

SHOW YOUR WORK FOR FULL CREDIT!

Problem	Max. Points	Your Points
1-5	5	
6	6	
7	7	
8	11	
9	11	
10	4	
11	6	
Total	50	

1. When are p-values negative?

- a. when the test statistic is negative.
- b. when the sample statistic is smaller than the hypothesized value of the parameter
- c. when the confidence interval includes only negative values
- d. when we fail to reject the null hypothesis
- ☒ e. never

2. The average growth of a certain variety of pine tree is 10.1 inches in three years. A biologist claims that a new variety will have a greater three-year growth. A random sample of 45 of the new variety has an average three-year growth of 10.8 inches and a standard deviation of 2.1 inches. The appropriate null and alternate hypotheses to test the biologist's claim are:

- ☒ a. $H_0: \mu = 10.1$ against $H_a: \mu > 10.1$
- b. $H_0: \mu = 10.1$ against $H_a: \mu \neq 10.1$
- c. $H_0: \mu = 10.8$ against $H_a: \mu > 10.8$
- d. $H_0: \mu = 10.1$ against $H_a: \mu < 10.1$
- e. $H_0: \mu = 10.8$ against $H_a: \mu \neq 10.8$

3. Which of the following is statistical inference about p ?

- ☒ a. Drawing conclusions about a population proportion based on information contained in a sample.
- b. Drawing conclusions about a sample proportion based on information contained in a population.
- c. Drawing conclusions about a sample proportion based on the measurements in that sample.
- d. Drawing conclusions about the population mean based on information contained in the sample.
- e. Drawing conclusions about the sample mean based on information contained in a population.

4. The 90% confidence interval for a population mean is (1.2, 5.2)

- a. Then the population mean is 3.2, and the margin of error is 2.
- ☒ b. Then the sample mean is 3.2, and the margin of error is 2.
- c. Then the sample mean is 2, and the margin of error is 3.2.
- d. Then the sample mean is 1.2, and the margin of error is 5.2.

5. Which of the following is true about p-values?

- ☒ a. The p -value for a specific statistical test is the probability (assuming H_0 is true) that the test statistic will take a value at least as extreme as that actually observed.
- b. The p -value for a specific statistical test is the probability (assuming H_0 is true) that the alternative hypothesis is true.
- c. The p -value for a specific statistical test is the probability (assuming H_a is true) that the test statistic will take a value at least as extreme as that actually observed.
- d. All of the above statements are true.
- e. None of the above statements are true.

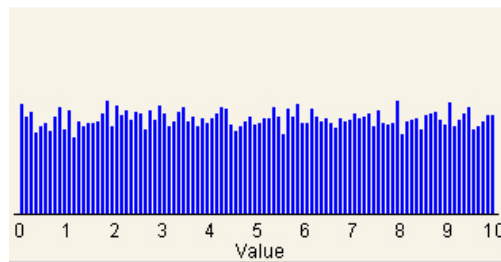
6. For each statement below decide whether it is true or false.

According to the Central Limit Theorem:

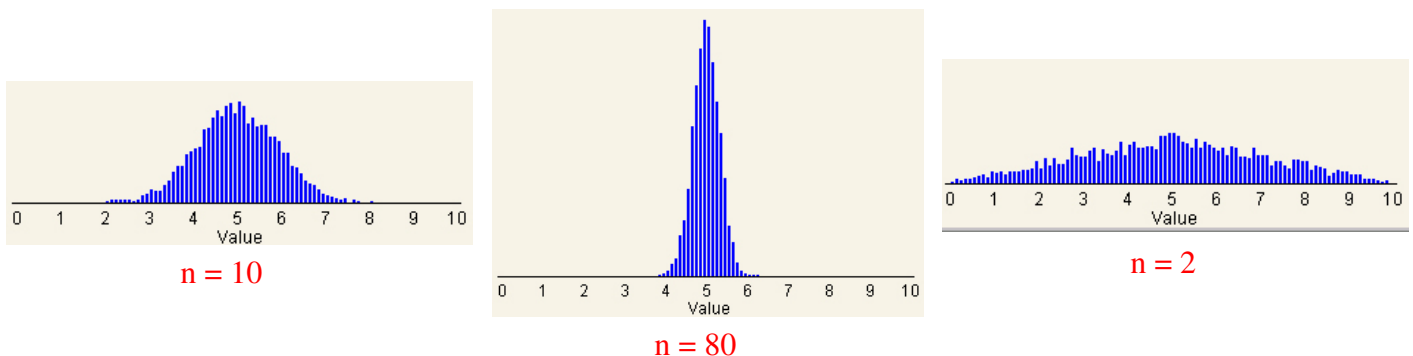
- An increase in sample size from $n = 16$ to $n = 25$ will produce a sampling distribution with a smaller standard deviation. **True** **False**
- The mean of a sampling distribution of sample means is equal to the population mean divided by the square root of the sample size. **True** **False**
- The larger the sample size, the more the sampling distribution of sample means resembles the shape of the population distribution. **True** **False**
- The mean of the sampling distribution of sample means for samples of size $n = 15$ will be the same as the mean of the sampling distribution for samples of size $n = 100$. **True** **False**
- The larger the sample size, the more the sampling distribution of sample means will resemble a normal distribution, regardless of the shape of the population distribution. **True** **False**
- If the shape of the population distribution is itself normal, then the sampling distribution of sample means will resemble a normal distribution for *any* sample size. **True** **False**

7. The histogram below shows the distribution of all values in a given population.

Number of values in population = 16,000; mean $\mu = 4.99$ and standard deviation $\sigma = 2.88$.



A computer simulation produced three sampling distributions of the sample means for $n = 2$, $n = 10$, and $n = 80$.



- Match each distribution with the given sample sizes.
- Calculate the theoretical mean and standard deviation of the sampling distribution for $n = 10$.

Mean = 4.99 standard deviation = $\frac{\sigma}{\sqrt{n}} = \frac{2.88}{\sqrt{10}} = 0.91$

8. The sodium content of a simple random sample of 40 “reduced sodium” hot dogs were collected, and their mean and standard deviation were calculated. Based on the sample results, a 90% confidence interval was calculated: (300.41mg, 319.59mg).

a. In this study, what is the parameter we want to estimate? Be specific.

We want to estimate the MEAN sodium content in ALL “reduced sodium” hot dogs.

b. Interpret the calculated confidence interval in context.

We are 90% confident that the mean sodium content in ALL “reduced sodium” hot dogs is between 300.41 mg and 319.59 mg.

c. What is exactly in the middle of the given confidence interval? Circle ALL the correct answers:

point estimate

sample proportion

sample mean

population proportion

population mean

90%

310mg

margin of error

$$\frac{300.41 + 319.59}{40}$$

d. What could be done to reduce the margin of error? List two ways to achieve this goal.

- Increasing the sample size
- Decreasing the confidence level

e. **True or False?**

If they had taken 100 samples of 40 randomly selected “reduced sodium” hot dogs and had calculated a 90% confidence interval from each sample, probably about 90 of these confidence intervals would not contain the population parameter, and about 10% would contain it.

9. A team of eye surgeons has developed a new technique for a risky eye operation to restore the sight of people blinded from a certain disease. Under the old method, it is known that 31% of the patients who undergo this operation recover their eyesight. Surgeons in various hospitals have performed a total of 225 operations using the new method and 88 have been successful (the patients fully recover their sight). Based on these data, can we justify the claim that the new method is better than the old method?

a. Specify the null and alternative hypotheses for this test, using the correct symbols and numbers.

Null: $p = 0.31$

Alternative: $p > 0.31$

b. Check the conditions for a hypothesis test.

Random ???

$$np_0 = 225(0.31) = 69.75 > 10$$

$$n(1 - p_0) = 225(1 - 0.31) = 155.25 > 10$$

c. Determine the value of the test statistic, and the p-value.

Using the formula for z, or the calculator: $z = 2.63$

p-value = 0.0043

d. Which one of the following statements is true at the 1% level of significance?

- (i) the results are significant, and so we can reject the null hypothesis.
- (ii) the results are not significant, and so we can reject the null hypothesis.
- (iii) the results are significant, and so we cannot reject the null hypothesis.
- (iv) the results are not significant, and so we cannot reject the null hypothesis.

e. State your conclusion in context.

Since the p-value is less than 1%, we can reject the null hypothesis at the 1% significance level. That is, we can reject the claim that 31% of patients who undergo this new operation recover their eyesight. We can conclude that the new method is better than the old.

10. As the newly hired manager of a company that provides cell phone service, you want to determine the percentage of adults in your state who live in a household with cell phones and no land-line phones. You want to be 95% confident that the sample percentage is within 5% of the true population percentage. How many adults you must survey?

Since we don't have a preliminary estimate of $\hat{p}$, we use 0.5.

$$n = \left(\frac{z^*}{m}\right)^2 \hat{p}(1 - \hat{p}) = \left(\frac{1.96}{0.05}\right)^2 0.5(1 - 0.5) = 384.16$$

Thus, we need to survey at least 385 adults.

11. In healthy adults the mean pH level of the blood is 7.4. A new drug for arthritis has been developed. However, it is thought that this drug may change blood pH level. If the drug changes the mean pH level of the blood, FDA will not approve it.

- a. State the null and alternative hypotheses in symbols.

$$H_0: \mu = 7.4$$

$$H_a: \mu \neq 7.4$$

- b. In a double-blind clinical study, a random sample of 60 arthritis patients took the drug for three months, and yielded the following results.

T Confidence Intervals

Variable	N	Mean	StDev	95.0 % CI
DIAMETER	60	7.87	1.62	(7.4515, 8.2885)

T-Test of the Mean

Variable	N	Mean	StDev	T	P
DIAMETER	60	7.87	1.62	2.247	0.0284

If you were one of the quality FDA personnel, what would be your recommendation? In your discussion use both the results from the T-Test, and from the T-Confidence Interval. You must explain your decision, and write your conclusion in context.

Since the CLAIMED value, 7.4 is not in the 95% confidence interval, it's not a plausible value. So we can reject it. Also, since the p-value is less than 5% (remember, the confidence level and the significance level must add up to 100%), we have enough evidence to reject the null hypothesis. So in both ways we can conclude that we have enough evidence to reject the claim that the mean blood pH level is 7.4. So we can conclude that the new drug changes the blood pH level. Therefore, I would not recommend the approval of the new drug.

- c. Is the confidence interval valid if the distribution from which the 60 measurements were taken is not normally distributed? Explain briefly.

Yes, it's valid because the sample size is large, $n = 60 > 30$. By the CLT, we know that the shape of the sampling distribution will be approximately normal.

In general: $\hat{p} = \frac{x}{n}$

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n+4}} \quad \text{where} \quad \hat{p} = \frac{x+2}{n+4}$$

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

$$n = \left(\frac{z^*}{m} \right)^2 \hat{p}(1-\hat{p})$$

$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

$$\bar{x} \pm t^* \frac{s}{\sqrt{n}}$$

z* for 90% confidence: 1.645

z* for 95% confidence: 1.96

z* for 99% confidence: 2.576