Stat 201:
Introduction to Statistics

Standard 28: Significance Test - Means

Confidence Intervals to Testing

- We've seen earlier that we can come up with interesting observations of our confidence intervals

- Next we will learn how to formally test whether or not the population mean is a particular value based off our sample mean

Hypothesis Test for Means: Step 1

- **State Hypotheses:** it's usually easier to write the alternative hypothesis first
  - **Null hypothesis**: that the population mean equals some μ_o
    - $H_o\colon \mu \leq \mu_o$ (one sided test)
    - $H_o\colon \mu \geq \mu_o$ (one sided test)
    - $H_o\colon \mu = \mu_o$ (two sided test)
  - **Alternative hypothesis:** What we're interested in
    - $H_a\colon \mu > \mu_o$ (one sided test)
    - $H_a\colon \mu < \mu_o$ (one sided test)
    - $Ha\colon \mu \neq \mu_o$ (two sided test)

Hypothesis Test for Means: Step 2

- Check the assumptions
  - The variable must be quantitative
  - The data are obtained using randomization
  - We're dealing with data from the normal distribution
    - If n>30
    - If a histogram of the data is approximately normal which indicates that the probability is normal

Hypothesis Test for Means: Step 3

- **<u>When we don't know $\boldsymbol{\sigma_x}$</u>**
- Calculate Test Statistic, t*
  – The test statistic measures how different the sample mean we have is from the null hypothesis
  – We calculate the t-statistic by assuming that μ_0 is the population mean

$$t^* = \frac{(\bar{x} - \mu_o)}{\frac{s_x}{\sqrt{n}}}$$

Hypothesis Test for Means: Step 3

- **When we know $\boldsymbol{\sigma_x}$**
- Calculate Test Statistic, z*
  - The test statistic measures how different the sample mean we have is from the null hypothesis
  - We calculate the t-statistic by assuming that μ_0 is the population mean

$$z = \frac{(\bar{x} - \mu_o)}{\frac{\sigma_x}{\sqrt{n}}}$$

Hypothesis Test for Means: Step 4

- **When we don't know σ_x**
- Determine the P-value
  - The P-value describes how unusual the data would be if H_o were true.
  - We will use software or your calculator to find this, or I will give it to you.

Alternative Hypothesis	Probability	Formula for the P-value
$H_a: \mu > \mu_o$	Right tail	P(T>t*)=1-P(T<t*)
$H_a: \mu < \mu_0$	Left tail	P(T<t*)
$H_a: \mu \neq \mu_o$	Two-tail	2*P(T<-|t*|)

Hypothesis Test for Proportions: Step 4

- **<u>When we know σ_x</u>**
- **Determine the P-value**
  - The P-value describes how unusual the sample data would be if H_o were true, which is what we're assuming($\mu = \mu_0$).
  - z* is the test statistic from step 3

Alternative Hypothesis	Probability	Formula for the P-value
$H_a: \mu > \mu_o$	Right tail	P(Z>z*)=1-P(Z<z*)
$H_a: \mu < \mu_0$	Left tail	P(Z<z*)
$H_a: \mu \neq \mu_o$	Two-tail	2*P(Z<-|z*|)

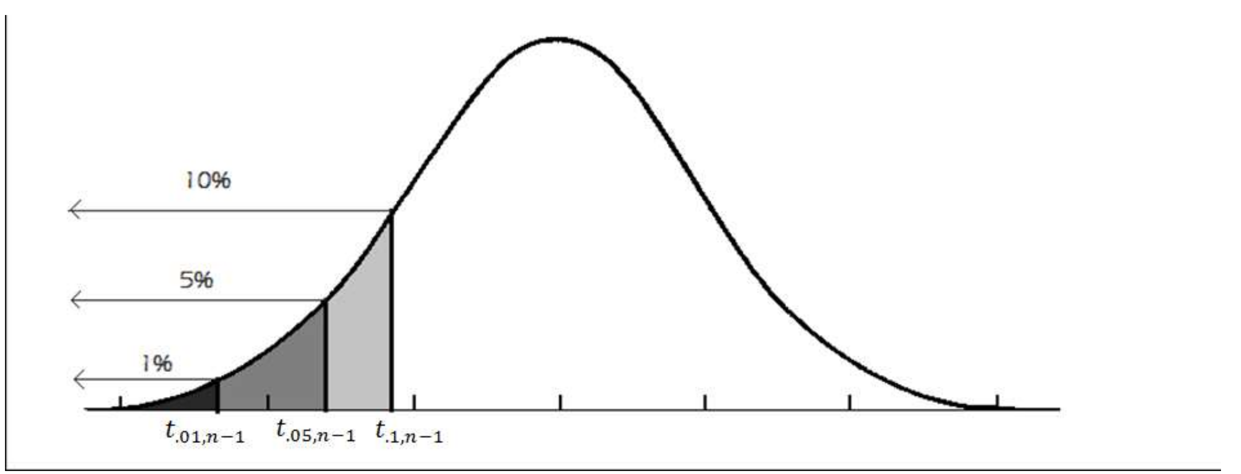
Hypothesis Test for Means: Step 5

- Summarize the test by reporting and interpreting the P-value
  - Smaller p-values give stronger evidence against H_o
- If p-value$\leq (1 - confidence) = \alpha$
  - Reject H_o, with a p-value = ____, we have sufficient evidence that the alternative hypothesis might be true
- If p-value$> (1 - confidence) = \alpha$
  - Fail to reject H_o, with a p-value = ____, we do not have sufficient evidence that the alternative hypothesis might be true

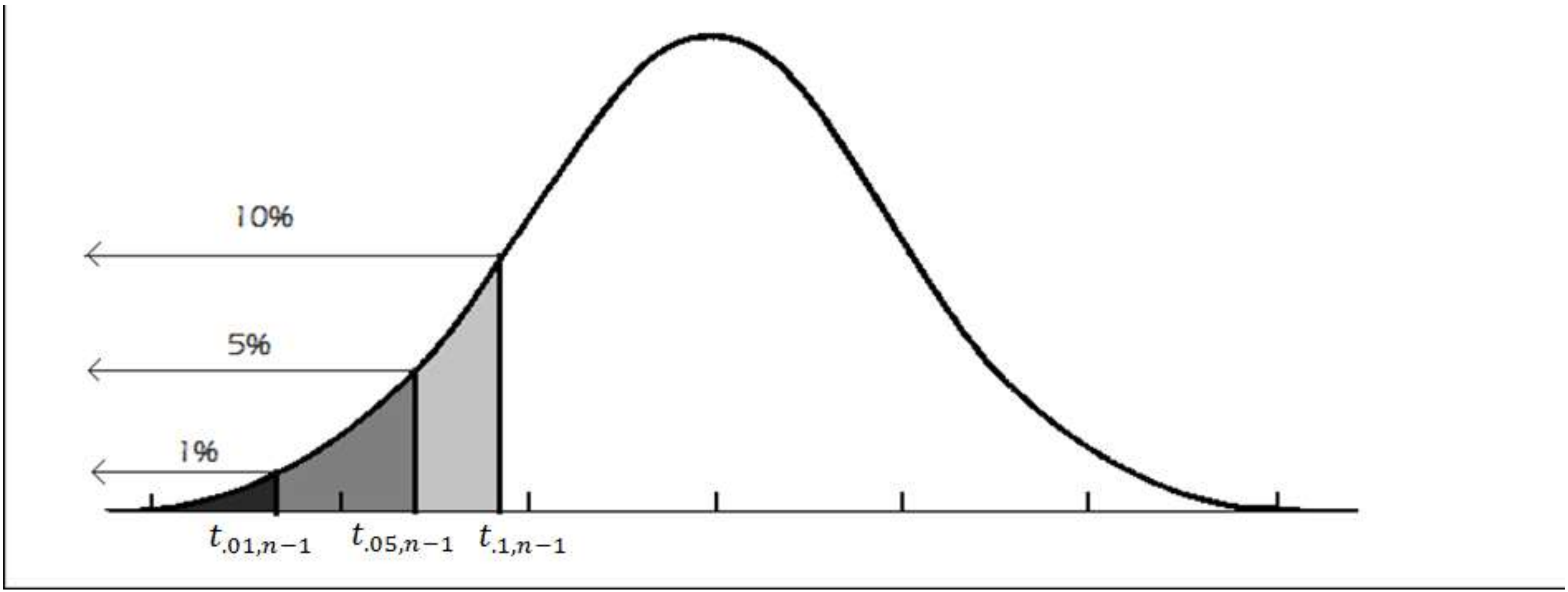
Hypothesis Test for Means– Step Five with Pictures

- For a left tailed test: $H_a: \mu < \mu_o$ → We have rejection regions for H_o are as follows

Confidence	Reject (test stat)	Reject (p-value)
0.90	Test-stat<-$t_{.10,n-1}$	P-value<.1
0.95	Test-stat<-$t_{.05,n-1}$	P-value<.05
0.99	Test-stat<-$t_{.01,n-1}$	P-value<.01



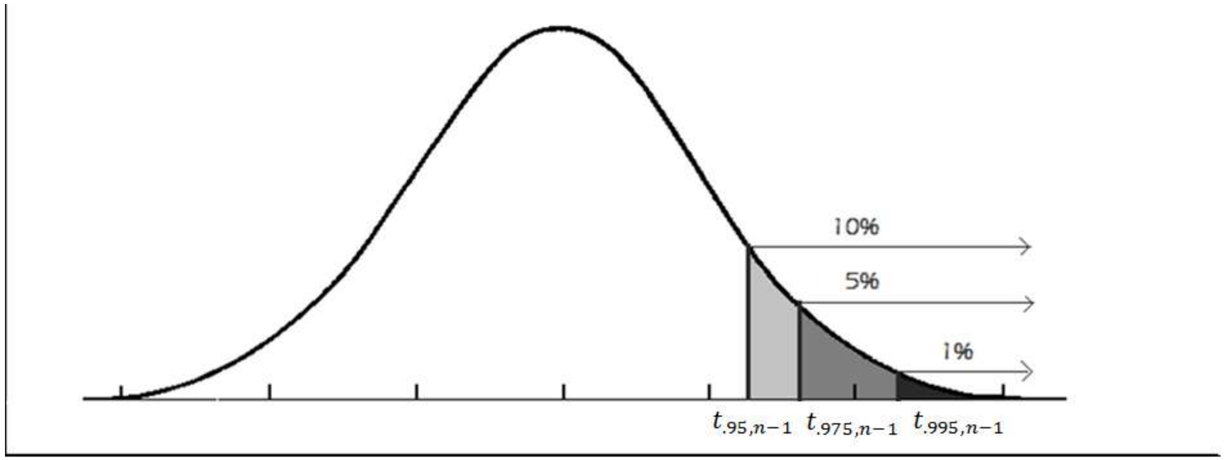
Zoom In



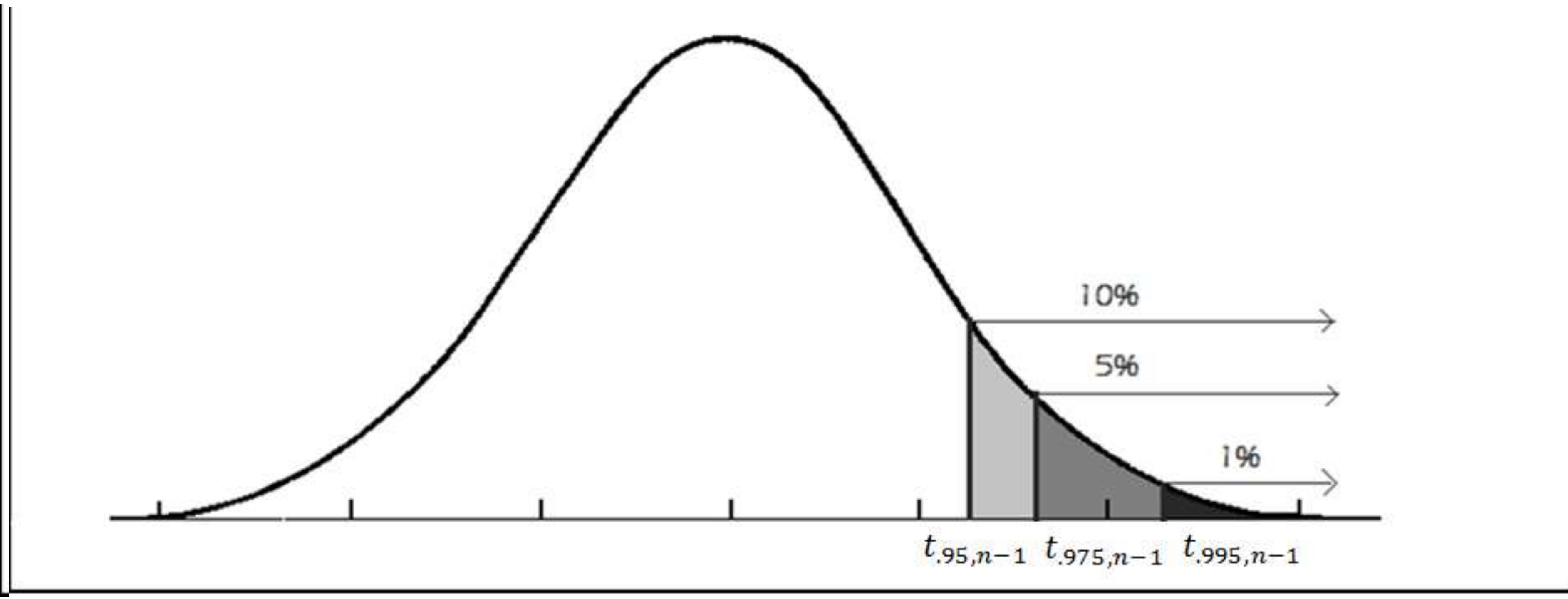
Hypothesis Test for Means– Step Five with Pictures

- For a left tailed test: $H_a: \mu > \mu_o \rightarrow$ We have rejection regions for H_o are as follows

Confidence	Reject (test stat)	Reject (p-value)
0.90	Test-stat<-$t_{.90,n-1}$	P-value<.1
0.95	Test-stat<-$t_{.95,n-1}$	P-value<.05
0.99	Test-stat<-$t_{.99,n-1}$	P-value<.01



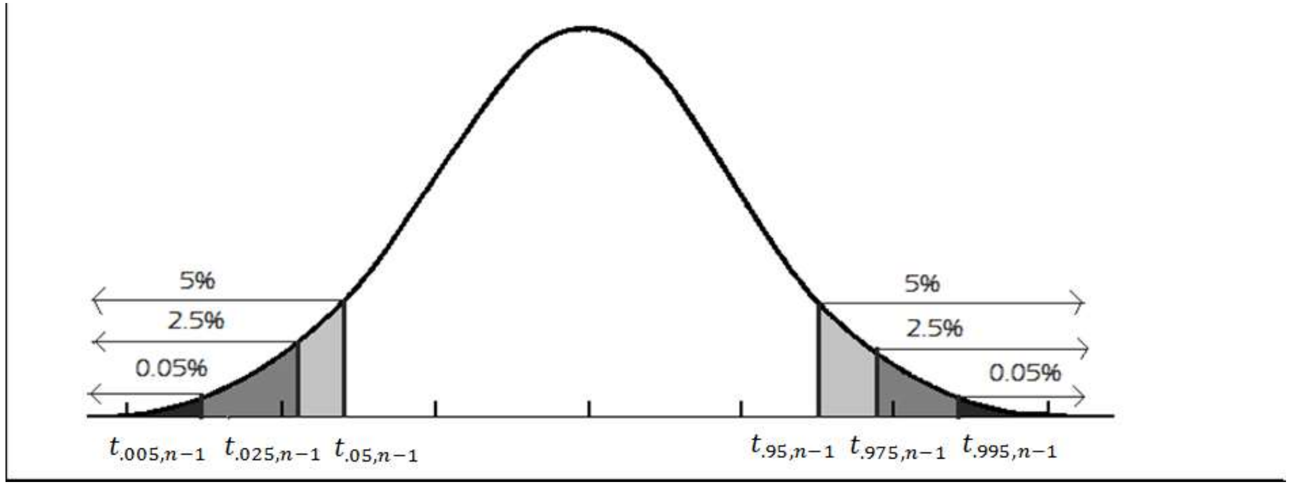
Zoom In



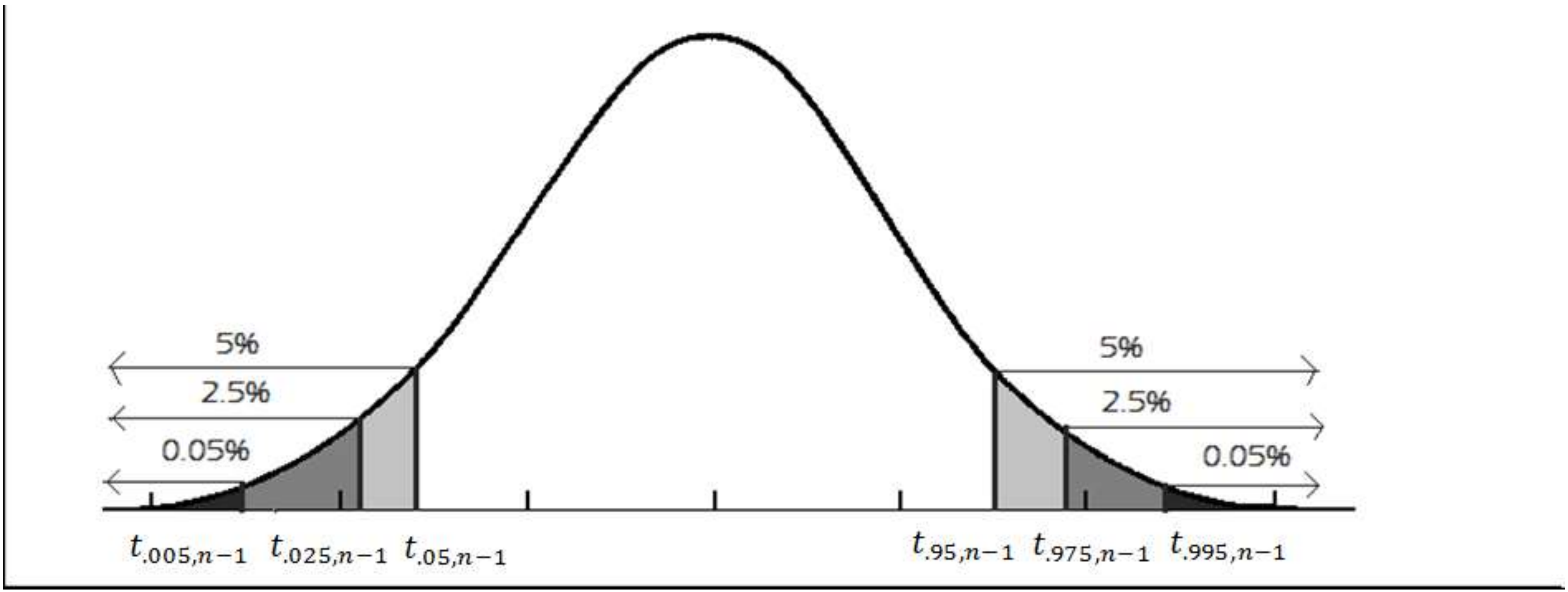
Hypothesis Test for Means– Step Five with Pictures

- For a two tailed test: $H_a: \mu \neq \mu_o$ → We have rejection regions for H_o are as follows

Confidence	Reject (test stat)	Reject (p-value)
0.90	|Test-stat|<-$t_{.90,n-1}$	P-value<.1
0.95	|Test-stat|<-$t_{.95,n-1}$	P-value<.05
0.99	|Test-stat|<-$t_{.99,n-1}$	P-value<.01



Zoom In



Hypothesis Test for Means– Step Five

- The pictures are the same when we know z as they are for proportions.

- In almost all feasible cases we will not know $\boldsymbol{\sigma_x}$ as is t is usually unrealistic to know it

Example 1

- It is often hypothesized that Velociraptors were warm blooded creatures, some scientists guessed their normal blood temperature was 87.5 degrees. Test whether or not the mean differs from 87.5 degrees at a .05 significance level, or 95% confidence.
- A random sample of thirteen Velociraptors during the shooting of Jurassic Park gave the data below

88.6, 86.4, 87.2, 87.4, 87.2, 87.6, 86.8, 86.1, 87.4, 87.3, 86.4, 86.6, 87.1

Example 1 – Step One

- A random sample of thirteen Velociraptors during the shooting of Jurassic Park gave the data below

88.6, 86.4, 87.2, 87.4, 87.2, 87.6, 86.8, 86.1, 87.4, 87.3, 86.4, 86.6, 87.1

- State the Hypotheses: we are interested in whether or not the mean is **not equal to 87.5 degrees**
  - H_o: $\mu = 87.5$
  - H_a: $\mu \neq 87.5$

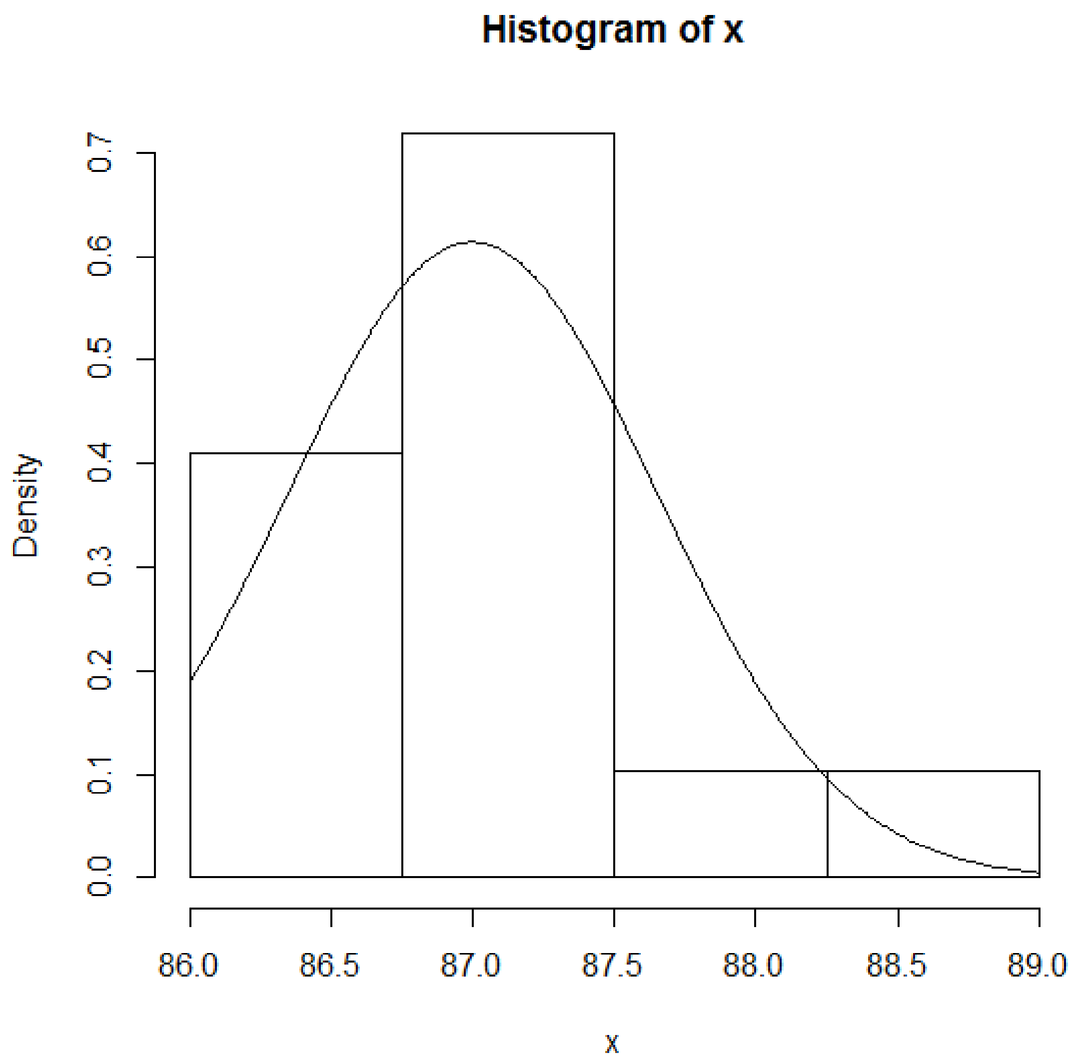
Example 1 – Step Two

- A random sample of thirteen Velociraptors during the shooting of Jurassic Park gave the data below

88.6, 86.4, 87.2, 87.4, 87.2, 87.6, 86.8, 86.1, 87.4, 87.3, 86.4, 86.6, 87.1

- Check Assumptions:
  - The data is quantitative
  - The sample is randomly selected
  - Even though $n<30$, a histogram of the data shows approximately normal

See, I told you. (close enough for us)



Example 1 – Step Three

- Calculate Test Statistic

Variable	Sample Mean $(\overline{x})$	Standard Deviation (s_x)	Standard Error $(s_{\overline{x}})$
Blood Temperature	87.0846	.6492	.1800

$$t = \frac{(\bar{x} - \mu_o)}{\frac{s}{\sqrt{n}}} = \frac{87.08 - 87.5}{\frac{.6492}{\sqrt{13}}} = \frac{.42}{.1800} = -2.\overline{33}$$

Example 1 – Step Four

- Determine P-value

  P-value from software is .0397

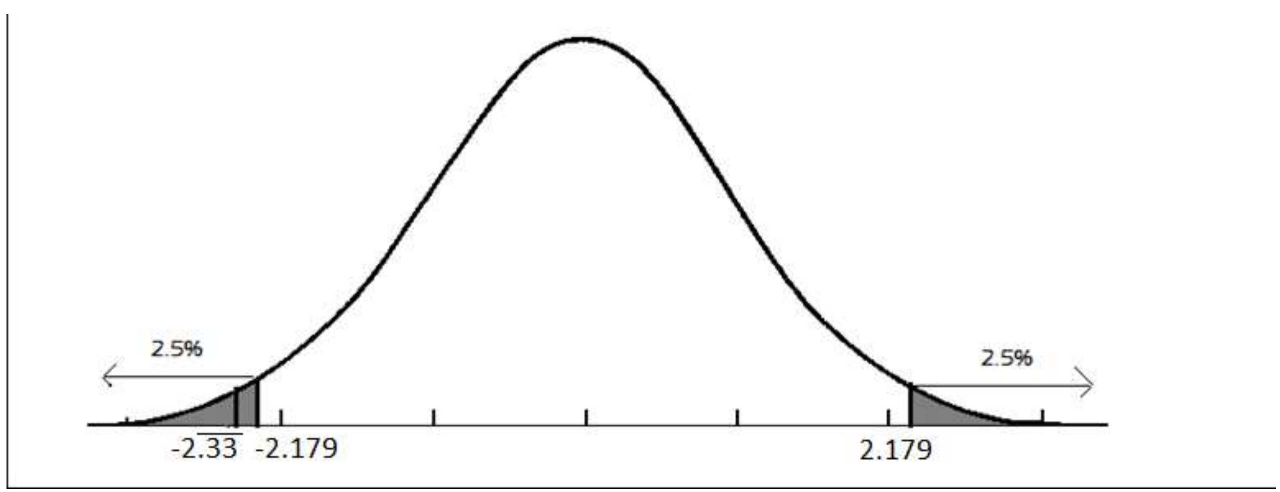
Example 1 – Step Five

- State Conclusion
  – Since .0397<.05 we reject H_o

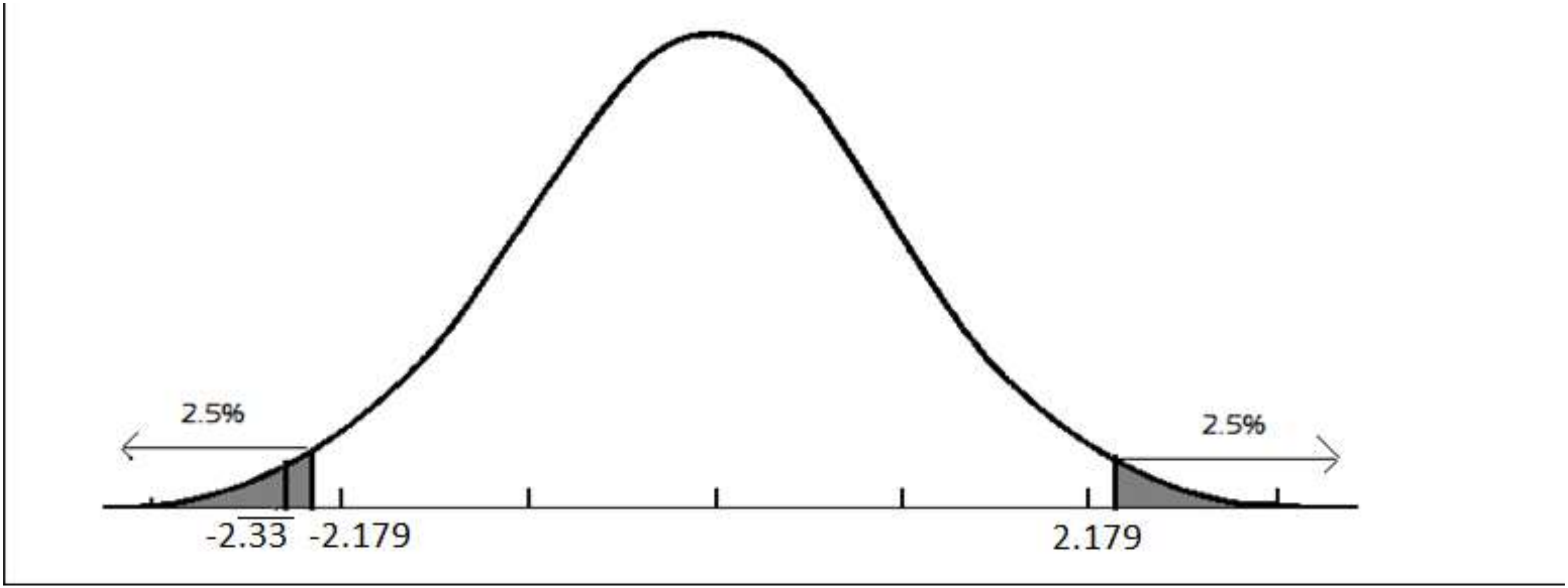
  At the .05 level of significance, or 95% confidence level, there is sufficient evidence that the mean blood temperature is different than 87.5.

Example 1 – Step Five with pictures

- State Conclusion
  - Anything with a p-value<.05 or a |t-value|>$t_{1-\frac{\alpha}{2},n-1}$ = $t_{.975,12}$ = 2.179 will be in the rejection region
  - Since .2932>.05 we fail to reject H_o



Zoom In



Example

- Suppose a random sample of 38 yearly average temperature measures in New Haven, CT. Among the sampled years the **sample mean temperature was 51.0474** degrees Fahrenheit with a **sample standard deviation of 1.3112.**
- Test whether or not the population mean differs from 50 degrees at a .05 significance level, or 95% confidence.

Example – Step One

- State the Hypotheses: we are interested in whether or not the mean is **not equal to 50 degrees**
  - H_o: $\mu = 50$
  - H_a: $\mu \neq 50$

Example – Step Two

- Check Assumptions:
  - The data is quantitative
  - The sample is randomly selected
  - $n>30$ so it is safe to assume the sampling distribution for the sample mean is normal

Example – Step Three

- Calculate Test Statistic

Variable	Sample Mean $(\bar{x})$	Standard Deviation (s_x)	Standard Error $(s_{\bar{x}})$
Yearly Temperature	51.0474	1.3112	.2127047

$$t = \frac{(\bar{x} - \mu_o)}{\frac{s}{\sqrt{n}}} = \frac{51.0474 - 50}{\frac{1.3112}{\sqrt{38}}} = \frac{1.0474}{.2127047} = 4.924198$$

Example – Step Four

- Determine P-value

$$
\begin{aligned}
P\ value &= 2 * P(T < -|t^*|) \\
&= 2 * P(T < -|4.924198\,|) \\
&= 2 * P(T < -4.924198) \\
&= .00001782519
\end{aligned}
$$

Example – Step Five

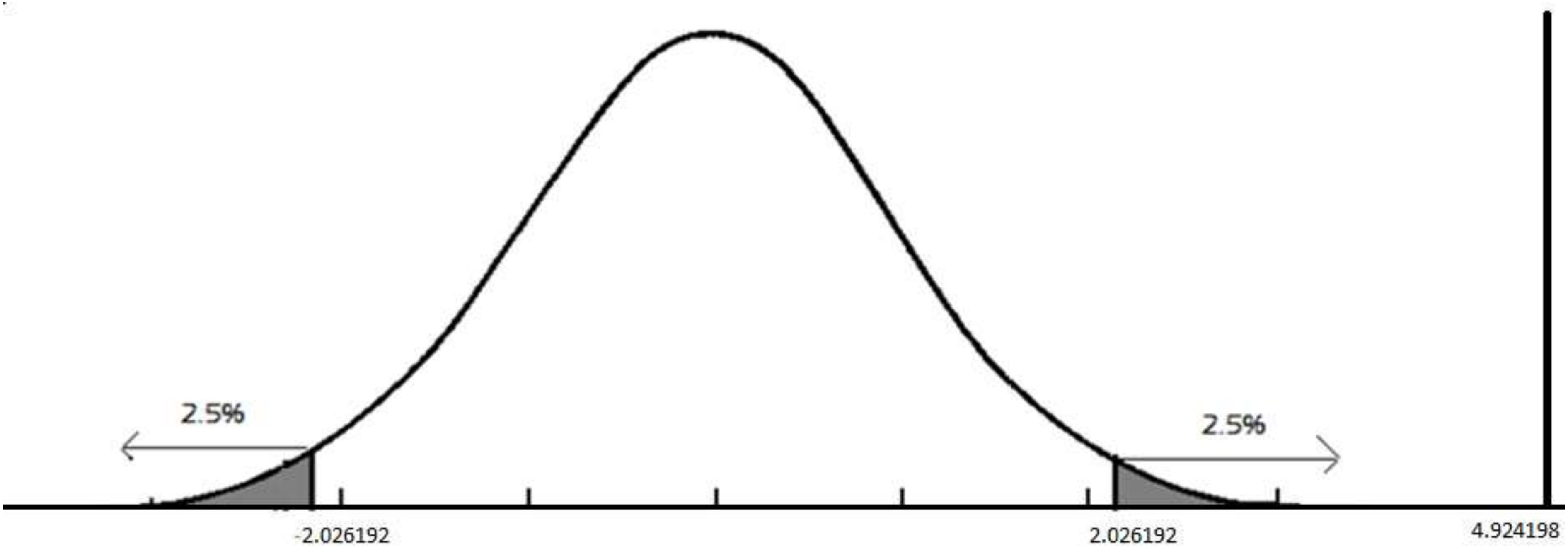
- State Conclusion
  - Since $.00001782519 < .05$ we reject H_o

  At the .05 level of significance, or 95% confidence level, there is sufficient evidence that the mean yearly temperature is different than 50 degrees.

Example – Step Five with pictures

- State Conclusion
  - Anything with a p-value<.05 or a |t-value|> $t_{1-\frac{\alpha}{2},n-1}$ = $t_{.975,37}$ = 2.026192 will be in the rejection region
  - By P-value:
    - Since $.00001782519$ >.05 we reject H_o
  - By T-statistic:
    - Since $|4.924198|$>2.026192 we reject H_o

Zoom In



Hypothesis Testing for Means in Your TI Calculator

- Hypothesis testing for means
  - https://www.youtube.com/watch?v=StpX5_AHKSs
  - https://www.youtube.com/watch?v=31fFfsSmuK8
  - https://www.youtube.com/watch?v=dyj0Mjvu_mQ

Hypothesis Testing for Means in Your TI Calculator

- **When we don't know σ_x, with data**
- **<u>INPUT:</u>**
  1. Press STAT
  2. Press → to TESTS
  3. Highlight '2: T-Test' and Press ENTER
  4. **With Data**
     1. Enter the we're interested in next to 'μ_0:'
     2. You should have your data table entered in L1
        - If you forgot: Press STAT, Press ENTER with 'Edit' highlighted, Enter the data into the L1 col.
     3. Next to 'List:' press 2^{nd} then press 1
     4. Set 'Frequency' to 1
     5. Select the appropriate alternative hypothesis on the 'μ:' line by highlighting the correct inequality and pressing ENTER
     6. Highlight Calculate and press ENTER

Hypothesis Testing for Means in Your TI Calculator

- **When we don't know σ_x, with data**
- **<u>Output:</u>**
  - Confirm the first line shows the hypothesis you would like to test
  - t = the test statistic for our hypothesis test
  - p = the p-value for this test
    - We make our decision based on this
  - $\bar{x}$ is the sample mean for the problem and should match the number you entered
  - s_x is the sample standard deviation for the problem
  - n is the sample size and should match the number you entered

Hypothesis Testing for Means in Your TI Calculator

- **When we don't know σ_x, with stats**
- **INPUT:**
  1. Press STAT
  2. Press → to TESTS
  3. Highlight '2: T-Test' and Press ENTER
  4. **With Stats**
     1. Enter the we're interested in next to 'μ_0:'
     2. Put the sample mean next to '$\bar{x}$:'
     3. Enter the sample standard deviation next to 's_x:'
     4. Put the sample size next to 'n:'
     5. Select the appropriate alternative hypothesis on the 'μ:' line by highlighting the correct inequality and press ENTER
     6. Highlight Calculate and press ENTER

Hypothesis Testing for Means in Your TI Calculator

- **When we don't know σ_x, with stats**
- **<u>Output:</u>**
  - Confirm the first line shows the hypothesis you would like to test
  - t = the test statistic for our hypothesis test
  - p = the p-value for this test
    - We make our decision based on this
  - $\bar{x}$ is the sample mean for the problem and should match the number you entered
  - s_x is the sample standard deviation for the problem
  - n is the sample size and should match the number you entered

Hypothesis Testing for Means in Your TI Calculator

- **When we know $\boldsymbol{\sigma_x}$, with data**
- **<u>INPUT:</u>**
    1. Press STAT
    2. Press → to TESTS
    3. Highlight '1: Z-Test' and Press ENTER
    4. **With Data**
        1. Enter the we're interested in next to 'μ_0:'
        2. Enter the population standard deviation next to 'σ:'
        3. You should have your data table entered in L1
            - If you forgot: Press STAT, Press ENTER with 'Edit' highlighted, Enter the data into the L1 col.
        4. Next to 'List:' press 2^{nd} then press 1
        5. Set 'Frequency' to 1
        6. Select the appropriate alternative hypothesis on the 'μ:' line by highlighting the correct inequality and pressing ENTER
        7. Highlight Calculate and press ENTER

Hypothesis Testing for Means in Your TI Calculator

- **When we know $\boldsymbol{\sigma_x}$, with data**
- **<u>Output:</u>**
  - Confirm the first line shows the hypothesis you would like to test
  - z = the test statistic for our hypothesis test
  - p = the p-value for this test
    - We make our decision based on this
  - $\bar{x}$ is the sample mean for the problem and should match the number you entered
  - n is the sample size and should match the number you entered

Hypothesis Testing for Means in Your TI Calculator

- **When we know σ_x, with stats**
- **<u>INPUT:</u>**
  1. Press STAT
  2. Press → to TESTS
  3. Highlight ‘2: T-Test’ and Press ENTER
  4. **<u>With Stats</u>**
     1. Enter the we’re interested in next to ‘μ_0:’
     2. Enter the population standard deviation next to ‘σ:’
     3. Put the sample mean next to ‘$\bar{x}$:’
     4. Put the sample size next to ‘n:’
     5. Select the appropriate alternative hypothesis on the ‘μ:’ line by highlighting the correct inequality and press ENTER
     6. Highlight Calculate and press ENTER

Hypothesis Testing for Means in Your TI Calculator

- **When we know σ_x, with data**
- **<u>Output:</u>**
  - Confirm the first line shows the hypothesis you would like to test
  - z = the test statistic for our hypothesis test
  - p = the p-value for this test
    - We make our decision based on this
  - $\bar{x}$ is the sample mean for the problem and should match the number you entered
  - n is the sample size and should match the number you entered

Confidence Intervals for Means

- **<u>StatCrunch Commands w/ data</u>**
  - Stat→T Stats→One Sample →with data (if you have the a list of data)→Choose the column→type the success value into the success box→ choose hypothesis→ enter the correct hypothesis→ Compute

- **<u>StatCrunch Commands w/ summaries</u>**
  - Stat→T Stats→One Sample →with summary (if you have the count) → enter the number of success and total observations→ enter the correct hypothesis→ Compute

Confidence Intervals known σ_x

Assumptions	Point Estimate	Margin of Error	Margin of Error
1. Random Sample *2.* $n > 30$ OR the population is bell shaped	$\bar{x}$	$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$	$\bar{x} \pm z_{\frac{\alpha}{2}}\left(\frac{\sigma_x}{\sqrt{n}}\right)$

- We are --% confident that the true population mean lays on the confidence interval.

Confidence Intervals unknown σ_x

Assumptions	Point Estimate	Margin of Error	Margin of Error
1. Random Sample *2.* $n > 30$ OR the population is bell shaped	$\bar{x}$	$\sigma_{\bar{x}} = \frac{s_x}{\sqrt{n}}$	$\bar{x} \pm t_{1-\frac{\alpha}{2},n-1}\left(\frac{s_x}{\sqrt{n}}\right)$

- We are --% confident that the true population mean lays on the confidence interval.

Hypothesis Testing known σ_x

Step One:	*1.* $H_0: \mu = \mu_0$ & $H_a: \mu \neq \mu_0$ *1.* $H_0: \mu \geq \mu_0$ & $H_a: \mu < \mu_0$ *2.* $H_0: \mu \leq \mu_0$ & $H_a: \mu > \mu_0$
Step Two:	1. Quantitative *2. Random Sample* *3.* $n > 30$ OR the population is bell shaped
Step Three:	$z^* = \frac{(\bar{x} - \mu_o)}{\frac{\sigma_x}{\sqrt{n}}}$
Step Four:	$H_a: \mu \neq \mu_0$ → p-value = 2*P(Z<-|z*|) $H_a: \mu < \mu_0$ → p-value = P(Z<z*) $H_a: \mu > \mu_0$ → p-value = P(Z>z*) = 1-P(Z<z*)
Step Five:	If p-value $\leq (1 - confidene) = \alpha$ → Reject H_0 If p-value $> (1 - confidence) = \alpha$ → Fail to Reject H_0

Hypothesis Testing unknown σ_x

Step One:	1. $H_0: \mu = \mu_0$ & $H_a: \mu \neq \mu_0$ 1. $H_0: \mu \geq \mu_0$ & $H_a: \mu < \mu_0$ 2. $H_0: \mu \leq \mu_0$ & $H_a: \mu > \mu_0$
Step Two:	1. Quantitative 2. *Random Sample* 3. $n > 30$ OR the population is bell shaped
Step Three:	$t^* = \frac{(\bar{x} - \mu_o)}{\frac{s_x}{\sqrt{n}}}$
Step Four:	$H_a: \mu \neq \mu_0$→ p-value = 2*P(T<-|t*|) $H_a: \mu < \mu_0$→ p-value = P(T<t*) $H_a: \mu > \mu_0$→ p-value = P(T>t*) = 1-P(T<t*)
Step Five:	If p-value$\leq (1 - confidene) = \alpha$ → Reject H_0 If p-value$> (1 - confidence) = \alpha$ → Fail to Reject H_0