DID I GET THIS: State the Hypotheses for a test for a population mean

For the following scenarios:

- Give the null and alternative hypotheses and
- State in words what μ represents in your hypotheses (define your **parameter** μ)

Question 1: The National Assessment of Educational Progress (NAEP) is administered annually to 4th, 8th, and 12th graders in the United States. On the math assessment, a score above 275 is considered an indication that a student has the skills to balance a checkbook. In a random sample of 500 young men between the ages of 18 and 20, the mean NAEP math score is 272. Do we have evidence to support the claim that young men nationwide have a mean score below 275?

CHECK ANSWER

Hypotheses:
H0: $\mu = 275$
Ha: $\mu < 275$
Where μ is the mean NAEP score for young men in the United States.

Question 2: The National Center for Health Statistics reports that the systolic blood pressure for males 35 to 44 years of age has a mean of 128. In a study of business executives, a random sample of 100 executives has a mean systolic blood pressure of 134. Do the data suggest that the mean systolic blood pressure for business executives is higher than 128?

CHECK ANSWER

Hypotheses:
H0: $\mu = 128$
Ha: $\mu > 128$
Where μ is the mean systolic blood pressure for business executives.

Question 3: An analytical chemistry lab is conducting quality control tests on a drug. A single dosage of the drug should contain 8 mg of active ingredient. Of course, there will be a small amount of variability due to imperfections in the production process, but the mean of all dosages produced should be 8 mg. In 20 random samples, the mean amount of active ingredient is 7.7 mg. Do the data suggest that the mean amount of active ingredient in all dosages produced is different from 8 mg?

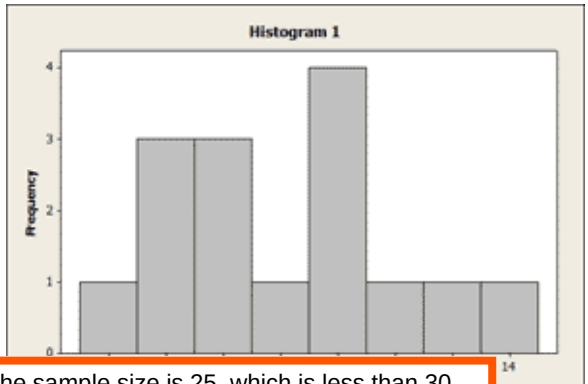
CHECK ANSWER

Hypotheses:
H0: $\mu = 8$
Ha: $\mu \neq 8$
Where μ is the mean amount of active ingredient in all drug dosages produced.

DID I GET THIS: Checking Conditions for Hypothesis Testing for the Population Mean

Use the histograms provided to answer the following questions.

QUESTION 1: The sample size is 25. Using the histogram, decide if we should proceed with a hypothesis test for the population mean and explain why or why not.



(A) is Incorrect. The sample size is 25, which is less than 30.

SELECT (A) Yes, the sample size is large enough.

SELECT (B) Yes, although the sample size < 30, there are no outliers.

(B) is Incorrect. When the sample size is less than 30, what else do we need to know besides that there are no outliers?.

SELECT (C) Yes, although the sample size < 30, the distribution is not very far from normal in shape, with no outliers.

(C) is Correct. Sample size is less than 30, but distribution is not very far from normal in shape, with no outliers.

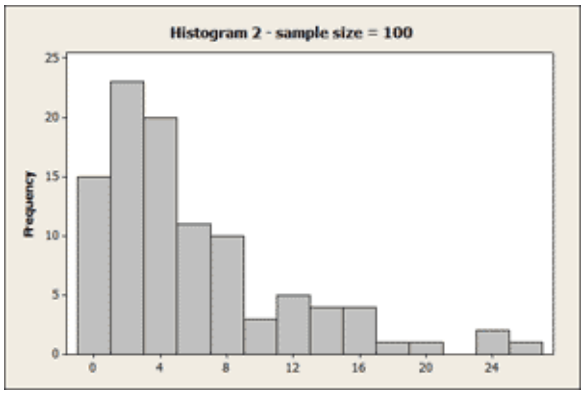
SELECT (D) No, the sample size is not large enough.

SELECT (E) No, the sample size is < 30 and there are outliers.

(D) is Incorrect. Even though the sample size is less than 30, we can proceed if certain conditions are met.

(E) is Incorrect. Even though the sample size is less than 30, we can proceed if certain conditions are met.

QUESTION 2:



(A) Correct. The sample size is greater than 30, so sample means will be normally distributed. Therefore, the t-test can be used.

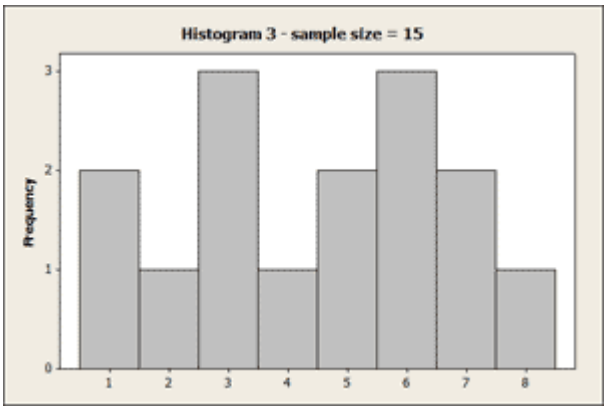
SELECT (A) Conditions are met; it is safe to proceed with the *t*-test.

SELECT (B) Conditions may not be met; conduct the *t*-test but include a disclaimer with the results.

SELECT (C) Conditions are not met; do not use the *t*-test.

(B and C) Incorrect. Look at the size of the sample and recheck the conditions for the t-test given in the table.

QUESTION 3:



(A) Incorrect. The sample size is only 15, so we can't be sure that the Central Limit Theorem applies.

(B) is Correct. This is a borderline case. The sample size is only 15, so we can't be sure that the Central Limit Theorem applies. However, the sample is not heavily skewed and there are no outliers, so we will assume that this sample could have come from a population in which the variable is normally distributed. We will use the t-test with caution.

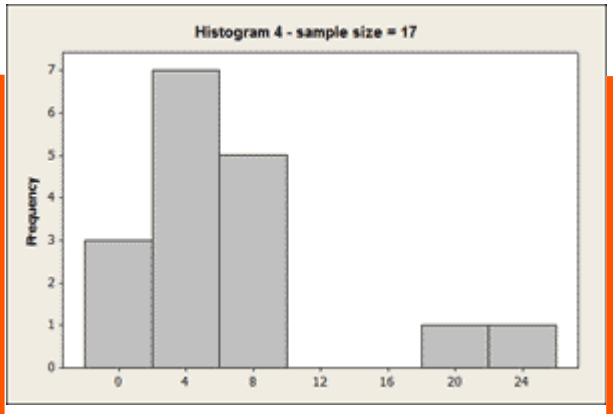
SELECT (A) Conditions are met; it is safe to proceed with the *t*-test.

SELECT (B) Conditions may not be met; conduct the *t*-test, but include a disclaimer with the results.

SELECT (C) Conditions are not met; do not use the *t*-test.

(C) is Incorrect. You are right to say the conditions are not met. However, this is a borderline case, since the sample is not heavily skewed and does not contain outliers.

QUESTION 4:



(A) is Incorrect. The sample size is only 17 and the data are not normally distributed in the sample. There are outliers. This suggests that the variable in the population might be skewed. So the Central Limit Theorem does not apply. We cannot assume that the sample means are normally

(B) is Incorrect. The sample size is only 17 and the data are not normally distributed in the sample. There are outliers. This suggests that the variable in the population might be skewed. So the Central Limit Theorem does not apply. We cannot assume that the sample means are normally distributed.

SELECT (A) Conditions are met; it is safe to proceed with the *t*-test.

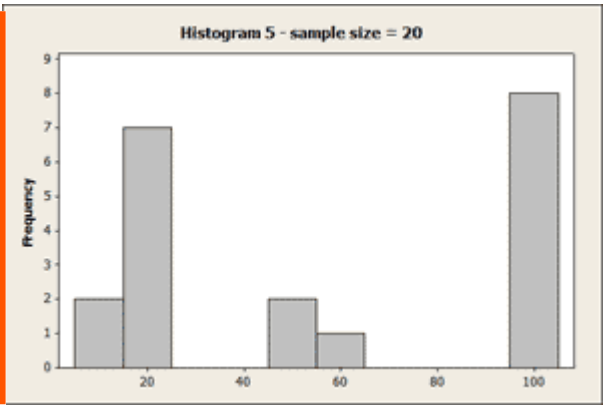
SELECT (B) Conditions may not be met; conduct the *t*-test but include a disclaimer with the results.

SELECT (C) Conditions are not met; do not use the *t*-test.

(C) is Correct. The sample size is only 17, and it contains outliers. This suggests that the variable in the population might be skewed. So the Central Limit Theorem does not apply. We cannot assume that the sample means are normally distributed.

QUESTION 5:

(A) is Incorrect. The sample size is only 20 and the data are not normally distributed in the sample. They are clustered in two humps. This suggests that the variable in the population is not normally distributed. So the Central Limit Theorem does not apply. We cannot assume that the sample means are normally distributed.



(B) is Incorrect. The sample size is only 20 and the data are not normally distributed in the sample. They are clustered in two humps. This suggests that the variable in the population is not normally distributed. So the Central Limit Theorem does not apply. We cannot assume that the sample means are normally distributed.

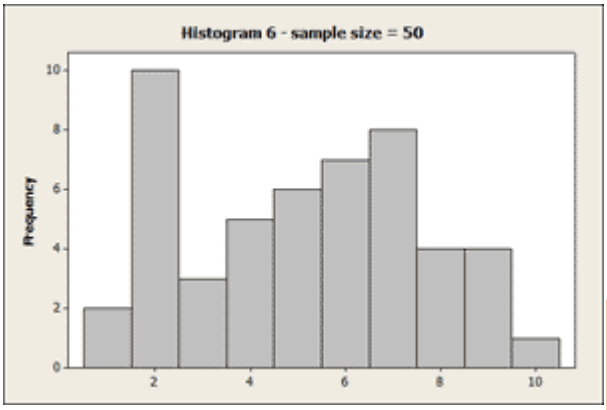
SELECT (A) Conditions are met; it is safe to proceed with the *t*-test.

SELECT (B) Conditions may not be met; conduct the *t*-test but include a disclaimer with the results.

SELECT (C) Conditions are not met; do not use the *t*-test.

(C) is Correct. The sample size is only 20 and the data are not normally distributed in the sample. This suggests that the variable in the population is not normally distributed. So the Central Limit Theorem does not apply. We cannot assume that the sample means are normally distributed.

QUESTION 6:



(A) is Correct. The sample size is greater than 30, so sample means will be normally distributed. Therefore, the t-test can be used.

(B) is Incorrect. Look at the size of the sample and recheck the conditions for the t-test given in the table.

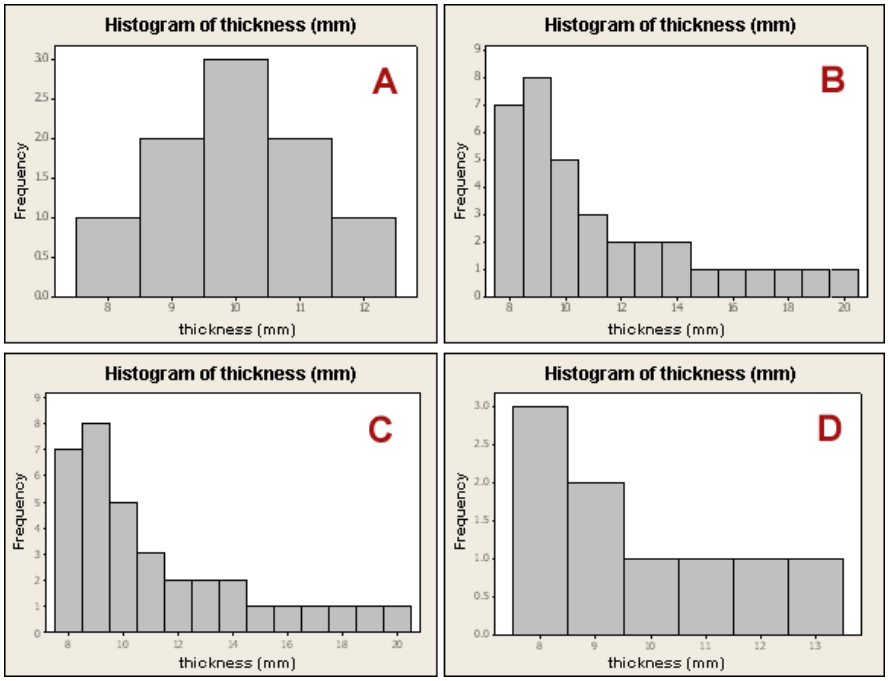
SELECT (A) Conditions are met; it is safe to proceed with the *t*-test.

SELECT (B) Conditions may not be met; conduct the *t*-test but include a disclaimer with the results.

SELECT (C) Conditions are not met; do not use the *t*-test.

(C) is Incorrect. Look at the size of the sample and recheck the conditions for the t-test given in the table.

QUESTION 7: Suppose that Intel is testing a brand new manufacturing process, for which prior information wasn't available. In particular, for this new process, the population distribution's shape isn't known. Use the following histograms to help you answer the question below.



To test whether or not the mean circuit board thickness is 12 mm with the new process, for which one of the following would the t-test NOT be justified?

SELECT (A) A random sample of 9 circuit boards is tested; the histogram of the data is as shown above in histogram A.

SELECT (B) A random sample of 40 circuit boards is tested; the histogram of the data is as shown above in histogram B.

SELECT (C) A random sample of 35 circuit boards is tested; the histogram of the data is as shown above in histogram C.

SELECT (D) A random sample of 9 circuit boards is tested; the histogram of the data is as shown above in histogram D.

(A) is Incorrect. In this case, although the sample size is relatively small (only 9 boards), notice that the population can be reasonably assumed to be normal (because the data distribution is reasonably unimodal and symmetric). So in this case, a t-test would be justified.

(B) is Incorrect. In this case, although the population cannot be assumed to be normal in shape (since the data distribution shows severe skewness), notice that the sample size is reasonably large. So in this case, a t-test would be justified.

(C) is Incorrect. In this case, although the population cannot be assumed to be normal in shape (since the data distribution shows severe skewness), notice that the sample size is greater than 30. So in this case, by the Central Limit Theorem, a t-test would be justified.

(D) is Correct. In this case, the sample size is relatively small (only 9 boards) and the population cannot be reasonably assumed to vary normally (because the data distribution is noticeably skewed). So in this case, the t-test would NOT be justified to conduct formal statistical inference. (A different sort of test called a "non-parametric" test might be justified)

LEARN BY DOING: State the Hypotheses for a test for a population mean

A research study measured the pulse rates of 57 college men and found a mean pulse rate of 70 beats per minute with a standard deviation of 9.85 beats per minute.

Researchers want to know if the mean pulse rate for all college men is different from the current standard of 72 beats per minute.

The null and alternative hypotheses in this case are: HINT

HINT: We want to test whether the mean pulse rate for college men is different from the current standard (72 beats per minute).

SELECT (A) Ho: $\mu = 70$, and Ha: $\mu < 70$

SELECT (B) Ho: $\mu = 72$, and Ha: $\mu < 72$

SELECT (C) Ho: $\mu = 70$, and Ha: $\mu > 70$

SELECT (D) Ho: $\mu = 72$, and Ha: $\mu > 72$

SELECT (E) Ho: $\mu = 70$, and Ha: $\mu \neq 70$

SELECT (F) Ho: $\mu = 72$, and Ha: $\mu \neq 72$

(A) is Incorrect. 70 is the sample mean pulse rate, and not the null value. Also, note that we want to test whether the mean pulse rate for college men is different from the current standard.

(B) is Incorrect. While it is true that the null value is 72, note that we want to test whether the mean pulse rate for college men is different from the current standard.

(C) is Incorrect. 70 is the sample mean pulse rate, and not the null value. Also, note that we want to test whether the mean pulse rate for college men is different from the current standard.

(D) is Incorrect. While it is true that the null value is 72, note that we want to test whether the mean pulse rate for college men is different from the current standard.

(E) is Incorrect. While the direction of the alternative hypothesis is correct (using $\neq$), 70 is the sample mean pulse rate, and not the null value.

(F) is Correct! We want to test whether the mean pulse rate for college men is different from the current standard (72 beats per minute).

LEARN BY DOING: Checking Conditions for Hypothesis Testing for the Population Mean

The purpose of this activity is to discuss how in some cases exploratory data analysis can help you determine whether the conditions that allow us to use the *t*-test for the population mean (μ) are met.

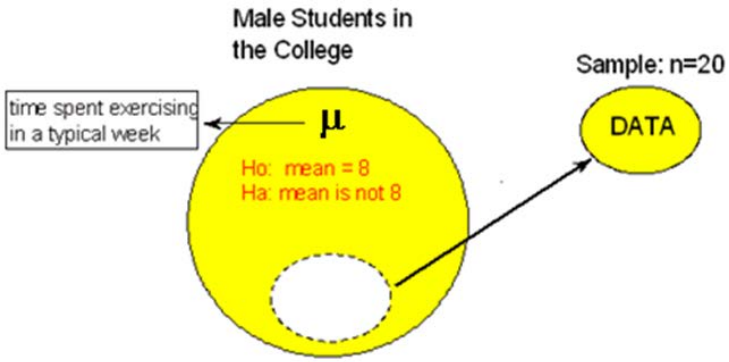
Background:

- In the Exploratory Data Analysis unit, we stressed that in general, it is always a good idea to **look at your data** (if the actual data are given).
- Moreover, related to our discussion now, looking at the data can be very helpful when trying to determine whether you can reliably use the test.
- Often in courses, data summaries (sample size, sample mean) are given rather than the raw data. We must trust what we are told in the problem. In practice, you often work with the raw data.

Now imagine the following situation:

- A health educator at a small college wants to determine whether the exercise habits of male students in the college are similar to the exercise habits of male college students in general.
- The educator chooses a random sample of 20 male students and records the time they spend exercising in a typical week.
- Do the data provide evidence that the mean time male students in the college spend exercising in a typical week differs from the mean time for male college students in general (which is 8 hours)?

NOTE: Whether σ is known or not is really not relevant to this activity.



Here is a situation in which we do not have any information about whether the variable of interest, "time" (time spent exercising in a typical week) varies normally or not, **and** the sample size (n = 20) is not really large enough for us to be certain that the Central Limit Theorem applies.

Recall from our discussion on the Central Limit Theorem that unless the distribution of "time" is extremely skewed and/or has extreme outliers, a sample of size 20 should be fine.

However, how can we be sure that is, indeed, the case?

- **If only the data summaries are given, there is really not a lot that can be done. You can say something like: "I'll proceed with the test assuming that the distribution of the variable "time" is not extremely skewed and does not have extreme outliers."**

- **If the actual data are given, you can make a more informed decision by looking at the data using a histogram. Even though the histogram of a sample of size 20 will not paint the exact picture of how the variable is distributed in the population, it could give a rough idea.**

Now we will look at a few different samples representing data in this scenario.

The data are available if you wish to create the graphs yourself but we also provide the graphs directly if you wish to skip analyzing the data yourself for this activity.

- **DATA:** EXCEL format, CSV format
- **OUTPUT:** GRAPHS (PDF)

For each sample, view the histogram and QQ-plot (or create them yourself). Comment on whether you think it would be safe to proceed with the test had those been the actual data in the problem above.

CHECK ANSWER

OUR ANSWER

Time 1—The histogram displays a roughly normal shape. For a sample of size 20, the shape is definitely normal enough for us to assume that the variable varies normally in the population and therefore it is safe to proceed with the test.

Time 2—The histogram displays a distribution that is slightly skewed and does not have any outliers. The histogram, therefore, does not give us any reason to be concerned that for a sample of size 20 the Central Limit Theorem will not kick in. We can therefore proceed with the test.

Time 3—The distribution does not have any "special" shape, and has one small outlier which is not very extreme (although it is arguable whether you would classify it as an outlier). Again, the histogram does not give us any reason to be concerned that for a sample of size 20 the Central Limit Theorem will not kick in. We can therefore proceed with the test.

Time 4—The distribution is extremely skewed to the right, and has one pretty extreme high outlier. Based on this histogram, we should be cautious about proceeding with the test, because assuming that this histogram "paints" at least a rough picture of how the variable varies in the population, a sample of size 20 might not be large enough for the Central Limit Theorem to kick in and ensure that x-bar has a normal distribution.

LEARN BY DOING: Calculate the Test Statistic for a Test for a Population Mean

Normal body temperature for healthy, at-rest human beings has always been said to be 98.6°F.

A doctor has seen a lot of patients who had a lower or higher body temperature when they were not ill.

So, he collected 50 randomly selected temperatures that had a mean of 98.4°F with a standard deviation of 0.35°F and wants to test the hypotheses.

- Ho: $\mu = 98.6$
- Ha: $\mu \neq 98.6$

Calculate the test statistic for this sample.

SELECT (A)	-0.34	(A) is Incorrect. Please review the formula and try again.
SELECT (B)	-0.57	(B) is Incorrect. Did you remember to use the sample size in your calculation?
SELECT (C)	-2.39	(C) is Incorrect. Did you take the square root of the entire denominator? (It is incorrect to do so.)
SELECT (D)	-4.04	(D) is Correct. The test statistic is (98.4 - 98.6) / (0.35 / sqrt(50)).
SELECT (E)	4.04	(E) is almost correct. Be sure to subtract the null value from the sample mean.

LEARN BY DOING: Hypothesis Testing for the Population Mean

The purpose of this activity is to give you guided practice in the process of a *t*-test for the population mean.

Background:

A group of 75 college students from a certain liberal arts college were randomly sampled and asked about the number of alcoholic drinks they have in a typical week. The file containing the data is linked below. The purpose of this study was to compare the drinking habits of the students at the college to the drinking habits of college students in general.

In particular, the dean of students, who initiated this study, would like to **check whether the mean number of alcoholic drinks that students at his college have in a typical week differs from the mean of U.S. college students in general, which is estimated to be 4.73.**

Let μ be the mean number of alcoholic beverages that students in THIS college drink in a typical week.

Answer the following questions using this SPSS OUTPUT (PDF), SAS OUTPUT (PDF), this SAS CODE

The raw data are also provided if you are interested in looking at it yourself: EXCEL or CSV

LEARN BY DOING: Hypothesis Testing for the Population Mean

The purpose of this activity is to give you guided practice in the process of a *t*-test for the population mean.

Background:

A group of 75 college students from a certain liberal arts college were randomly sampled and asked about the number of alcoholic drinks they have in a typical week. The file containing the data is linked below. The purpose of this study was to compare the drinking habits of the students at the college to the drinking habits of college students in general.

In particular, the dean of students, who initiated this study, would like to **check whether the mean number of alcoholic drinks that students at his college have in a typical week differs from the mean of U.S. college students in general, which is estimated to be 4.73.**

Let μ be the mean number of alcoholic beverages that students in THIS college drink in a typical week.

QUESTION 1:

State the hypotheses that are being tested in this problem.

CHECK ANSWER

We would like to test whether or not μ is the same as the national mean (4.73) therefore we are testing:
Ho: μ = 4.73
Ha: μ ≠ 4.73

QUESTION 2:

a. Look at the data using a histogram and describe the shape of the distribution.

b. Why are the conditions allowing us to safely use the *t*-test are met regardless of your findings in (a)?

CHECK ANSWER

a. Review output to see histogram. The distribution of is extremely skewed right and has high outliers.
b. The conditions are met since:
(i) The sample is random and (ii) The sample size is large enough (n = 75) for the Central Limit Theorem to apply and ensure that sample means vary normally, despite the fact that the distribution of the variable of interest (number of alcoholic drinks per week) is extremely skewed right, with high outliers.

QUESTION 3:

State the test statistic, interpret its value and show how it was found.

CHECK ANSWER

x-bar = 3.93 and s = 3.775. We already know that μo = 4.73 and n = 75 from reading the problem, so we can calculate: t = (3.93 - 4.73) / (3.775 / SQRT(75)) = -1.835
From the output we find t = -1.828. The difference is rounding error and the output value is more accurate. What this tells us is that the sample mean is 1.83 standard errors below the null value.

QUESTION 4:

Based on the p-value, draw your conclusions in context.

CHECK ANSWER

From the output provided, the p-value is 0.072, which at the 0.05 significance level indicates that the results are not significant. The data, therefore, do not provide enough evidence to reject Ho.

There is NOT ENOUGH evidence that the mean number of alcoholic drinks students at the college consume in a typical week is different from 4.73, the mean of college students in general.

QUESTION 5:

What would your conclusions be if the dean of students suspected that the mean number of alcoholic drinks that students in the college consume in a typical week is lower than the mean of U.S. college students in general? In other words, if this were a test of the hypotheses:

H0: μ = 4.73 drinks per week

Ha: μ < 4.73 drinks per week

HINT: Consider how a one-sided versus a two-sided test affects the p-value.

HINT

CHECK ANSWER

In this case, we would test:
H0: μ = 4.73
Ha: μ < 4.73
The p-value for this test would be 0.072/2 = 0.036 (half that of the two-sided test). At the 0.05 significance level we would therefore reject H0 and conclude that the data provide evidence that the mean number of alcoholic drinks that students in the college consume in a typical week is indeed lower than the mean of U.S. college students in general (as the dean suspected).

QUESTION 6:

Now suppose that instead of the 75 students having been randomly selected from the entire student body, the 75 students had been randomly selected only from the engineering classes at the college (for the sake of convenience).

Address the following two issues regarding the effect of such a change in the study design:

a. Would we still be mathematically justified in using the t-test for obtaining conclusions, as we did previously?

b. Would the resulting conclusions still address the question of interest (which, remember, was to investigate the drinking habits of the students at the college as whole)?

CHECK ANSWER

OUR ANSWER

a. Yes, the t-procedure is still mathematically justified, since the sample size is still 75 (and inference is still justified because the sample is still random).

b. No, the resulting conclusion would no longer address the question of interest, because the sampling frame (all students in the engineering classes) no longer matches the population of interest (all students in the entire college). The inferential conclusion from the t-test would only allow us to make conclusions about students in the engineering classes (i.e., about the group from which the sample was randomly selected). Art majors or humanities majors, for instance, probably don't take engineering and therefore would not have been represented in the (altered) study. So the conclusions from this (altered) study design would not be valid for the college as a whole.