9.5: Hypothesis Tests for One Population Mean When σ is Unknown

In practice, when we are using a sample to make inferences about the population mean, it is rare for us to know the population standard deviation.

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

Instead, we must use the sample standard deviation, *s*, as a point estimate of the population standard deviation, σ.

When using *s* as an estimate for σ, we cannot use a *z*-test, because $\frac{\bar{x}-\mu_{\bar{x}}}{\frac{s}{\sqrt{n}}}$ is not normally distributed.

$z = \frac{\bar{x}-\mu_{\bar{x}}}{\sigma_{\bar{x}}}$

The *t*-test for one population mean:

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

When using *s* as an estimate for σ, we use the Student *t*-distribution.

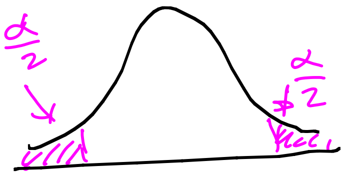
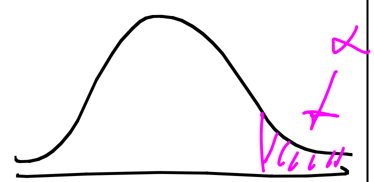
In order to use this procedure, we need to know (or be able to reasonably assume) that the variable of interest follows a normal distribution, or we must have a large sample size ($n \geq 30$).

In addition, the sample should be randomly obtained, observations within the sample must be independent of one another. This means that if we have a sample size that is more than 5% of the population, we should multiply the standard error by a finite population correction factor, $\sqrt{\frac{N-n}{n-1}}$. (In this class, I do not anticipate that we will encounter this situation.)

<u>Hypothesis Testing for a Population Mean:</u>

<u>Step 1</u>: Determine the significance level α .

<u>Step 2</u>: Determine the null and alternative hypotheses.

Two-Tailed Test (most common)	Left-Tailed Test (rare)	Right-Tailed Test (rare)
$H_0 : \mu = \mu_0$ $H_1 : \mu \neq \mu_0$	$H_0 : \mu = \mu_0$ $H_1 : \mu < \mu_0$	$H_0 : \mu = \mu_0$ $H_1 : \mu > \mu_0$
 Rejection Region	 Rejection Region	 Rejection Region

Note: One tailed tests assume that the scenario not listed ($\mu > \mu_0$ for a left-tailed test or $\mu < \mu_0$ for a right-tailed test) is not possible or is of zero interest.

<u>Step 3</u>: Use your α level and hypotheses, sketch the rejection region.

<u>Step 4</u>: Compute the test statistic $t = \dfrac{\overline{x} - \mu}{\dfrac{s}{\sqrt{n}}}$.

<u>Step 5</u>: Use a table (T~~able IV, on page A-13~~) to determine the <u>critical value for *t*</u> associated with your rejection region. A-6, A-7

<u>Step 6</u>: Determine whether the value of *t* calculated from your sample (in Step 3) is in the rejection region.

- If *t* is in the rejection region, reject the null hypothesis.
- If *t* is not in the rejection region, do not reject the null hypothesis.

<u>Step 7</u>: State your conclusion.

Example 1: The normal human body temperature is widely accepted to be 98.6° F and can be assumed to follow a normal distribution. A medical researcher wants to know whether a certain geographical community of Native Alaskans has a mean body temperature of 98.6° F. A sample of 20 members of the Native Alaskan geographical community resulted in a mean body temperature of 98.3° F with a standard deviation of 0.7° F. Perform an appropriate hypothesis test at the 95% confidence level.

Sample info:

$n = 20$

$\bar{x} = 98.3°F$

$s = 0.7°F$

$df = n - 1 = 19$

From table, critical value of t is

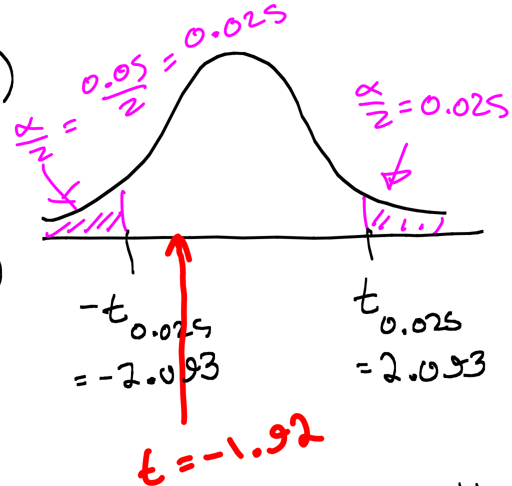
$t_{0.025} = 2.093$

$H_0: \mu = 98.6°F$

$H_a: \mu \neq 98.6°F$ (2-tailed)

$\alpha = 1 - 0.95 = 0.05$

assumptions met? Yes (temp. is normally distributed)



Calculate t for our sample:

Std error: $\sigma_{\bar{x}} \approx \frac{s}{\sqrt{n}} = \frac{0.7}{\sqrt{20}} \approx 0.1565$

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{98.3 - 98.6}{0.1565} = -1.92$$

This is not in the rejection region, so $\boxed{\text{Do not Reject } H_0}$.

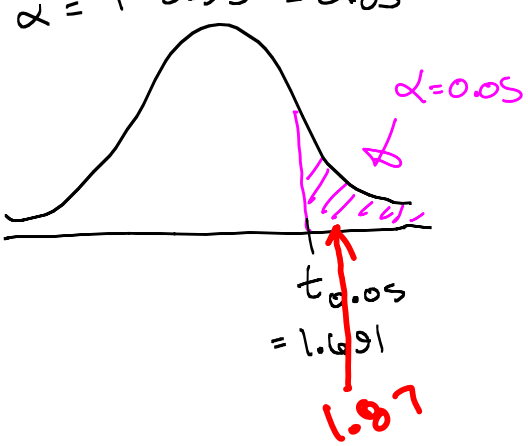
Conclusion: This sample does not provide evidence that the temp. differs from 98.6°F

Example 2: The average amount of lead in the blood of young children is 2 micrograms per deciliter (mcg/dL). A city has recently changed its water supply, and there have been widespread reports of increased lead levels in the water. A concerned doctor wants to dig into the city's medical records to find out whether the children in the city have blood lead levels above the average level of 2 mcg/dL. In a sample of 35 children, she found a mean lead level of 2.60 mcg/dL with a standard deviation of 1.9 mcg/dL. Perform an appropriate hypothesis test at the 95% confidence level.

$H_0: \mu = 2 mcg/dL$

$H_a: \mu > 2 mcg/dL$

$\alpha = 1 - 0.95 = 0.05$



Sample info

$n = 35$

$\bar{x} = 2.60$

$s = 1.9$

$df = n - 1 = 34$

assumptions met? Yes, $n \geq 30$

From table, critical value is $t_{0.05} = 1.691$

Calculate t for our sample

std error: $\sigma_{\bar{x}} \approx \frac{s}{\sqrt{n}} = \frac{1.9}{\sqrt{35}} \approx 0.32116$

$$t = \frac{\bar{x} - \mu_0}{\sigma_{\bar{x}}} = \frac{2.60 - 2}{0.32116} = 1.87$$

t is in the rejection region, so $\boxed{\text{Reject } H_0}$.

This sample provides evidence that the population of kids have avg lead level above 2.

This sample provides evidence that the pop. mean differs from 8 oz.

Example 3: Suppose a manufacturer claims on the label that a package contains 8 ounces of potato chips. A customer (or a FDA analyst) buys 50 bags of chips, weighs them on a high-accuracy scale, and obtains a sample mean of 7.89 ounces with a sample standard deviation of 0.2 ounces. Does this sample provide evidence that the manufacturer's labeling may be inaccurate? Use the $\alpha = 0.10$ level of significance.

Sample info

$n = 50$

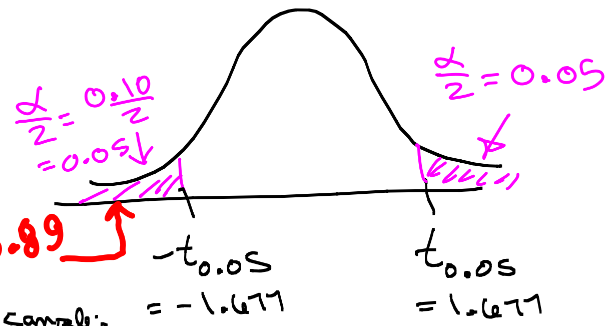
$\bar{x} = 7.89\text{oz}$

$s = 0.20$

$H_0: \mu = 8$

$H_a: \mu \neq 8$

$\alpha = 0.10$



Find critical value of t

$df = n - 1 = 49$

For $df = 49$, and tail area $= 0.05$, critical value is $t_{0.05} = 1.677$

Calculate t for our sample:

Std error: $\frac{s}{\sqrt{n}} = \frac{0.2}{\sqrt{50}} \approx 0.002828$

Calculate t: $t = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{7.89 - 8}{0.02828} = -3.89$

This is in the rejection region, so we $\boxed{\text{Reject } H_0}$

Example 4: Suppose a manufacturer claims on the label that a package contains 8 ounces of potato chips. Again, a customer (or a FDA analyst) wonders whether the package size is accurate. This time, the analysist only buys 10 bags of chips, and obtains a sample mean of 7.89 ounces with a sample standard deviation of 0.2 ounces. Does this sample provide evidence that the manufacturer's labeling may be inaccurate? Use the $\alpha = 0.10$ level of significance.