Test 2

This is a 75-minute, closed-book, closed-note exam. You may use a non-programmable, non-graphing calculator. A formula sheet and normal probability table will accompany this exam. Simplify each answer according to the capabilities of your calculator.

1. Suppose X is the number of times a student is late to an 8 am class held 42 times during the semester, and the probability of being late on any given day is 20%.

   (a) Use the normal approximation to the binomial to determine the probability that the student is late for five or fewer classes. Be sure to apply a continuity correction.

   (b) Is the normal approximation to the binomial appropriate in this situation? Explain why.

2. Why do many statisticians consider it inappropriate to "accept the null hypothesis"?

3. (a) *Undergraduates only.* Define the p-value of a hypothesis test.

   (b) *Graduates only.* Explain the following statement from ASA's recent statement on p-values:
   "Decisions should not be be based solely on whether a p-value is less than or equal to α."

4. (a) What are the assumptions for constructing a confidence interval for the ratio of two population variances?

   (b) The confidence interval for the ratio of two population variances is not robust against violations of normality. How could we demonstrate this by simulation?

5. A study of the difference in housing costs for on-campus housing versus off-campus housing sampled each of 20 students. A 95% confidence interval for the mean one-semester difference in on-campus housing versus off-campus housing was calculated as (-$151, $307).

   (a) Carefully interpret the 95% confidence interval

   (b) Based on the confidence interval, would you reject or fail to reject a two-sided size $\alpha = 0.05$ test of $H_0 : \mu_{On} - \mu_{Off} = 0$? Explain why.

6. A survey of 1400 South Carolina citizens found that 360 supported legalization of marijuana.

   (a) Construct and interpret a 95% confidence interval for the proportion of South Carolinians who favor legalization of marijuana.

   (b) Suppose you want the sampling error SE for the 95% confidence interval to be $\pm 2\%$. How big should n be?

7. In a study of safety belt use, 91% of SC drivers are assumed to use their belts. If n=50 drivers are randomly sampled for safety belt compliance, would you recommend using a confidence interval constructed using $\hat{p}$ or $\tilde{p}$? Explain.

8. (a) For tests that assume observations are drawn from a normal random sample, how do you check normality before proceeding with the test?

   (b) When conducting a two-sample t-test, explain how you choose between Welch's test (unequal variances) and the pooled t-test.

9. Consider testing $H_0 : \mu = 14$ vs. $H_A : \mu < 14$ at the $\alpha = 0.10$ level. A normal sample of size 15 has a mean of 12.6 and a standard deviation of 3.5 Test the hypothesis and report your conclusion. Graduate students should explain how they would compute a p-value.

10. A state is testing the accuracy of a a vendor's breathalyzer, used to measure blood alcohol content. Suppose calibrated samples with known BAC of 0.08 are to be tested. Consider a test of $H_0 : \mu = 0.08$ vs. $H_A : \mu > 0.08$

(a) In words, describe a Type I error and explain its consequences.

(b) In words, describe a Type II error and explain its consequences.

(c) *Graduate students only* Based on the consequences of a Type I error, which alternative do you think is most appropriate for the state to test: $H_a : \mu < 0.08$, $H_a : \mu \neq 0.08$, of $H_a : \mu > 0.08$? Explain.