Carrying Out a Test for a Population Mean Quiz

1. For a certain brand of canned corn, the company claims that the mean weight of the contents of the cans is 15.25 ounces. A random sample of 36 cans were selected. The sample was found to have mean 15.18 ounces and standard deviation 0.12 ounce. A hypothesis test will be conducted to investigate whether there is evidence to support the belief that the mean is less than 15.25 ounces.

Which of the following is the correct test statistic for the hypothesis test?

(A) $t = \frac{15.18-15.25}{\frac{0.12}{6}}$ ✓

(B) $t = \frac{15.18-15.25}{\frac{0.12}{36}}$

(C) $t = \frac{15.25-15.18}{\frac{0.12}{6}}$

(D) $t = \frac{15.25-15.18}{\frac{0.12}{36}}$

(E) $t = \frac{15.25-15.18}{0.12}$

Answer A

Correct. The numerator of the test statistic is the sample mean (15.18) minus the assumed population mean (15.25). The denominator of the test statistic is the sample standard deviation (0.12) divided by the square root of the sample size $\left(\sqrt{36} = 6\right)$.

2. A magazine article reported that college students spend an average of \$100 on a first date. A university sociologist believed that number was too high for the students at the university. The sociologist surveyed 32 randomly selected students from the university and obtained a sample mean of \$92.23 for the most recent first dates. A one-sample t -test resulted in a p-value of 0.026.

Which of the following is a correct interpretation of the p-value?

(A) The probability is 0.026 that the mean amount of money students from the university spend on a first date is less than \$100.

(B) The probability is 0.026 that the mean amount of money students from the university spend on a first date is less than \$92.23.

(C) The probability is 0.026 that the mean amount of money students from the university spend on a first date is more than \$92.23.

(D) If the mean amount of money that students from the university spend on a first date is \$100, the probability is 0.026 that a randomly selected group of 32 students from the university would spend a mean of \$92.23 or less on their most recent first dates. ✓

(E) If the mean amount of money that students from the university spend on a first date is less than \$100, the probability is 0.026 that a randomly selected group of 32 students from the university would spend a mean of \$92.23 or less on their most recent first dates.

Carrying Out a Test for a Population Mean Quiz

Answer D

Correct. The p-value is the probability that, if many samples of size 32 were selected from the students of the university, we would observe a sample mean of \$92.23 or less, given that the population mean is \$100.

3. A fast-food restaurant claims that a small order of french fries contains 120 calories. A nutritionist is concerned that the true average calorie count is higher than that. The nutritionist randomly selects 35 small orders of french fries and determines their calories. The resulting sample mean is 155.6 calories, and the p-value for the hypothesis test is 0.00093.

Which of the following is a correct interpretation of the p-value?

(A) If the population mean is 120 calories, the p-value of 0.00093 is the probability of observing a sample mean of 155.6 calories or more. ✓

(B) If the population mean is 120 calories, the p-value of 0.00093 is the probability of observing a sample mean of 155.6 calories or less.

(C) If the population mean is 120 calories, the p-value of 0.00093 is the probability of observing a sample mean of 155.6 calories or more, or a sample mean of 84.4 calories or less.

(D) If the population mean is 155.6 calories, the p-value of 0.00093 is the probability of observing a sample mean of 120 calories or more.

(E) If the population mean is 155.6 calories, the p-value of 0.00093 is the probability of observing a sample mean of 120 calories or less.

Answer A

Correct. If the null hypothesis is true, the p-value is the probability of observing the sample mean (155.6) or more.

Carrying Out a Test for a Population Mean Quiz

4. Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

A recent study reported that high school students spend an average of 94 minutes per day texting. Jenna claims that the average for the students at her large high school is greater than 94 minutes. She will conduct a study to investigate this claim.

(a) To collect data, Jenna will select a sample of size 32 from the population.

(i) State Jenna's population of interest.

(ii) Name and describe a sampling method Jenna could use that will satisfy the conditions needed for inference to the population.

(b) Based on a sample of 32 students, Jenna calculated a sample mean of 96.5 minutes and a sample standard deviation of 6.3 minutes. Assume all conditions for inference are met. At the significance level of $\alpha = 0.05$, do the data provide convincing statistical evidence to support Jenna's claim? Complete an appropriate inference procedure to support your answer.

4-part Inference scoring

The primary goals of this question are to assess a student's ability to (1) state a population of interest; (2) describe a random sampling procedure; and (3) perform a significance test and make a conclusion

Scoring

Part (a) is scored as essentially correct (E), partially correct (P), or incorrect (I). Part (b) has three scoring steps. Those scoring steps 1, 2, 3 are each scored as essentially correct (E), partially correct (P), or incorrect (I). So, part (a) is worth 1 E and part (b) is worth 3 Es.

Each essentially correct (E) part counts as 1 point.

Each partially correct (P) part counts as ½ point.

If a response is between two scores (for example, 2½ points), use a holistic approach to decide whether to score up or down, depending on the overall strength of the response and communication.

Reasons to score up:

· *All notation is correct and clearly marked*

· *All explanations are clear*

· *No wrong information is included that was not part of the scoring (for example, saying sample size must be greater than 30 when that has nothing to do with the problem)*

· *No minor calculation errors are made, if they are not part of the scoring*

· *Interpretation parts are especially strong*

Carrying Out a Test for a Population Mean Quiz

Reasons to score down:

· Notation is not wrong, but is spotty and not clearly marked

· Explanations are not wrong, but are hard to follow

· Wrong or extraneous information is included but not part of scoring

· Minor calculation errors that are not part of the scoring are made

Interpretation parts are scored an E but are considered a weak E

0	1	2	3	4 ✓

Scoring steps 1, 2, 3 and part (a) sum to 4 points

OR

Scoring steps 1, 2, 3 and part (a) sum to 3½ points AND a holistic approach is used to decide to score up

- ☐ Part (a) essentially correct
- ☐ Part (a) partially correct
- ☐ Part (a) incorrect
- ☐ Scoring Step 1 Part (b) essentially correct
- ☐ Scoring Step 1 Part (b) partially correct
- ☐ Scoring Step 1 Part (b) incorrect
- ☐ Scoring Step 2 Part (b) essentially correct
- ☐ Scoring Step 2 Part (b) partially correct
- ☐ Scoring Step 2 Part (b) incorrect
- ☐ Scoring Step 3 Part (b) essentially correct
- ☐ Scoring Step 3 Part (b) partially correct
- ☐ Scoring Step 3 Part (b) incorrect

Solution

Part (a)

(i) The population of interest is all students at Jenna's high school.

(ii) One possible method is to use a simple random sample. Obtain (or create) a numbered list of the names of all students at the school, where each student has a unique number. Using a random number generator, generate 32 different random numbers from that list. The students whose names are associated with the selected numbers are used as the sample.

OR

Carrying Out a Test for a Population Mean Quiz

One possible method is to use a stratified random sample. Using class year as strata, obtain (or create) a numbered list of the names of all freshmen at the school, where each freshmen has a unique number. Create similar lists for the other class years. Then use a random number generator to select a random sample of size 8 from each of the classes of freshmen, sophomores, juniors, and seniors.

Scoring

Part (a) is scored as follows:

Essentially correct (E) if the response satisfies the following three components:

· In part (a-i), the response gives a correct description of the population.

· In part (a-ii), the response identifies that the individuals in the population will be represented in the sample by name or number.

· In part (a-ii), a random sampling method is identified and an explanation is given for how the sampling will be conducted.

Partially correct (P) if the response satisfies only two of the three components

Incorrect (I) if the response does not meet the criteria for E or P

Note:

A response that gives a benefit of stratifying is considered a positive in holistic scoring. For example, one possible benefit is that there may be an association between the class year and the amount of texting per day, so the stratification will help to address this variability in texting between class years.

Solution

Part (b)

Scoring step 1: Identification and hypotheses

The appropriate test is the one-sample *t*-test for a population mean. The hypotheses are

$H_0 : \mu = 94$
$H_a : \mu > 94$

where μ represents the population mean number of minutes texting per day for students at Jenna's school.

Scoring

Part (b)

Scoring step (1) is scored as follows:

Essentially correct (E) if the response satisfies the following three components:

· Correctly states, either by name or by formula, the *t*-test for a population mean: $\bar{x} \pm t^* \frac{s}{\sqrt{n}}$

· Population parameter correctly defined in context

Carrying Out a Test for a Population Mean Quiz

· Correct hypotheses

Partially correct (P) if the response includes only two of the three components

Incorrect (I) if the response does not meet the criteria for E or P

Note: Component 1 can be satisfied in scoring step 2 if the correct formula is shown.

Solution

Part (b)

Scoring step 2: Calculation

We are told all conditions for inference have been met. With a sample size of 32, there are $n - 1 = 32 - 1 = 31$ degrees of freedom.

The test statistic is $t = \frac{96.5-94}{\frac{6.3}{\sqrt{32}}} \approx 2.24$

Using a *t*-distribution with 31 degrees of freedom, $P(t > 2.24) = 0.016$

So, the *p*-value is 0.016.

Scoring

Part (b)

Scoring step (2) is scored as follows:

Essentially correct (E) if the response satisfies the following two components.

· The correct test statistic.

· The correct *p*-value *or* correct critical value $t^* = 1.696$

Partially correct (P) if the response satisfies only one of the two components

Incorrect (I) if the response does not meet the criteria for E or P

Note: The *p*-value in component 2 is satisfied if it follows correctly from an incorrectly calculated value of the test statistic.

Solution

Part (b)

Scoring step 3: Conclusion with justification

The *p*-value of 0.016 is less than the significance level 0.05. (Or, the test statistic *t* = 2.24 is greater than the critical value t^* of 1.696.) The null hypothesis is rejected. There is convincing statistical evidence that the mean number of minutes that all students at Jenna's high school spend per day on texting is greater than 94 minutes.

Scoring

Carrying Out a Test for a Population Mean Quiz

Part (b) is scored as follows.

Scoring step (3) is scored as follows:

Essentially correct (E) if the response satisfies the following three components.

· Comparison of *p*-value to alpha (or test statistic to critical value).

· Correct decision to reject the null hypothesis

· States conclusion about the population parameter in context.

Partially correct (P) if the response satisfies only two of the three components

Incorrect (I) if the response does not meet the criteria for E or P

5. A century ago, the average height of adult women in the United States was 63 inches. Researchers believe that the average might be greater today. A random sample of 40 adult women was selected from the population. The sample had mean 64.2 inches and standard deviation 2.9 inches. Assuming all conditions for inference are met, the researchers will perform an appropriate hypothesis test to investigate their belief.

Which of the following is the correct test statistic for the hypothesis test?

(A) $t = \frac{63-64.2}{2.9}$

(B) $t = \frac{63-64.2}{\frac{2.9}{\sqrt{39}}}$

(C) $t = \frac{63-64.2}{\frac{2.9}{\sqrt{40}}}$

(D) $t = \frac{64.2-63}{\frac{2.9}{\sqrt{39}}}$

(E) $t = \frac{64.2-63}{\frac{2.9}{\sqrt{40}}}$ ✓

Answer E

Correct. The numerator of the test statistic is the sample mean (64.2) minus the hypothesized mean (63). The denominator of the test statistic is the sample standard deviation divided by the square root of the sample size $\left(\sqrt{40}\right)$.

6. The distribution of mass for United States pennies minted since 1982 is approximately normal with mean 2.5 grams. A random sample of 10 pennies minted since 1982 was selected. The sample had a mean mass of 2.47 grams and a standard deviation of 0.04 gram.

The test statistic for the population mean has which of the following distributions?

Carrying Out a Test for a Population Mean Quiz

(A) A normal distribution with mean 0 and standard deviation 1

(B) A normal distribution with mean 2.5 and standard deviation 0.04

(C) A normal distribution with mean 2.47 and standard deviation 0.04

(D) A t-distribution with 9 degrees of freedom ✓

(E) A t-distribution with 10 degrees of freedom

Answer D

Correct. When the sample standard deviation is used instead of the population standard deviation, the test statistic follows a t-distribution with degrees of freedom equal to the sample size minus 1. In this case, the sample size was 10, so there are 9 degrees of freedom.

7. Most dermatologists recommend that the ideal shower lasts approximately 10 minutes. A researcher suspects that the average shower length of high school students is greater than 10 minutes. To test the belief, the researcher surveyed 125 randomly selected high school students and found that their average shower length was 14.7 minutes. With all conditions for inference met, a hypothesis test was conducted at the significance level of $\alpha = 0.05$, and the test produced a p-value of 0.0000. Which of the following is an appropriate conclusion?

(A) The test was flawed because the p-value cannot equal 0.

(B) The researcher has statistical evidence to conclude that the sample mean shower length for high school students is greater than 10 minutes.

(C) The researcher does not have statistical evidence to conclude that the sample mean shower length for high school students is greater than 10 minutes.

(D) The researcher has statistical evidence to conclude that the population mean shower length for high school students is greater than 10 minutes. ✓

(E) The researcher does not have statistical evidence to conclude that the population mean shower length for high school students is greater than 10 minutes.

Answer D

Correct. The p-value of 0.0000 is less than the significance level of 0.05, and the correct decision is to reject the null hypothesis. There is statistical evidence to support the alternative hypothesis and conclude that the population mean shower length for high school students is greater than 10 minutes.

Carrying Out a Test for a Population Mean Quiz

8. A car company claims that its new car, the GoFast2000, has a gas mileage of 35 miles per gallon (mpg). A consumer group suspects that the true mean gas mileage of the new cars is less than 35 mpg. The group tests 50 randomly selected GoFast2000 cars and finds a sample mean of 34.8 mpg. With all assumptions for inference met, a hypothesis test resulted in a p-value of 0.324.

For a significance level of $\alpha = 0.05$, which of the following is a correct conclusion?

(A) The p-value is less than 0.05, and the null hypothesis is rejected. There is convincing statistical evidence that the mean is less than 35 mpg.

(B) The p-value is less than 0.05, and the null hypothesis is not rejected. There is not convincing statistical evidence that the mean is less than 35 mpg.

(C) The p-value is greater than 0.05, and the null hypothesis is rejected. There is not convincing statistical evidence that the mean is less than 35 mpg.

(D) The p-value is greater than 0.05, and the null hypothesis is not rejected. There is convincing statistical evidence that the mean is 35 mpg.

(E) The p-value is greater than 0.05, and the null hypothesis is not rejected. There is not convincing statistical evidence that the mean is less than 35 mpg. ✓

Answer E

Correct. The null hypothesis is not rejected because the p-value of 0.324 is greater than the significance level of 0.05. There is not convincing statistical evidence to support the alternative hypothesis that the gas mileage is less than 35 mpg.

9. A bank manager wants the average time that a customer waits in line to be at most 3 minutes. Customers at the bank have complained about the long wait times. To test whether the average wait time at the bank is greater than 3 minutes, 60 customers were randomly selected as they entered the bank and their wait times were recorded. The mean wait time was 4.7 minutes. A one-sample t-test resulted in a p-value of 0.00031.

Which of the following is an appropriate interpretation of the p-value?

(A) The probability that the population mean wait time is greater than 3 minutes is 0.00031.

(B) The probability that the sample mean wait time is greater than 3 minutes is 0.00031.

(C) If the population mean wait time is greater than 3 minutes, the probability of observing a sample mean wait time of 4.7 minutes or more is 0.00031.

(D) If the population mean wait time is 3 minutes, the probability of observing a sample mean wait time of 4.7 minutes is 0.00031.

(E) If the population mean wait time is 3 minutes, the probability of observing a sample mean wait time of 4.7 minutes or more is 0.00031. ✓

Carrying Out a Test for a Population Mean Quiz

Answer E

Correct. If the null hypothesis is true and the mean is 3 minutes, the p-value is the probability of observing the sample mean of 4.7 minutes or more.

10. For a certain brand of tomato seeds, the seed package claims that it takes 87 days after planting for the tomato plants to produce fruit. Sarah, a botanist, wanted to know whether the mean number of days for the plants to produce fruit where she lives is different from 87 days. She planted 40 seeds and recorded the number of days for each plant to produce fruit. With all conditions for inference met, the hypothesis test was conducted at the significance level $\alpha = 0.05$, and the test resulted in a p-value of 0.0752.

Which of the following is a correct conclusion?

(A) Sarah has convincing statistical evidence to conclude that the population mean number of days for the plants to produce fruit is greater than 87 days.

(B) Sarah has convincing statistical evidence to conclude that the population mean number of days for the plants to produce fruit is different from 87 days.

(C) Sarah does not have convincing statistical evidence to conclude that the population mean number of days for the plants to produce fruit is greater than 87 days.

(D) Sarah does not have convincing statistical evidence to conclude that the population mean number of days for the plants to produce fruit is different from 87 days. ✓

(E) Sarah does not have convincing statistical evidence to conclude that the population mean number of days for the plants to produce fruit is equal to 87 days.

Answer D

Correct. The p-value of 0.0752 is greater than the significance level of 0.05, and the null hypothesis is not rejected. There is not convincing statistical evidence to support the alternative hypothesis that the mean time is different from 87 days.