), determine the probability of type I error. Produce a graph of these probabilities as a function of c.
b) Repeat a) for the probability of type II error.
c) Produce a graph of the sum of these error probabilities as a function of c. Determine the value of c that minimizes this sum.
d) Suppose that you regard type I error as 3 times more serious than type II error, so you want to minimize 3*Pr(type I error) + Pr(type II error). Produce a graph of this quantity as a function of c. Determine the value of c that minimizes this quantity.

6. Suppose that the continuous random variable $X \sim \text{Uniform}(0, \theta)$. Consider testing the simple hypotheses H_0: $\theta = 15$ vs. H_1: $\theta = 20$ with a test procedure that rejects H_0 whenever $X \geq 10$.
a) Calculate the probability of type I error.
b) Calculate the probability of type II error.
Now consider test procedures that reject H_0 whenever $X \geq k$, where k can be any positive real number.
c) Express the probability of type I error in terms of k.
d) Express the probability of type II error in terms of k.
e) Determine the value of k that minimizes the sum of the two error probabilities. Also report the minimum value of that sum.

7. Suppose that $X_1, X_2, \ldots, X_{10}$ are i.i.d. Poisson random variables with parameter λ, and you want to test H_0: $\lambda \geq 1.0$ vs. H_1: $\lambda < 1.0$. Consider test procedures that reject H_0 when $X_1 + X_2 + \ldots + X_{10} \leq k$.
a) Determine the value of k for a significance level of .05. [*Hint*: First determine the probability distribution of $X_1 + X_2 + \ldots + X_{10}$.]
b) Report the size of this test.
c) Determine and graph the power function of this test.
d) Suppose that the context for this test is counting the number of typographical errors on 10 pages of a textbook. Suppose that a total of 4 errors are found. Determine the p-value of the test.

8. D&S, page 461, #1 Also graph the power function in a).

9. D&S, page 462, #8, 10