Chapter 10—Hypothesis Testing

> *Definition*: A **statistical hypothesis** is a statement or claim regarding a population parameter (e.g., μ=500).

To determine the truth or falsity of a statistical hypothesis with 100% accuracy, you would need to examine the entire population. But, that is not possible because it would take too much time and cost too much to look at all the observations in a population.

In practice, we take a sample and use the information in the sample to decide whether we believe the hypothesis.

> *Definition*: **Statistical test** is a statistical procedure or decision rule that leads to establishing the truth or falsity of a statistical hypothesis.

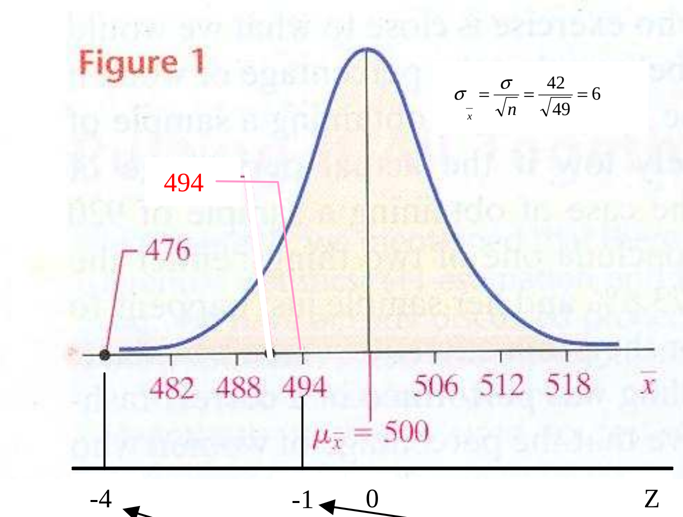
> *Definitions*: The **null hypothesis** denoted H_0 (read "H-naught"), is a statement to be tested (e.g., H_0: μ=500 hours). The null hypothesis is assumed true until evidence indicates otherwise.
> The **alternative hypothesis,** denoted H_1 (read "H-one"), is what is considered to be true if the null hypothesis is rejected (H_1: $\mu \neq$ 500 hours). The alternative hypothesis is the claim that we seek evidence for.

Steps in Hypothesis Testing (p. 454)	
1) A claim is made about a population parameter. >>>>>>>>>>>>>	*Statistical hypothesis* is stated.
2) Evidence (sample data) is collected to test the claim. >>>>>>>>>	*Random sample* is drawn.
3) The data are analyzed in order to support or refute the claim. >>	*Statistical test* is applied.

The Logic of Hypothesis Testing (pp. 462-65)

Problem—The packaging on a lightbulb states that the bulb will last 500 hours under normal use. A consumer advocate would like to know if the mean lifetime of a bulb is less than 500 hours (a claim regarding the population mean). A random sample of n=49 lightbulbs is burned to determine how long a lightbulb lasts. Assume we know the population standard deviation is σ = 42.

H_0: μ = 500 hours versus H_1: μ < 500 hours



$$Z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{476 - 500}{\frac{42}{\sqrt{49}}} = \frac{-24}{6} = -4.00 \qquad Z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{494 - 500}{\frac{42}{\sqrt{49}}} = \frac{-6}{6} = -1.00$$

If $\bar{x} = 494$, then the sample mean is one standard deviation below 500 (the claim regarding the population mean).

- $P(\bar{X} \le 494) = 0.1587$, which would happen 16% of the time under H_0.
- In this case, we <u>do not reject</u> H_0.
- Note: We would only reject H_0 in the event of obtaining an "unusual" sample (i.e., a sample that occurs with low probability under the null hypothesis).

If $\bar{x} = 476$, then the sample mean is four standard deviations below 500.

- $P(\bar{X} \le 476) = 0.0$, which says there is essentially no chance of finding a sample mean of 476 when H_0 is true.
- In this case, we would <u>reject</u> the null hypothesis.
- In this case, we are inclined to believe that the sample mean came from a population whose mean is less than 500 (draw this distribution on the graph above).

Logic of hypothesis testing (pp. 462-65): We reject the null hypothesis if the sample mean is "too many" standard deviations away from the null hypothesis (H_0) (i.e., an unusual event that happens <5% of the time).

Or, stated another way, if the sample data result in a statistic ($\bar{x}$) that is not likely under the assumption that the null hypothesis is true, we reject the null hypothesis.

Important points to remember about hypothesis tests (refer back to these points after you have worked some hypothesis-test problems):

1. A statistical hypothesis ALWAYS includes a **parameter** (e.g., H_0: $\mu=500$), and never a statistic ($\bar{x}$).

2. The null hypothesis (or simply hypothesis) ALWAYS includes an **equals sign** (=).

3. In a hypothesis test, we ALWAYS begin the test assuming the null hypothesis is true.

4. The claim we seek evidence for is the alternative hypothesis.

5. The alternative hypothesis includes <, >, or ≠ . Look for key phrases in the claim. For example, "more than" means >; "different from" means ≠ ; and "less than" means <.

6. The alternative hypothesis can be one-tailed or two-tailed.

   One-tailed → Left-tailed
   (H_1: $\mu<500$)

   → Right-tailed
   (H_1: $\mu>500$)

   Two-tailed (H_1: $\mu \neq 500$)

7. If you are unsure whether to use a one-tailed or two-tailed hypothesis test, ALWAYS use a two-tailed test.

8. α = significance level = probability of a Type I error (reject a true H_0)
   β = probability of a Type II error (do not reject a false H_0)

9. When testing a statistical hypothesis, there is always a possibility that your conclusion will be wrong (Type I error or Type II error). And, to make matters worse, you won't know whether you are wrong or not.

10. You can **never say the null hypothesis is TRUE** unless you have access to all the population data (and that never occurs). Rather, we say we **do not reject** the null hypothesis.

11. A **significant result** occurs when you reject the null hypothesis.

12. P-value is the probability that the test statistic takes a value equal to or more extreme than the value actually observed (in both directions for a two-tail test), assuming H_0 is true.

13. A "large" P-value is evidence for H_0 and a "small" P-value is evidence against H_0.

14. An **extremely small P-value (<0.01)** means that H_0 is **strongly rejected** or the result is **highly statistically significant**.

Three approaches are available for testing a statistical hypothesis:

1) Classical
2) P-Value
3) Confidence-Interval

The approaches will be illustrated below. For each approach you will find a “template” that shows the 4 steps of a hypothesis test and an example illustrating the application of the hypothesis test. In the case of the P-value approach, an explanation is provided of how to use Excel to find P-values.

Classical Method using the Z-Distribution—Hypothesis Test Regarding μ with σ Known

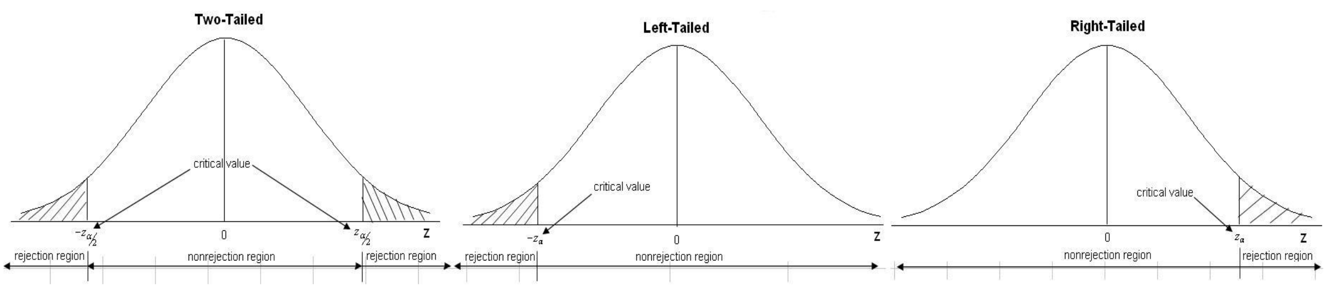
Assumptions:

- The sample is obtained using simple random sampling
- The population, from which the sample is drawn, is normally distributed or the sample size, n, is "large" ($n \geq 30$).

Step 1: A claim is made regarding the population mean, μ. The null and alternative hypotheses can be structured in three ways:

Two-Tailed	Left-Tailed	Right-Tailed
H_0: $\mu = \mu_0$	H_0: $\mu = \mu_0$	H_0: $\mu = \mu_0$
H_1: $\mu \neq \mu_0$	H_1: $\mu < \mu_0$	H_1: $\mu > \mu_0$
$\mu<\mu_0$ or $\mu>\mu_0$		

Step 2: Select a **level of significance**, **α**, which is generally chosen to be 0.10, 0.05, or 0.01. The significance level is used to determine the *critical value*. **Critical value** is the Z-value that separates the rejection and nonrejection regions. The **rejection region** (or **critical region**) is the set of all values of the test statistic (defined in step 3) such that the null hypothesis is rejected.



Step 3: Calculate the **test statistic or calculated Z-value**.

$$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

The test statistic (Z) measures the number of standard deviations that the actual sample mean, $\bar{x}$, is from the assumed population mean, μ_0.

Step 4: Draw a conclusion:

- Compare the calculated Z-value (or test statistic) to the critical Z-value and state whether or not H_0 is rejected at the specified α.

Two-Tailed	Left-Tailed	Right-Tailed
If $Z < -z_{\alpha/2}$ *or* $Z > z_{\alpha/2}$ *reject the null hypothesis*	*If* $Z < -z_\alpha$ *reject the null hypothesis.*	*If* $Z > z_\alpha$ *reject the null hypothesis.*

- Interpret the conclusion in the context of the problem.

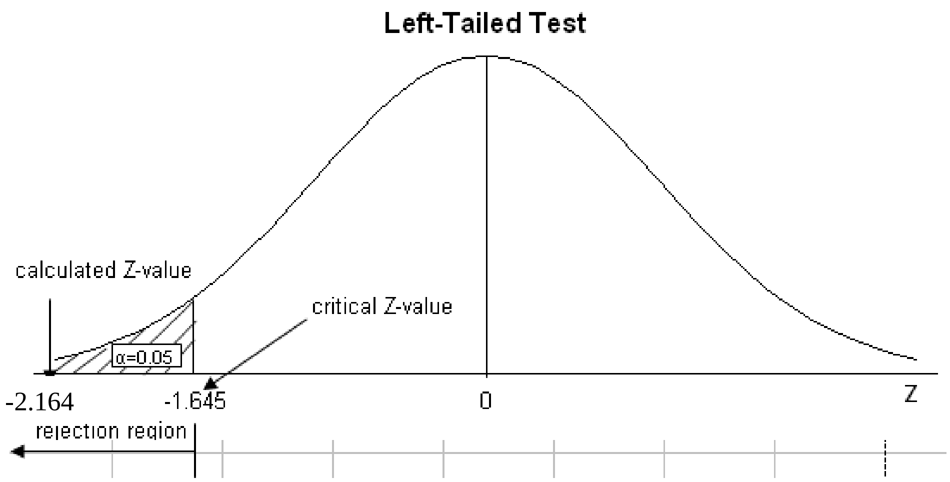
Classical Method using the Z-Distribution—Hypothesis Test Regarding μ with σ Known

Problem: A feed dealer buys 20% protein feed from a feed manufacturer and **resells** the feed to local ranchers. The feed dealer is interested in checking to make certain that the feed does not average less than 20% protein. Carry out a hypothesis test of the relevant null hypothesis at the 5% significance level. Show all of your calculations and justify your conclusion.

Step 1: Null and Alternative Hypotheses:

H_0: $\mu = 20\%$ protein
H_1: $\mu < 20\%$ protein

Step 2: Select $\alpha = 0.05$ and find the critical value of Z.



Step 3: Draw a random sample of n=10 bags of feed and calculate the test statistic, Z. Assume we know the population standard deviation is σ=0.19.

Protein content of 10 bags of feed (%) X
19.60
19.95
20.15
19.90
20.00
19.82
19.85
20.04
19.79
19.60

$\bar{x} = 19.87$

$$Z = \frac{(\bar{x} - \mu_0)}{\frac{\sigma}{\sqrt{n}}} = \frac{(19.87 - 20.)}{\frac{0.19}{\sqrt{10}}} = \frac{-0.13}{0.060} = -2.164$$

Step 4: Conclusion—Because the calculated Z = -2.164 is less than the critical z = -1.645 (and in the rejection region), reject H_0 at the 0.05 significance level. The mean protein level of the feed is significantly less than 20%.

P-Value Method using the Z-Distribution—Hypothesis Test Regarding μ with σ Known

Assumptions:
- The sample is obtained using simple random sampling
- The population, from which the sample is drawn, is normally distributed or the sample size, n, is "large" ($n \geq 30$).

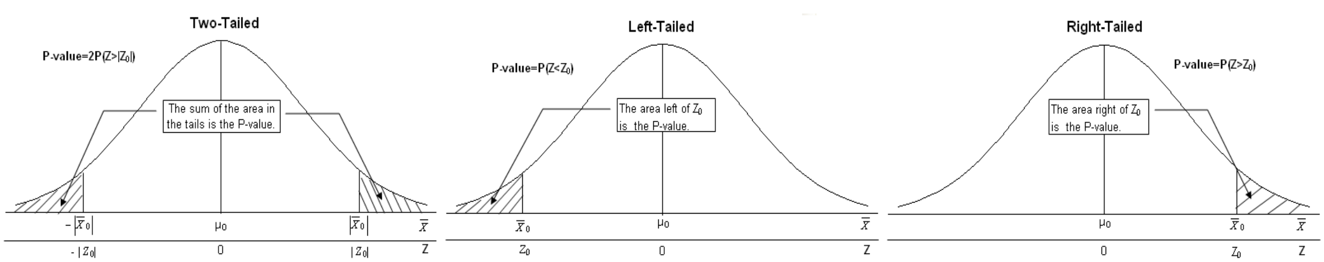
Step 1: A claim is made regarding the population mean, μ. The null and alternative hypotheses can be structured in three ways:

Two-Tailed	Left-Tailed	Right-Tailed
H_0: $\mu = \mu_0$	H_0: $\mu = \mu_0$	H_0: $\mu = \mu_0$
H_1: $\mu \neq \mu_0$	H_1: $\mu < \mu_0$	H_1: $\mu > \mu_0$
$\mu<\mu_0$ or $\mu>\mu_0$		

Step 2: Select a **level of significance**, **α**.

Step 3: Calculate the **test statistic**, $Z_0 = \dfrac{\bar{x}_0 - \mu_0}{\sigma/\sqrt{n}}$, and determine the **P-value** using Table III or Excel (see backside of this page).

P-value is the probability of observing a test statistic as extreme or more extreme than the one observed under the assumption that the null hypothesis is TRUE (for a two-tailed test, *extreme* includes both directions).



Step 4: Draw a conclusion:

- Compare the calculated P-value to the significance level and state whether or not H_0 is rejected at the specified α.

If P-value > α, do not reject H_0
If P-value ≤ α, reject H_0

***Note that this decision rule applies to one-tailed and two-tailed tests.

- Interpret the conclusion in the context of the problem.

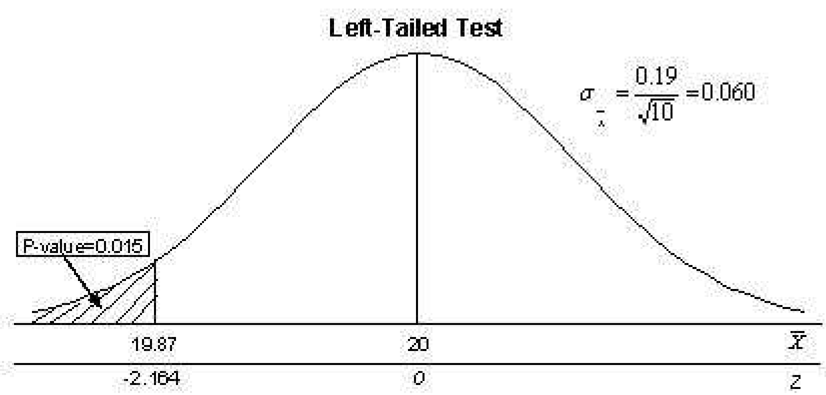
P-Value Method using the Z-Distribution—Hypothesis Test Regarding μ with σ Known

Problem: A feed dealer buys 20% protein feed from a feed manufacturer and **resells** the feed to local ranchers. The feed dealer is interested in checking to make certain that the feed does not average less than 20% protein. Carry out a hypothesis test of the relevant null hypothesis at the 5% significance level. Show all of your calculations and justify your conclusion.

Step 1: Null and Alternative Hypotheses:

H_0: μ = 20% protein
H_1: μ < 20% protein

Step 2: Select $\alpha = 0.05$.



Step 3: Draw a random sample of n=10 bags of feed. Calculate the test statistic, Z_0, and the P-value. Assume we know the population standard deviation is σ=0.19.

Protein content of 10 bags of feed (%) X
19.60
19.95
20.15
19.90
20.00
19.82
19.85
20.04
19.79
19.60

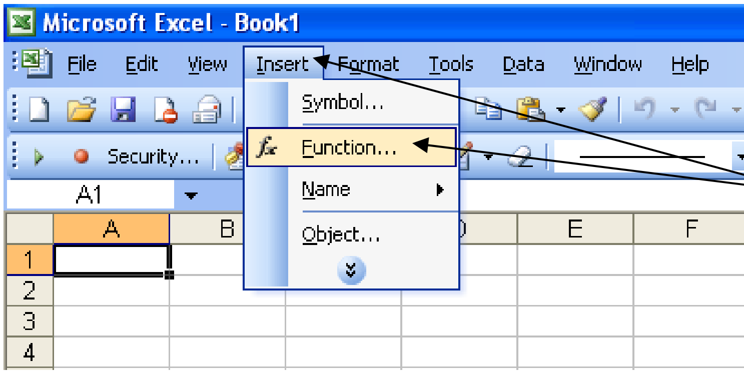
$$\bar{x} = 19.87$$

$$Z_0 = \frac{(\bar{x} - \mu_0)}{\frac{\sigma}{\sqrt{n}}} = \frac{(19.87 - 20.)}{\frac{0.19}{\sqrt{10}}} = \frac{-0.13}{0.060} = -2.164$$

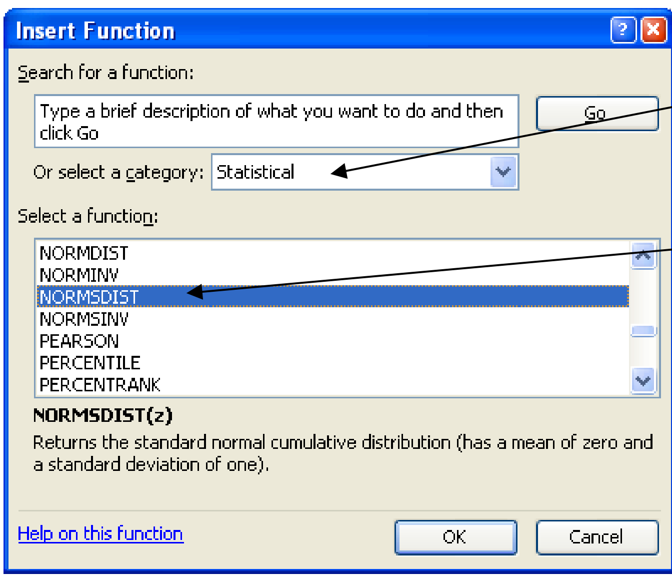
Step 4: Conclusion—Because the P-value = 0.015 is less than α=0.05, reject H_0 at the 0.05 significance level. The mean protein level of the feed is significantly less than 20%.

Excel: Finding P-values for a Z-distribution.

Step 1: Select **Insert/Function** ($\boldsymbol{f_x}$) from the Windows menubar. In the **Function Category**, select "Statistical." In the **Function Name**, select "NORMSDIST."
Step2: Enter the test statistic **Z**=$|Z_0|$ and click **OK**. To obtain the P-value for a one-tailed test, subtract the "Formula result" (at the bottom of the **Function Arguments** window shown below) from 1.0 to find P(Z>$|Z_0|$). The P-value for a two-tailed test is calculated as 2P(Z>$|Z_0|$).

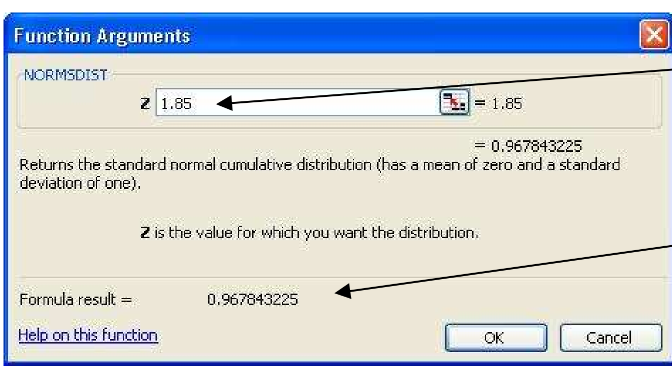


From the **Insert** menu select **Function**



Select a Category: Statistical.

Select a Function: NORMDIST



Enter the Z value

Get the result

Confidence-Interval Method using the Z-Distribution—Hypothesis Test Regarding μ with σ Known

Assumptions:

- The sample is obtained using simple random sampling
- The population, from which the sample is drawn, is normally distributed or the sample size, n, is "large" ($n \geq 30$).

Step 1: A claim is made regarding the population mean, μ. The null and alternative hypotheses can be structured in three ways:

Two-Tailed	Left-Tailed	Right-Tailed
H_0: $\mu = \mu_0$	H_0: $\mu = \mu_0$	H_0: $\mu = \mu_0$
H_1: $\mu \neq \mu_0$	H_1: $\mu < \mu_0$	H_1: $\mu > \mu_0$
$\mu<\mu_0$ or $\mu>\mu_0$		

Step 2: Select a **level of significance**, $\boldsymbol{\alpha}$, and calculate a **$(1-\alpha)\cdot 100\%$ confidence interval**, **σ known**.

$(1-\alpha)\cdot 100\%$ Confidence Interval: $\overline{X} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

Step 3: Draw a conclusion:

- Compare μ_0 to the confidence interval bounds and state whether or not H_0 is rejected at the specified α.

Two-Tailed
If the confidence interval contains μ_0, do not reject the null hypothesis.

- Interpret the conclusion in the context of the problem.

Confidence-Interval Method using the Z-Distribution—Hypothesis Test Regarding μ with σ Known

Problem: A feed manufacturer produces and sells 20% protein to local ranchers. The feed manufacturer is interested in checking to determine if the feed includes a mean of 20% protein. Carry out a hypothesis test of the relevant null hypothesis at the 5% significance level. Show all of your calculations and justify your conclusion.

Step 1: Null and Alternative Hypotheses:

H_0: $\mu = 20\%$ protein
H_1: $\mu \neq 20\%$ protein, i.e., $\mu < 20\%$ protein or $\mu > 20\%$ protein

Step 2: Select $\alpha=0.05$. Draw a random sample of n=10 bags of feed and calculate a 95% confidence interval. Assume we know the population standard deviation is $\sigma=0.19$.

Protein content of 10 bags of feed (%) X
19.60
19.95
20.15
19.90
20.00
19.82
19.85
20.04
19.79
19.60

$\bar{x} = 19.87$

95% Confidence Interval:

$$\bar{X} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$19.87 \pm 1.96 \cdot \frac{0.19}{\sqrt{10}}$$

$$19.87 \pm 0.118$$

$$19.752 \textit{ to } 19.988$$

Step 3: Conclusion—Because the 95% confidence interval does not include the hypothesized value of the mean, 20%, reject H_0 at the 0.05 significance level. The mean protein level of the feed is significantly different (less) than 20%.

Classical Method using the t-Distribution—Hypothesis Test Regarding μ with σ Unknown

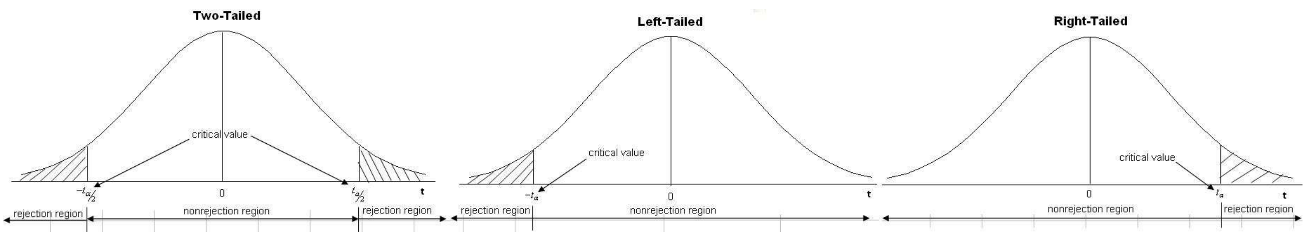
Assumptions:

- The sample is obtained using simple random sampling
- The population, from which the sample is drawn, is normally distributed or the sample size, n, is "large" ($n \geq 30$).

Step 1: A claim is made regarding the population mean, μ. The null and alternative hypotheses can be structured in three ways:

Two-Tailed	Left-Tailed	Right-Tailed
H_0: $\mu = \mu_0$	H_0: $\mu = \mu_0$	H_0: $\mu = \mu_0$
H_1: $\mu \neq \mu_0$	H_1: $\mu < \mu_0$	H_1: $\mu > \mu_0$
$\mu<\mu_0$ or $\mu>\mu_0$		

Step 2: Select a **level of significance**, **α**, which is generally chosen to be 0.10, 0.05, or 0.01. The significance level is used to determine the *critical value*. **Critical value** is the t-value that separates the rejection and nonrejection regions. The **rejection region** (or **critical region**) is the set of all values of the test statistic (defined in step 3) such that the null hypothesis is rejected.



Step 3: Calculate the **test statistic** or **calculated t-value**.

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

which follows Student's t-distribution with df=n-1. The test statistic (t) measures the number of standard deviations that the sample mean, $\bar{x}$, is from the assumed population mean, μ_0.

Step 4: Draw a conclusion:

- Compare the calculated t-value (or test statistic) to the critical t-value and state whether or not H_0 is rejected at the specified α.

Two-Tailed	Left-Tailed	Right-Tailed
If $t < -t_{\alpha/2}$ *or* $t > t_{\alpha/2}$ *reject the null hypothesis*	*If* $t < -t_\alpha$ *reject the null hypothesis.*	*If* $t > t_\alpha$ *reject the null hypothesis.*

- Interpret the conclusion in the context of the problem.

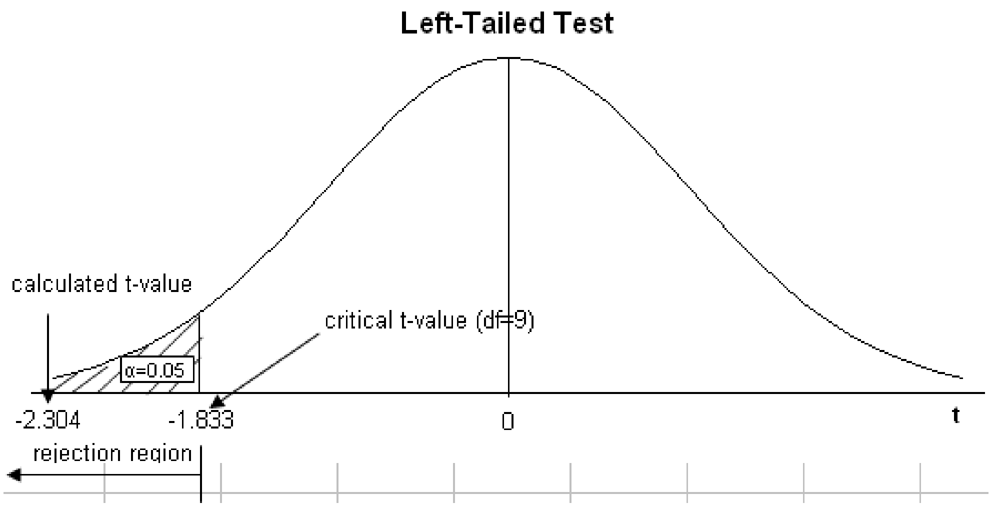
Classical Method using the t-Distribution—Hypothesis Test Regarding μ with σ Unknown

Problem: A feed dealer buys 20% protein feed from a feed manufacturer and **resells** the feed to local ranchers. The feed dealer is interested in checking to make certain that the feed does not average less than 20% protein. Carry out a hypothesis test of the relevant null hypothesis at the 5% significance level. Show all of your calculations and justify your conclusion.

Step 1: Null and Alternative Hypotheses:

H_0: μ = 20% protein
H_1: μ < 20% protein

Step 2: Select $\alpha = 0.05$ and find the critical value of t (df=9).



Step 3: Draw a random sample of n=10 bags of feed. Calculate the sample standard deviation, s, and the test statistic, *t*.

Protein content of 10 bags of feed (%) X
19.60
19.95
20.15
19.90
20.00
19.82
19.85
20.04
19.79
19.60

$\bar{x} = 19.87$

$$t = \frac{(\bar{x} - \mu_0)}{\frac{s}{\sqrt{n}}} = \frac{(19.87 - 20.)}{\frac{0.178}{\sqrt{10}}} = \frac{-0.13}{0.056} = -2.304$$

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{0.287}{10-1}} = 0.178$$

Step 4: Conclusion—Because the calculated t = -2.304 is less than the critical t = -1.833 (and in the rejection region), reject H_0 at the 0.05 significance level. The mean protein level of the feed is significantly less than 20%.

P-Value Method using the t-Distribution—Hypothesis Test Regarding μ with σ Unknown

Assumptions:
- The sample is obtained using simple random sampling
- The population, from which the sample is drawn, is normally distributed or the sample size, n, is "large" ($n \geq 30$).

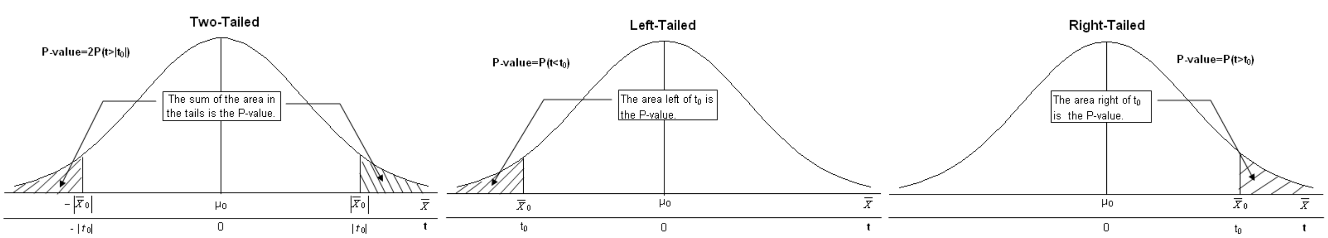
Step 1: A claim is made regarding the population mean, μ. The null and alternative hypotheses can be structured in three ways:

Two-Tailed	Left-Tailed	Right-Tailed
H_0: $\mu = \mu_0$	H_0: $\mu = \mu_0$	H_0: $\mu = \mu_0$
H_1: $\mu \neq \mu_0$	H_1: $\mu < \mu_0$	H_1: $\mu > \mu_0$
$\mu<\mu_0$ or $\mu>\mu_0$		

Step 2: Select a **level of significance**, **α**.

Step 3: Calculate the **test statistic**, $t_0 = \dfrac{\bar{x}_0 - \mu_0}{s/\sqrt{n}}$, and determine the **P-value** using Table III (explained on p. 412) or Excel (see backside of this page).

P-value is the probability of observing a test statistic as extreme or more extreme than the one observed under the assumption that the null hypothesis is TRUE (for a two-tailed test, extreme includes both directions).



Step 4: Draw a conclusion:

- Compare the calculated P-value to the significance level and state whether or not H_0 is rejected at the specified α.

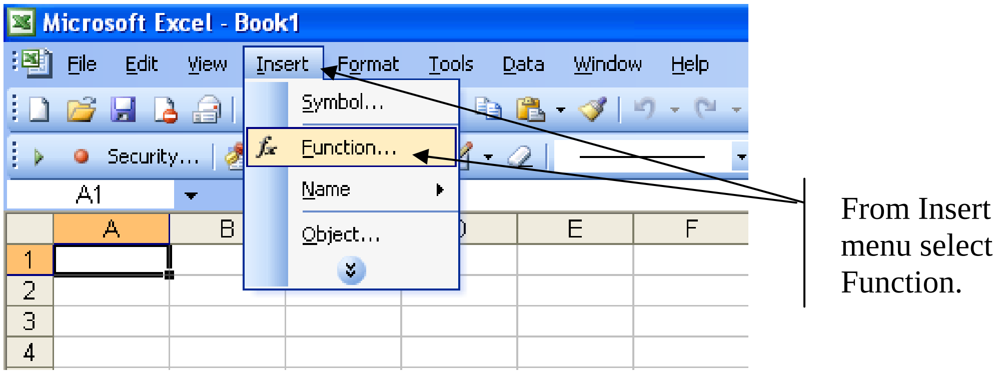
  If P-value > α, do not reject H_0
  If P-value ≤ α, reject H_0

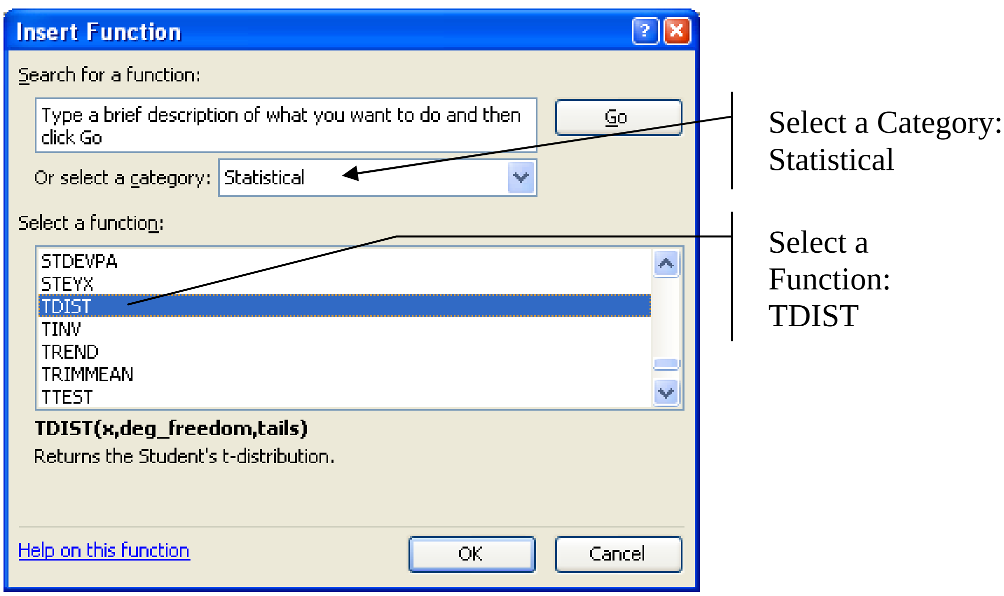
  ***Note that this decision rule applies to one-tailed and two-tailed tests.

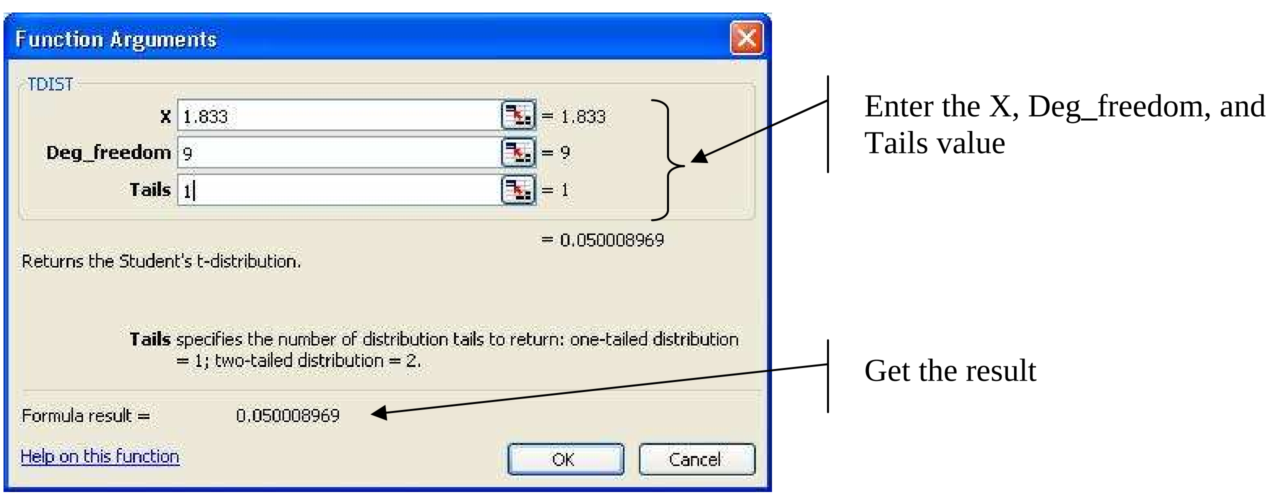
- Interpret the conclusion in the context of the problem.

Excel: Finding P-values for a t-distribution:

Step 1: Select **Insert/Function** (f_x) from the Windows menubar. In the **Function Category**, select "Statistical." In the **Function Name**, select "TDIST."
Step2: Enter the test statistic **X** = $|t_0|$, **Deg_freedom** = n-1, and **Tails** equals 1 or 2, depending on whether a one- or two-tailed test is used. Click **OK**. Read the P-value from the "Formula result" at the bottom of the **Function Arguments** window.







P-Value Method using the t-Distribution—Hypothesis Test Regarding μ with σ Unknown

Problem: A feed dealer buys 20% protein feed from a feed manufacturer and **resells** the feed to local ranchers. The feed dealer is interested in checking to make certain that the feed does not average less than 20% protein. Carry out a hypothesis test of the relevant null hypothesis at the 5% significance level. Show all of your calculations and justify your conclusion.

Step 1: Null and Alternative Hypotheses:

H_0: $\mu = 20\%$ protein
H_1: $\mu < 20\%$ protein

Step 2: Select $\alpha = 0.05$.

Step 3: Draw a random sample of n=10 bags of feed. Calculate the sample standard deviation, s. Calculate the test statistic, t_0, and the P-value.

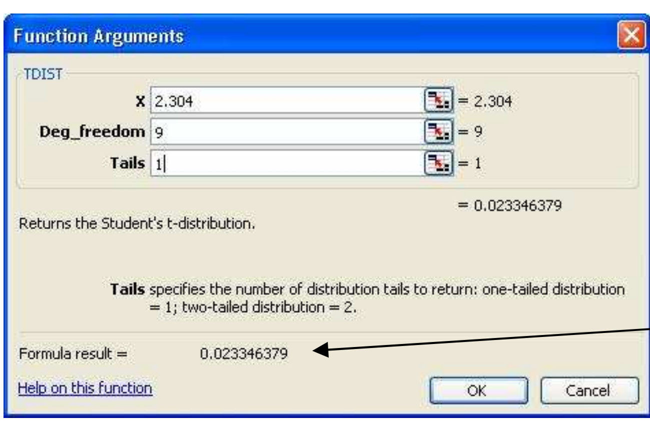
Protein content of 10 bags of feed (%) X
19.60
19.95
20.15
19.90
20.00
19.82
19.85
20.04
19.79
19.60

$\bar{x} = 19.87$

$$t_0 = \frac{(\bar{x}-\mu_0)}{\dfrac{s}{\sqrt{n}}} = \frac{(19.87-20.)}{\dfrac{0.178}{\sqrt{10}}} = \frac{-0.13}{0.056} = -2.304$$

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{0.287}{10-1}} = 0.178$$

Using Table III, the P-value is as follows: 0.02 < P-value < 0.025. When σ is unknown, exact P-values can be found only via technology (e.g., Excel P-value of 0.023 is shown below).



Select **Insert/Function ($\mathbf{f_x}$)** from the Windows menu. In the **Function Category**, select "Statistical." In the **Function Name**, select "TDIST."

Enter the following information:
X = 2.304 (absolute value of the t-statistic from **Step 3**)
df = 9
Tails = 1 (for area in one tail)

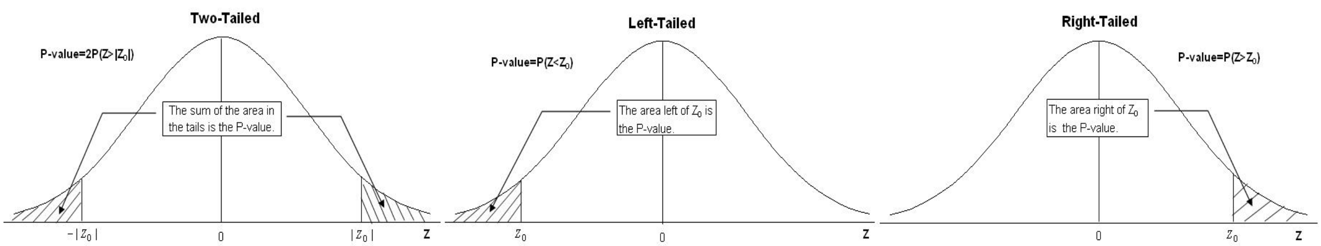
Step 4: Conclusion—Because the P-value = 0.023 is less than α=0.05, reject H_0 at the 0.05 significance level. The mean protein level of the feed is significantly less than 20%.

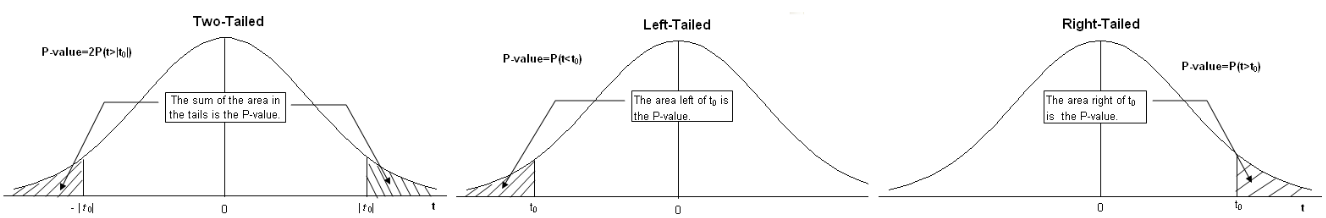
Junk below here:

Step 5: State the conclusion, with specific reference to the question under consideration.

The critical value represents the maximum number of standard deviations the sample mean can be from μ_0 before the null hypothesis is rejected.

Two-Tailed	Left-Tailed	Right-Tailed
If $Z < -z_{\alpha/2}$ or $Z > z_{\alpha/2}$ reject the null hypothesis	If $Z < -z_{\alpha}$ reject the null hypothesis.	If $Z > z_{\alpha}$ reject the null hypothesis.





The feed dealer is interested in checking to determine if the feed includes an average of 20% (or higher) protein.