STAT 3470 Exam 2 Autumn 2019

Instructions: This exam consists of 12 questions and there are 50 points possible. Where applicable, you must define events and random variables (and give the distribution for the latter), use probability statements, show all your work, and justify your answer in order to receive full credit. You must state and check the assumptions for all inference procedures. Please round to 4 decimal digits for probability calculations and at least 2 decimal digits for other calculations. You work must be clear and organized. You may use a calculator (no internet connected devices/cellphones, no sharing of calculators) and you are permitted one page of notes (handwritten, front and back). There are a couple of pages of blank paper at the end for you to use as scratch paper if needed.

Lecturer: Daryl Swartzentruber

Name: ___________________________ .#: __________________________

Section (Circle One): 12:40 3:00

Statement of Adherence to OSU's Code of Student Conduct:

"By signing below I certify that I have neither given nor received help on this exam and that I am in compliance with the academic misconduct policies laid out in OSU's Code of Student Conduct."

Signature: _________________________________

Please do not turn over until you are instructed to do so.

Multiple Choice Questions (3 points each, please circle the letter for the best option)

1. Which of the following statements about the t-distribution is FALSE?
   a. It has heavier tails than the standard normal distribution to account for the added variability that comes from estimating σ with s.
   b. It is slightly skewed because we don't know the population standard deviation.
   c. It is approximately equal to the standard normal distribution when n is very large
   d. It has a parameter called "degrees of freedom" which get larger as n gets larger

2. Suppose you want to estimate the true proportion of OSU students who plan on buying student football tickets next year. If you want the margin of error to be at most .05 with 98% confidence, how many students do you need to sample? Use the conservative approach.
   a. 33
   b. 43
   c. 421
   d. 543

True/False Questions (2 points each, please circle T or F for each statement)

3. T F The margin of error for a prediction interval is larger than for a confidence interval for the same data.
4. T F The p-value is the probability that we would get a sample result at least as unusual as the one we got, if the alternative hypothesis is really true.
5. T F If the p-value is very high, we should conclude that we are very confident the null hypothesis is true.

For problems 6-8, determine whether the width of the confidence interval would increase, decrease, or stay the same based on the given change. Assume all other things are held constant. Circle your answer. (2 points each)

6. Increasing the significance level	Increase	Decrease	Stay the Same
7. Increasing the sample size	Increase	Decrease	Stay the Same
8. Increasing the point estimate	Increase	Decrease	Stay the Same

9. (12 points) A company that makes cement bricks claims that the thermal conductivity of their bricks is 0.340. A sample of 35 bricks is taken, resulting in a mean thermal conductivity of 0.343 and a standard deviation of 0.010. You would like to test whether or not the company's claim is correct.

i. Circle the correct hypotheses for this test (2 points)

a. $H_0: \mu = 0.34, H_a: \mu > 0.34$ d. $H_0: \mu = 0.34, H_a: \mu \neq 0.34$

b. $H_0: \mu = 0.34, H_a: \mu < 0.34$ e. $H_0: \mu < 0.34, H_a: \mu = 0.34$

c. $H_0: \mu > 0.34, H_a: \mu = 0.34$ f. $H_0: \mu \neq 0.34, H_a: \mu = 0.34$

ii. What type of test is this? Why? (2 points)

a. This is a z-test, because we are given the value of the standard deviation.
b. This is a z-test, because our sample size isn't large enough for a t test.
c. This is a t-test, because we don't know the value of the population standard deviation.
d. This is a t-test, because 35>30.

iii. Conduct the hypothesis test you chose above. Use $\alpha = .05$. You do not have to check assumptions. (6 points total)

Test statistic: $t = \frac{.343 - .34}{\frac{.01}{\sqrt{35}}} = 1.77$

p-value: $2P(t > 1.77) = .0849$ (table gives .05<p<.10)

Since my p-value is .0849 > alpha=.05, I fail to reject the null hypothesis. I do not have statistically significant evidence that the true mean thermal conductivity of this company's bricks is different than .34.

1 point for correct test statistic
1 point for showing work
1 point for correct p-value (or a range of p-values)
1 point for correctly comparing p-value (or range of p-values) to alpha
1 point for correct conclusion ("fail to reject the null hypothesis") based on the calculated p-value
1 point for writing the interpretation/conclusion statement in context
Note: If they pick a z-test in part ii and then do it correctly here, give them full credit for this part

iv. Which of the following would be a Type I error for this test? (2 points)

a. You conclude that the true thermal conductivity is 0.340, when in reality it is different than 0.340
b. You forget to check one of the conditions and it causes the results of your test to be invalid.
c. You use the sample standard deviation instead of the population standard deviation.
d. You conclude that the true thermal conductivity isn't 0.340, when in reality it is 0.340.

10. (6 points) The Gamma(α, β) distribution has pdf: $f(x) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta}, & if\ x > 0 \\ 0, & otherwise \end{cases}$.

Recall that $E(X) = \alpha\beta$ and $Var(X) = \alpha\beta^2$. Suppose that we have a random sample of size n from a Gamma$(\alpha, 3)$ distribution.

i. What is the method of moments estimator of α? (3 points total)

$E(X) = 3\alpha$

So set $3\alpha = \bar{X}$ and solve

$\widehat{\alpha}_{MOM} = \frac{\bar{X}}{3}$

1 point for recognizing that $E(X) = 3\alpha$

1 point for setting up the equation

1 point for final answer

ii. What is the mean squared error (MSE) of the estimator you found in part a? (3 points total)

$MSE\left(\frac{\bar{X}}{3}\right) = Var\left(\frac{\bar{X}}{3}\right) + Bias\left(\frac{\bar{X}}{3}\right)^2$

Note that $E\left(\frac{\bar{X}}{3}\right) = \frac{1}{3}E(X) = \alpha$ so the bias is 0.

Thus $MSE\left(\frac{\bar{X}}{3}\right) = Var\left(\frac{\bar{X}}{3}\right) = \frac{1}{9}Var(\bar{X}) = \frac{1}{9} * \frac{Var(X)}{n} = \frac{1}{9} * \frac{9\alpha}{n} = \frac{\alpha}{n}$

1 point for recognizing/writing MSE formula

1 point for plugging in 0 for the bias

1 point for calculating variance (which is the final answer)

Note that they could use the other formula for MSE, so if they get the right answer and show appropriate work give full credit.

11. (6 points) Suppose $X_1, X_2, \dots, X_n$ are i.i.d. from a distribution with pdf:

$f(x) = \theta e^{\theta} x^{-\theta-1}$ for $x > e$

$f(x) = 0$ otherwise

i. What is the maximum likelihood estimator of θ? [You do not have to check the second derivative to verify that it is a maximum.]

$L(\theta) = \theta^n e^{n\theta} (\prod_{i=1}^{n} x_i)^{-\theta-1}$

$l(\theta) = nlog(\theta) + n\theta + (-\theta - 1) \sum_{i=1}^{n} \log(x_i)$

$l'^{(\theta)} = \frac{n}{\theta} + n - \sum_{i=1}^{n} \log(x_i)$

Set = 0: $\frac{n}{\theta} = \sum_{i=1}^{n} \log(x_i) - n$

$\hat{\theta}_{MLE} = \frac{n}{\sum_{i=1}^{n} \log(x_i) - n}$

1 point for L(theta)

1 point for ln(theta)

1 point for taking derivative

1 point for setting equal to 0

1 point for final answer

ii. What is the maximum likelihood estimator of e^{θ}?

$\widehat{e^{\theta}_{MLE}}$ =e^$\frac{n}{\sum_{i=1}^{n} \log(x_i) - n}$

1 point for raising e to whatever their answer in i was

12. (8 points) In a recent Pew Research poll, 54% out of a random sample of 3487 American adults approve of the House of Representatives' decision to conduct an impeachment inquiry.

i. Calculate a 99% confidence interval for the true proportion of American adults who approve of the inquiry. Make sure to check the assumptions for the interval. (4 points total)

We have a random sample and .54(3487)=1882.98, .46(3487)=1694.02 are both bigger than 10

$$.54 \pm 2.58\sqrt{\frac{(.54)(.46)}{3487}} = (.5182, .5618)$$

1 point for random sample condition
1 point for large sample size condition
1 point for work
1 point for final answer

ii. Which of the following is a correct interpretation of your interval? (Circle the letter of the best answer) (2 points)

a. There is a 99% chance the true proportion of American adults who approve of the inquiry is in my interval.
b. If I took repeated samples of the same size and calculated confidence intervals each time, about 99% of the intervals would contain the sample proportion of American adults who approve of the inquiry.
c. I am 99% confident the true proportion of American adults who approve of the inquiry is in my interval
d. The probability of capturing the true proportion of American adults who approve of the inquiry in my interval is .95

iii. Does this interval provide significant evidence that a majority of American adults approve of the inquiry? Explain.

Yes, because the entire interval is larger than 50%.

1 point for yes
1 point for explanation