Week Ten: Hypothesis Testing with Proportions

Week Ten Goals

- Vocabulary of Hypothesis Testing
- The Complete Testing Process ← you will be asked to show this process for each test.
- Population Step
  - Describe the Population and State the Goal of the Test
- Method Step
  - State the Null and Alternative Hypotheses (Left, Right, Two-Tailed)
  - State the Level of Significance
  - Indicate the Distribution that will be Used
- Sample Step
  - One Proportion Test—with StatHelper
  - One Proportion Test—with Minitab 19
  - Check Normality Assumption $n(p_0)(1 - p_0) \geq 10$
- Results Step
  - The Z - Test Statistic and Interpretation
  - The P-value and Interpretation
- Conclusion Step
  - Make a Decision
  - Write a Statistical Conclusion
  - State an Everyday-English Conclusion
- Examples
- Type I Error and Type II Error Examples

Vocabulary of Hypothesis Testing (Same as Week 9)

Parameter: A quantity that is calculated from data and describes a **population.**

Statistic: A quantity that is calculated from data and describes a **sample.**

Null Hypothesis: A general statement or default position that there is nothing new happening, like there is no association among groups, or no relationship between two measured phenomena or that nothing has changed…it is as we always thought it was. The status quo. The null hypothesis always includes an equal sign.

Alternative Hypothesis: A statement that indicates the change that the researcher is hoping to show. It is often the goal of the researcher. The alternative hypothesis is always written as either 'less than', 'greater than' or 'not equal'.

Hypothesized Value: The value that you are assuming is true in the null hypothesis.

Level of Significance: Also denoted as alpha or α, is the probability of rejecting the null hypothesis when it is true. For example, a **significance level** of 0.05 indicates a 5% risk of concluding that a difference exists when there is no actual difference.

Test Statistic: A numerical summary of a data-set that reduces the data to one value that can be used to perform the hypothesis test. The test statistic indicates the number of standard errors away that the sample statistic is from the hypothesized value.

Probability-value (P-value): The probability of a getting the statistic you get from sampling or more extreme under the assumption that the null hypothesis is true.

Reject the Null Hypothesis: When the sample data is sufficiently different than the hypothesized value, the calculated P-value will be smaller than or equal to alpha. As a result, we will decide that we CAN "reject the null hypothesis". In other words, we are no longer believing the statement in the null hypothesis; we are deciding that we have enough evidence to support the alternative hypothesis.

Fail to reject the null hypothesis or deciding we CANNOT reject the null hypothesis. In the event that our sample data is not sufficiently different from the hypothesized value, the calculated P-value will be greater than alpha. As a result, we will decide that we "CANNOT reject the null hypothesis". In other words, we do not have enough evidence to support the alternative hypothesis. We are not going to come right out and say we '*know that the null hypothesis is true*' but we certainly are not able to support a statement against the null.

Type I Error: Error that occurs when the sample data leads you to decide to reject the null, ***when the null is actually true***. The probability of Type I Error is alpha…the level of significance. The default alpha = 0.05.

Type II Error: Error that occurs when the sample data leads you to decide that we cannot reject the null, ***when the null should have been rejected***. The probability of Type II Error is beta. If beta decreases, alpha increases. If alpha decreases, beta increases; there is an inverse relationship between alpha and beta. We will not calculate beta in this course.

The Complete Testing Process: There are steps that Prof. Coffey wants to see for each hypothesis test. This process includes the following steps: Population Step, Method Step, Sample Step, Results Step and Conclusion Step. These are defined on the following pages and the process shown in all examples. When I ask for the complete testing process, I am asking for you to show these steps as taught. If you do not, then it is very clear to me that you have not read these typed notes.

Video Introducing you to the Vocabulary of hypothesis testing from Prof. Coffey

These videos are the same videos as Week Nine

Here is a video about vocabulary for hypothesis testing: https://youtu.be/n5Z1WBEKYdo

Here is the same video with interpreting: https://youtu.be/yBCkkdddmVA

The Complete Testing Process

Population Step -- Describe the population and state the goal.

1. Describe the variable being studied (also indicate if the variable is numerical or categorical; this week it is CATEGORICAL).
2. Define the parameter **p** = The true **proportion** of___________
3. State the GOAL of the analysis. ***The Goal ALWAYS matches the Ha.***

We will be testing to see if the population **proportion**, p, has increased, decreased or is different from the hypothesized **proportion**, p_0.

Method Step -- Indicate the method that will be used.

1. State the Ho and Ha.

$Ho: p = p_0$ $Ha: p < p_0$	Left-Tailed Test
$Ho: p = p_0$ $Ha: p > p_0$	Right-Tailed Test
$Ho: p = p_0$ $Ha: p \neq p_0$	Two-Tailed Test

2. State alpha (α). [α = 0.05 is the default; use it unless told otherwise]

NOTE: α is a level of significance. It is the probability of Type I error, i.e. the probability of rejecting the null (and concluding that the alternative hypothesis was true) when in fact, the null is true.

3. State the distribution with which you will work. When we study **categorical** data, we will work on the **z-curve**, the standard normal distribution.

Sample Step -- Analyze the sample data by running the test with technology; also check the normality assumption.

You need the number of events, x and the number of trials, n. Recall that this is used to calculate $\hat{p} = \frac{x}{n}$.

1. Run the hypothesis test using statistical technology and paste the output here.

2. Check to see if the normality assumption has been met. In order to determine if the sample size, n, is large enough, the following calculation will be used. NOTE: The hypothesized proportion, p_0 is used in the calculation (since it is sitting at the center of your Z-curve). We can assume a normal distribution if $\boldsymbol{n\,(p_0)(1 - p_0) \geq 10}$.

Results Step – Z Test statistic and P-value results <u>and</u> Interpretations

1. State the Z-test statistic and <u>interpret</u> the Z-test statistic.
2. State and <u>interpret</u> the P-value.

Conclusion Step -- Make a decision and a statistical conclusion; everyday conclusion encouraged

Determine if your P-value is less than or equal to alpha? (P-value $\leq \alpha$?) **Yes** **No**
1. State your decision about the null hypothesis (Ho). **Can Reject Ho** **Cannot Reject Ho**

A statistical conclusion is always a statement indicating that we **HAVE ENOUGH** support or that we **DO NOT HAVE ENOUGH** support for the Ha/goal. You can use a template such as the one below

2. Write the statistical conclusion:
"At the ___% level of significance, the sample data **DOES** or **DOES NOT** provide sufficient evidence to state that the true **proportion** of …………….restate the goal here………………."

A conclusion in everyday English:
Often, I will ask you to also write an 'everyday' conclusion; a conclusion that is written casually, in the context of the problem, with no statistical language and gets right to the point of the conclusion.

Population Step

Describe the Population and State the Goal of the Test:
It is important that we can…

- State the variable (and units) being studied and indicate if it is numerical or categorical.
- If the variable is numerical, we will define the parameter μ = the true/population mean..........describe variable...........
- If the variable is **categorical**, we will define the parameter **p** = the true/population **proportion** of ....describe variable...........
- State the GOAL of the hypothesis test: "Test to see if there is support for saying that μ / p has (increased, decreased, or changed) from ……describe μ_0 *or* p_o (*the hypothesized value*)…."

Method Step

State the Null and Alternative Hypotheses (Left, Right, Two-Tailed)

In Statistics, when testing **claims**, we use an objective method called hypothesis testing.
We call these **claims** hypotheses. Our starting point, the status quo, is called the **null hypothesis (H_o)** and the alternative claim is called the **alternative hypothesis (H_a).**
We may have some past information on a variable and wonder if the current data show that the value has statistically increased, decreased or is 'different' than the past information. Hypothesis testing can accomplish this.

The Null Hypothesis (H_o)

The Null Hypothesis, denoted by **H_o**, specifies a population parameter of interest (μ or p) and proposes a value for that parameter, called the hypothesized value (μ_0 or p_0). This statement is what we are assuming is true as we begin testing. It is generally a statement of 'status quo', meaning it is what we have always believed the value to be. The null hypothesis always is a statement of 'equality'; meaning there is always an EQUAL SIGN in the null hypothesis. It is possible for the Ho to read p ≤ ____ or p ≥ ___ (as long as the equal sign is there, it is a fair statement for the null). However, it is not vital that the null be anything other than an equal sign, so we can agree to always make the null statement a statement of equality.

The Alternative hypothesis (H_a)

The Alternative hypothesis, denoted by **H_a or H_1**, is the claim we are testing for. It is what the researcher is interested in proving. The alternative hypothesis is a statement that μ/p has decreased or μ/p has increased or that μ/p has changed.

$H_0: p = \_\_\_\_\_$ $H_a: p < \_\_\_\_\_$ OR	$H_0: p = \_\_\_\_\_$ $H_a: p > \_\_\_\_\_$ OR	$H_0: p = \_\_\_\_\_$ $H_a: p \neq \_\_\_\_\_$
If you are trying to prove that μ has decreased. **This is a left-tailed test.**	If you are trying to prove that μ has increased. **This is a right-tailed test.**	If you are trying to prove that μ is different (has changed). **This is called a two-tailed test.**

NOTE: When typing, you do not have to use subscripts, like I have done. I am fine with Ho and Ha.

State the Level of Significance

Every hypothesis test has a level of significance associated with it. Early on in the research process, before beginning to collect data, the researcher decides on this level of significance. It would be improper to change this level of significance after seeing your results! The Greek letter alpha α is used to represent level of significance. Most hypothesis tests use an alpha value of .05. When testing on life and death decisions is made, often an alpha value of .01 is chosen. Occasionally, we will see alpha = .10

Unless distinctly told otherwise, we will always choose an alpha = .05.

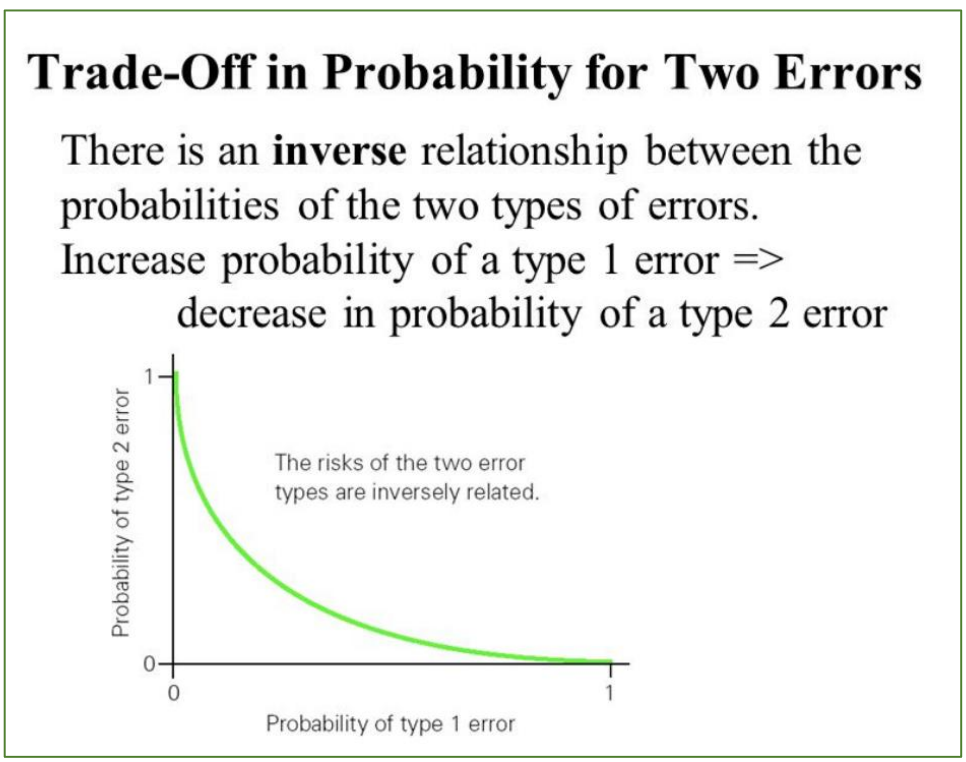
Alpha = α = level of significance/significance level = the probability of type 1 error

The level of significance, alpha, is the probability of Type 1 error. In other words, the level of significance, alpha, is the probability of rejecting the null hypothesis when it is true.

For example, a significance level of 0.05 indicates a 5% risk of concluding that a difference exists and we have enough evidence to support the alternative hypothesis (Ha), when there really is no difference.

Type I Error is the probability of deciding we **CAN** reject the H_o when, in fact, that decision to reject is actually a mistake. Alpha, α, is the probability of Type I Error.

Type II Error is the probability of deciding we **CANNOT** reject the H_o when, in fact, that decision is actually a mistake. Beta, β, is the probability of Type II Error. We will not learn to calculate beta in this course.



Indicate the Distribution that Will Be Used

When working with numerical data, we will always work on the students t-distribution to accomplish this and degrees of freedom df = n – 1.
This week, we are working with categorical data and will work on the standard normal distribution—**the Z-Curve**.

This is a left-tailed test.

$H_o: p =$ _____
$H_a: p <$ _____

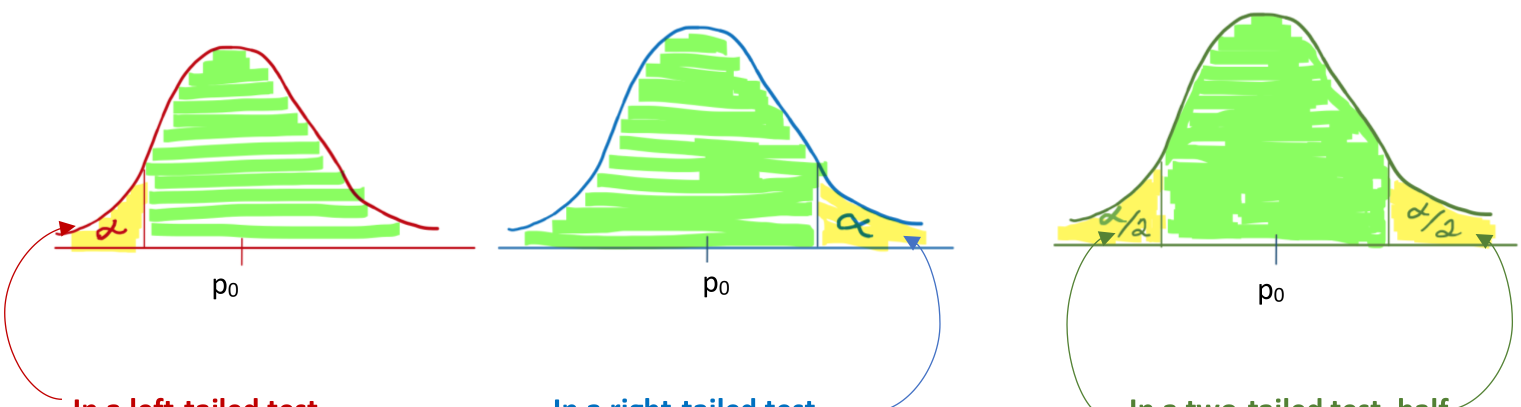
This is a right-tailed test.

$H_o: p =$ _____
$H_a: p >$ _____

This is called a two-tailed test.

$H_o: p =$ _____
$H_a: p \neq$ _____

These are Z-curves with the hypothesized proportion, p_o, is at the center.



In a left-tailed test, alpha is in the left tail.

In a right-tailed test, alpha is in the right-tail.

In a two-tailed test, half of alpha is in each tail.

Here is the plan: We will analyze the sample data and determine if the sample is data is consistent with the hypothesis.

If the sample proportion (p-hat) lands in the green highlighted region, then we conclude that the sample proportion is consistent with the hypothesis; i.e., the sample proportion is ***not statistically different*** from the hypothesized proportion, p_0. This green highlighted region is the region where we would say our sample data is really NOT THAT DIFFERENT from the hypothesized proportion (the proportion we were claiming in the null and at the center of the curve).

If the sample proportion (p-hat) lands in the yellow highlighted region, then we have evidence that the sample proportion is statistically different from the hypothesized proportion (the proportion we were claiming in the null and at the center of the curve). This yellow highlighted region is the region where we would say our sample data is ***statistically different*** from the hypothesized proportion.

To determine where the sample proportion lies relative to the hypothesized proportion, we will find the corresponding **Z-value**. This is called a **Z Test Statistic**.

Sample Step

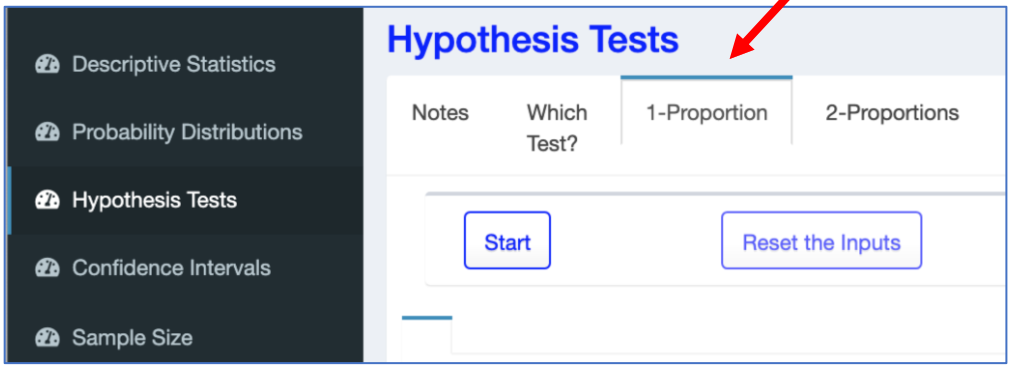
Analyze the Sample Data using Technology

Use statistical technology to conduct the hypothesis test you stated in the Method Step; StatHelper instructions and Minitab 19 instructions can be found on the next pages. Paste the output in your work. The test statistic and P-value outputs are needed.

Check to see if the Normality Assumption Has Been Met

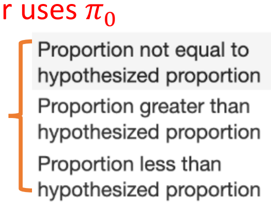
In order to determine if the sample size, n, is large enough, the following calculation will be used. NOTE: The hypothesized **proportion**, p_0 is used in the calculation (since it is sitting at the center of your Z-curve). We can assume a normal distribution if $\boldsymbol{n\,(p_0)(1-\,p_0) \geq 10}$**.**

One Sample Proportion Test –with StatHelper

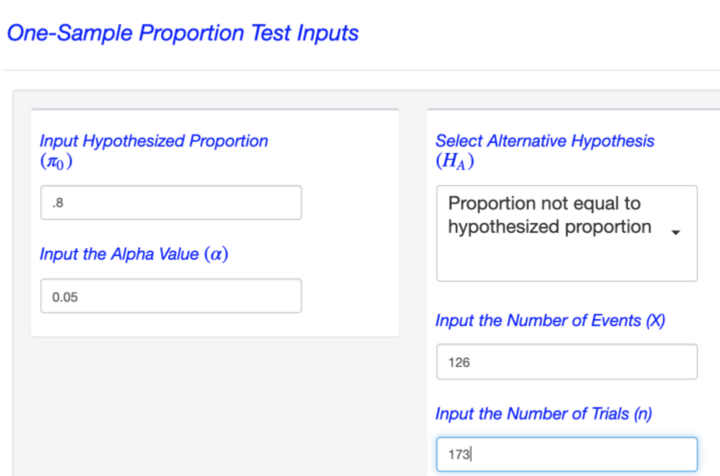


<u>StatHelper Instructions</u>

1. Choose Hypothesis Tests from the left side.
2. Choose **1-Proportion** from the options across the top.
3. Press Start.
4. Choose the form your data takes; Choose Summarized Data (input values)
5. Input the Hypothesized Proportion (p_o)—NOTE StatHelper uses π_0
6. Input the Alpha Value (α).
7. Select Alternative Hypothesis (H_A).



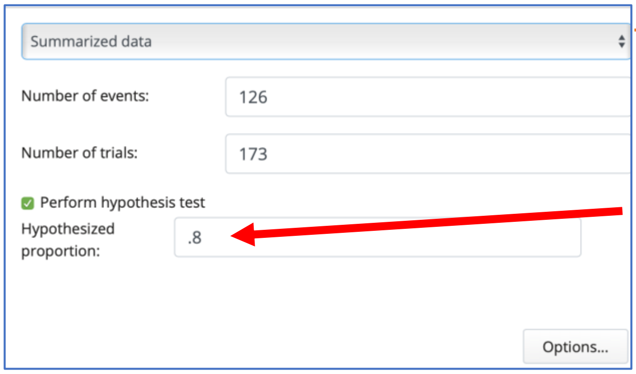
8. Input the Number of Events (X) and the Number of Trails (n)
9. Click RUN. The work is under the WORK tab. The decision and conclusion are under the INTERPRETATION tab



One-Proportion Test –with Minitab 19

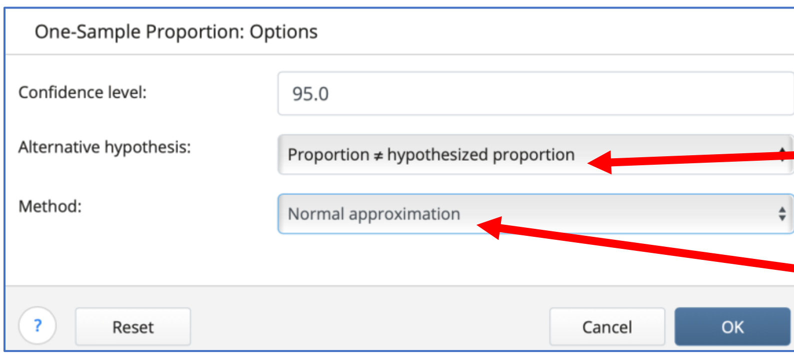
Minitab 19 keystrokes

Stat→Basic Statistics→ 1 Proportion



Choose summarized data.
We will not work with raw categorical data. Enter the number of events and the number of trials.

Select: Perform hypothesis test
Enter p_o, the hypothesized proportion.



Select Options…

(note: confidence level does not need to be adjusted in testing)

Choose the correct Alternative Hypothesis.

The test method MUST BE CHANGED to read Normal approximation.

Results Step

The Z - Test Statistic

A test statistic is a measure of location and represents the number of standard errors (SE) away your sample **proportion** is from what you are assuming in the null hypothesis (the hypothesized **proportion**, p_o).

The formula for the **Z test statistic** is below. Always round the SE calculation to at least 3 decimal places. Round the test statistic calculation to 2 decimal places (round to the nearest hundredth).

$$Z = \frac{\hat{p} - p_o}{\sqrt{\frac{(p_o)(1 - p_o)}{n}}}$$

A negative test statistic means that the sample proportion ($\hat{p}$) is ***<u>less than</u>*** the hypothesized proportion (p_0).

A positive test statistic means that the sample proportion ($\hat{p}$) is ***<u>greater than</u>*** the hypothesized proportion (p_0).

Interpret the Z - Test Statistic

It is important that you are able to interpret the Z - test statistic. You can use this template sentence (note that you will take the absolute value of the test statistic when you interpret it since the positive or negative sign indicates direction).

"My sample proportion is ______ standard errors above/below __________ ."

State $|Z\ test\ stat|$ the hypothesized proportion p_0

- Leave off the +/- sign when interpreting the t test statistic.
- If it is a positive t test statistic, say "above".
- If it is a negative t test statistic, say "below".

If we were to try to make a decision at this point, we would need to know if the test statistic we obtained is statistically far enough away from the center of the curve (the hypothesized mean). This approach is called the 'Critical Value' approach. StatHelper calls this the 'Rejection Region' approach. You can learn more here, if interested. **NOTE** We will <u>not</u> use this approach to make a decision in hypothesis testing.

We will ALWAYS let technology obtain the Probability value or **P-value** associated with the test statistic we calculated. This is called the 'P-value' approach. Our e-text uses this approach, so you can learn more by reading the e-text or keep reading in these notes!

In this course, we will ***<u>ALWAYS</u>*** use the **P-value approach** to making decisions in Hypothesis Testing.

The P-value

A P-value is the probability of getting a statistic at least as extreme as the one that was obtained through sampling, assuming the null hypothesis is true. In other words, the P-value is the probability that, if the null hypothesis is true, the results from another randomly selected sample will be as extreme or more extreme as the results obtained from the given sample.

Here is the most casual way I can say it: The P-value is the area in the tail(s) of the curve beyond the test statistic that you found.

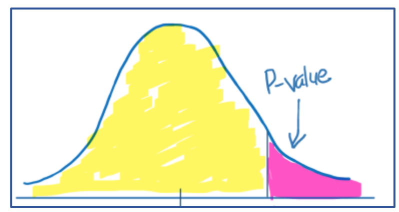
NOTE: Technology calculates the P-value for us. We *could* find it using Minitab's Standard Normal Distribution or Z-table. We do not have to. The P-value will be part of the hypothesis test output.

LEFT-TAILED TEST

If you are conducting a left-tailed test (the H_a has a 'less than' sign in it), then the P-value represents the area in the **left** tail beyond your test statistic.

To calculate a P-value on the Z curve, the statistical technology will find the area beyond the Z test statistic.

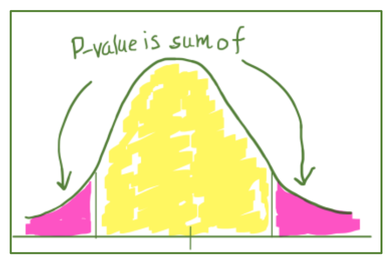
RIGHT-TAILED TEST



If you are conducting a right-tailed test (the H_a has a 'greater than' sign in it), then the P-value represents the area in the **right** tail beyond your test statistic.

To calculate a P-value on the Z curve, statistical technology will find the area beyond the Z test statistic.

TWO-TAILED TEST



If you are conducting a two-tailed test (the H_a has a 'NOT EQUAL TO' sign in it), then the P-value represents the area in the left and right tails beyond your test statistic. The areas are identical since the curve is symmetric. If you find one, just double it to get the P-value.

To calculate a P-value on the Z curve, statistical technology will find the area beyond the test statistic **in both directions.**

Interpret the P-value

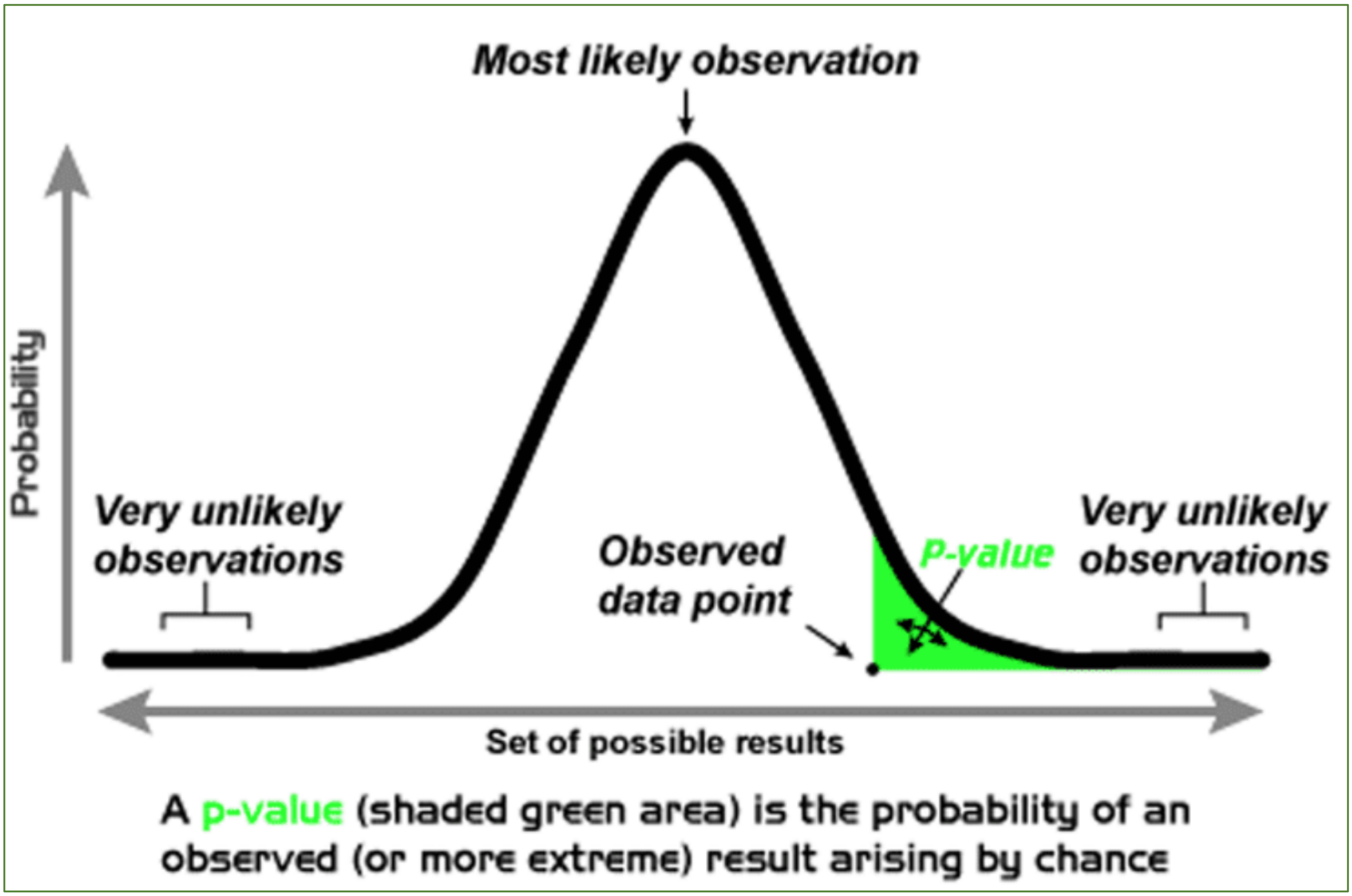
A P-value is the probability of obtaining a statistic ($\bar{x}$ or $\hat{p}$) at least as extreme as the one that was obtained through sampling, ***assuming the null hypothesis is true***.

Our textbook describes it: The *p*-value is the probability that, if the null hypothesis is true, the results from another randomly selected sample will be as extreme or more extreme as the results obtained from the given sample.

NOTE: Technology calculates the P-value for us. The P-value will be part of the hypothesis test output. **We need to be able to interpret the meaning behind the P-value.**

You may use the following P-value interpretation template:

"Assuming the true proportion equals…..p_o….., there is a ….P-value…. probability of getting a sample proportion ($\hat{p}$) at least as extreme as the one we got from sampling."


Most likely observation
Probability
Very unlikely observations
Observed data point
P-value
Very unlikely observations
Set of possible results
A p-value (shaded green area) is the probability of an observed (or more extreme) result arising by chance

Conclusion Step

Make a Decision

Statisticians developed the P-value approach to making a decision in hypothesis testing. [There is another method of making a decision. It is called the rejection region approach or critical value approach. We will not use this method in this course. NOTE: StatHelper does show both the P-value and rejection region approach.]

The P-value Approach

If the P-value $\leq \alpha$, then we will decide that we CAN REJECT the null hypothesis (in favor of the alternative hypothesis).

In other words:
If the P-value is small, then it would be considered a rare event to get a sample statistic ($\bar{x}$ *or* $\hat{p}$) like we did, if the null were true. This is evidence that the null is not true (hence, we CAN reject the null). In other words, if the P-value is less than the significance level (**P-value < α**), then we **can reject Ho**.

If the P-value > α, then we will decide that we CANNOT reject the null hypothesis (in other words, we FAIL TO reject the null or DO NOT reject the null)

In other words:
If the P-value is big enough, then it would **NOT** be considered a RARE event to get a sample statistic ($\bar{x}$ *or* $\hat{p}$) like we did, if the null were true. This is evidence that the null is probably true (hence, we cannot reject the null). In other words, if the P-value is greater than the significance level (**P-value > α**), then we **CANNOT reject Ho**.

To assist you with making a decision, you may do the following three things:

1. First ask and answer this question: Is your P-value ≤ α?

Example: In a test with α = 0.05 and P-value = 0.0397, we would say YES since 0.0397 < 0.05.
Example: In a test with α = 0.05 and P-value = 0.13, we would say NO since .13 > 0.05.

2. Make a Decision.

If your answer is **YES**, then the decision for this test is that **we CAN reject the null hypothesis (H_o).**
If your answer is **NO**, then the decision for this test is that **we CANNOT reject the null hypothesis (H_o).**

If P-value $\leq \alpha$, then we **CAN** reject Ho and, therefore, we will conclude that the sample **DOES** provide sufficient evidence to support the alternative hypothesis (Ha).

If P-value > α, then we **CANNOT** reject Ho and, therefore, we will conclude that the sample **DOES NOT** provide sufficient evidence to support the alternative hypothesis (Ha).

3. Write a Statistical Conclusion using this template.

The following template is a statistical conclusion written in the context of the problem and always indicating if we DO HAVE enough evidence to support the goal or if we DO NOT have enough evidence to support the goal.

"At the ____% level of significance, the sample data DOES / DOES NOT provide sufficient evidence to say that the true **proportion** of ……………{restate the goal in words}…………………".

Example 1 –Two-Tailed Test

The CEO of a large electric utility claims that 80% of his customers are VERY SATISFIED with customer service. To test this claim, the local newspaper randomly surveyed 173 customers and found 126 to be very satisfied. Complete the appropriate hypothesis test at the 5% level of significance. Show the complete testing process.

Population Step

The variable is whether or not they are very satisfied with customer service. It is CATEGORICAL.

p = the true proportion of utility customers that are very satisfied with customer service.

Goal: Test to see if the proportion of customers that are very satisfied is different from 80% ($p \neq .80$).

Method Step

H_o: p = .80

H_a: p ≠ .80

Alpha = 0.05

Z-curve

Sample Step

(173)(.80)(1 - .80) = 27.68 (which is bigger than 10). Based on this calculation, the sample size is large enough to assume a normal sampling distribution.

StatHelper Output (Work Tab) Minitab 19 Output

$Z_{Stat} = -2.36$

For the p-value approach, the p-value is:

$$\text{p-value} = 2P(Z \geq |z|)$$
$$\text{p-value} = 2(1 - P(Z \leq |-2.36|))$$
$$\text{p-value} = 2(1 - 0.9909)$$
$$\text{p-value} = 2 * 0.0091$$
$$\text{p-value} = 0.0183$$

WORKSHEET 1

Test and CI for One Proportion

Method

p: event proportion

Normal approximation method is used for this analysis.

Descriptive Statistics

N	Event	Sample p	95% CI for p
173	126	0.728324	(0.662039, 0.794608)

Test

Null hypothesis H_0: p = 0.8

Alternative hypothesis H_1: p ≠ 0.8

Z-Value	P-Value
-2.36	0.018

Results Step

The test statistic is Z = -2.36

My sample proportion is 2.36 standard errors below the hypothesized proportion of .80.

The P-value = 0.0183

Assuming the true proportion is .80, there is a 0.0183 probability of getting a sample proportion at least as extreme as the one we got from sampling.

Conclusion Step

Since P = 0.0183 is less than α = 0.05, Yes, we can reject the null.

Statistical Conclusion:

At the 5% level of significance, the sample data DOES provide sufficient evidence to say that the true proportion of customers that are very satisfied with customer service is different from .80.

Everyday Conclusion

We have evidence that the CEO is wrong. The percentage is statistically different from 80%. (In fact, it is less)!

Video Discussing Example 1 Hypothesis Test with Proportions from Prof. Coffey

Here is a video of example 1: https://youtu.be/72ycHL44XGM

Here is the same video with interpreting: https://youtu.be/jvFGxIRogWs

Example 2 – Left-Tailed Test

Research conducted a few years ago showed that 35% of UCLA students had travelled outside the US. With the threat of terrorism abroad, UCLA administrators wonder if the percentage that travel outside the US has decreased. A random sample of 100 students was taken and we found that 33 had travelled abroad. Complete the appropriate hypothesis test at the 5% level of significance. Show the complete testing process.

Population Step

The variable is whether or not they have travelled outside the US (abroad). It is CATEGORICAL.
p = the true proportion of UCLA students that have travelled outside the US.
Goal: Test to see if the proportion that travel outside the US has decreased (p < .35).

Method Step

H_o: p = .35
H_a: p < .35 left-tailed test
Alpha = 0.05
Z-curve

Sample Step

(100)(.35)(1 - .35) = 22.75. Based on this calculation, the sample size is large enough to assume a normal sampling distribution.

StatHelper Output (Work Tab)

$Z_{Stat} = -0.42$

For the p-value approach, the p-value is:

$\text{p-value} = P(Z \leq z)$
$\text{p-value} = P(Z \leq -0.42)$
$\text{p-value} = 0.3372$

Minitab 19 Output

WORKSHEET 1
Test and CI for One Proportion

Method

p: event proportion
Normal approximation method is used for this analysis.

Descriptive Statistics

N	Event	Sample p	95% Upper Bound for p
100	33	0.330000	0.407343

Test

Null hypothesis H_0: p = 0.35
Alternative hypothesis H_1: p < 0.35

Z-Value	P-Value
-0.42	0.337

Results Step

Z = -.42
My sample proportion is .42 standard errors below the hypothesized proportion of .35.
P-value = .337
Assuming the true proportion is .35, there is a .337 probability of getting a sample proportion at least as extreme as the one we got from sampling.

Conclusion Step

Compare the P-value to alpha; Is the P-value $\leq$ alpha? No, it is not.
We cannot reject the null hypothesis.

At the 5% level of significance, the sample data DOES NOT provide sufficient evidence to say that the true proportion of UCLA students that have travelled outside the US is less than .35

Everyday Conclusion

Yes, the percentage has decreased! (We are unable to say WHY it has decreased…we do not have that information.)

Example 3 – Right-Tailed Test

A potential side-effect of taking the drug Lipitor is flu-like symptoms. Most doctors believe that this is a very rare side-effect. A researcher hopes to show the medical community that over 1.5% of the people who take the drug Lipitor experience flu-like symptoms. A random sample of 1165 people who take Lipitor were asked whether or not they experience flu-like symptoms as a side effect and 25 indicated that they did. Show a complete hypothesis test using a 1% level of significance.

Population Step

The variable is whether or not they get the flu-like symptoms as a side effect. It is CATEGORICAL

p = the true proportion of people who take Lipitor that get flu-like symptoms

GOAL: To test whether there is support for the true proportion being more than .015.

Method Step

H_o: p = .015

H_a: p > .015 Right-tailed test

Alpha = 0.05

Z-curve

Sample Step

(1165)(.015)(1 - .015) = 17.21. Based on this calculation, the sample size is large enough to assume a normal sampling distribution.

StatHelper Output (Work Tab) Minitab 19 Output

$Z_{Stat} = 1.81$

For the p-value approach, the p-value is:

$$\text{p-value} = P(Z \geq z)$$
$$\text{p-value} = 1 - P(Z \leq 1.81)$$
$$\text{p-value} = 1 - 0.9649$$
$$\text{p-value} = 0.0351$$

WORKSHEET 1

Test and CI for One Proportion

Method

p: event proportion

Normal approximation method is used for this analysis.

Descriptive Statistics

N	Event	Sample p	95% Lower Bound for p
1165	25	0.021459	0.014476

Test

Null hypothesis H_0: p = 0.015

Alternative hypothesis H_1: p > 0.015

Z-Value	P-Value
1.81	0.035

Z = 1.81

My sample proportion is 1.81 standard errors above the hypothesized proportion of .015.

P-value = .035

Assuming the true proportion is 0.015, there is a .035 probability of getting a sample proportion at least as extreme as the one we got from sampling.

Conclusion Step

Compare the P-value to alpha; Is the P-value ≤ alpha 0.01? No, it is not.

We cannot reject the null hypothesis.

At the 1% level of significance, the sample data DOES NOT provide sufficient evidence to say that the true proportion of people who take Lipitor that get flu-like symptoms is greater than .015.

Everyday Conclusion

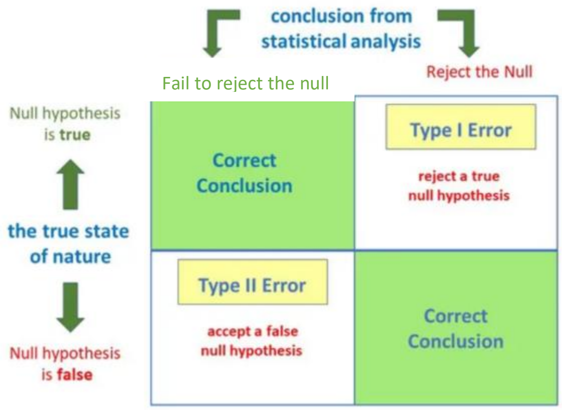
The researcher did not find the evidence to show that more than 1.5% of Lipitor users get flu-like symptoms.

Type I Error and Type II Error

This website has been referenced.

Type I Error is the probability of deciding we **CAN** reject the H_o when, in fact, that decision to reject is actually a mistake. Alpha, α, is the probability of Type I Error.

Type II Error is the probability of deciding we **CANNOT** reject the H_o when, in fact, that decision is actually a mistake. Beta, β, is the probability of Type II Error. We will not learn to calculate beta in this course.



How does type I error occur?
A type I error is also known as a false positive and occurs when a researcher incorrectly rejects a true null hypothesis. This means that your report that your findings are significant when in fact they have occurred by chance.

The probability of making a type I error is represented by your alpha level (α), which is the *p*-value below which you reject the null hypothesis. An alpha of 0.05 indicates that you are willing to accept a 5% chance that you are wrong when you reject the null hypothesis.

You can reduce your risk of committing a type I error by using a lower value for α. For example, an alpha-value of 0.01 would mean there is a 1% chance of committing a type I error. However, using a lower value for alpha means that you will be less likely to detect a true difference if one really exists (thus risking a type II error).

How does type II error occur?
A type II error is also known as a false negative and occurs when a researcher cannot reject a null hypothesis which is really false. Here a researcher concludes there is not a significant effect, when actually there really is. The probability of making a type II error is called Beta (β), and this is related to the power of the statistical test (power = 1- β). You can decrease your risk of committing a type II error by ensuring your test has enough power. You can do this by ensuring your sample size is large enough to detect a practical difference when one truly exists.

Why are type I and type II errors important?
The consequences of making a type I error mean that changes or interventions are made which are unnecessary, and thus waste time, resources, etc.
Type II errors typically lead to the preservation of the status quo (i.e. interventions remain the same) when change is needed.

Example 4 --Describe Type I and Type II Errors and Consequences

A particular compound is not hazardous in drinking water if it is present at a rate of 25ppm (or less). A watchdog group believes that a certain water source does not meet this standard.

μ: the true mean amount of the compound (in ppm)
H_0: $\mu = 25$
H_a: $\mu > 25$

If the watchdog group decides to gather data and formally conduct this test, describe type I and type II errors in the context of this scenario and the consequences of each.

Type I error

Stating that the evidence indicates the water is unsafe when, in fact, it is safe. The watchdog group will have potentially initiated a clean-up where none was required ($$ wasted).

Type II error

Stating that there is no evidence that the water is unsafe when, in fact, it is unsafe. The opportunity to note (and repair) a potential health risk will be missed.

Example 5 --Describe Type I and Type II Errors and Consequences

A lobbying group has a been advocating a particular ballot proposal. One week before the election, they are considering moving some of their advertising efforts to other issues. If the proposal has a support level of at least 55%, they will feel it's "safe" and move money to other campaigns.

p: the true proportion of people who support the proposal
H_0: $p = .55$
H_a: $p < .55$

If the lobbying group decides to gather data and formally conduct this test, describe type I and type II errors in the context of this scenario and the consequences of each.

Type I error

Stating that the evidence indicates the support level is less than 55% (and the proposal may be in jeopardy of failing) when that is not the case. The lobbying group will have kept advertising dollars aimed at this proposal when they could have been spent elsewhere.

Type II error

Stating that the proposal appears to have a "safe" level of support when that is not the case. The lobbying group would shift funds away from supporting this proposal even though it may still be in need of that support.

Video Discussing Type I and II Errors from Prof. Coffey

Here is a video about type I and II errors: https://youtu.be/daa_fDuUIA8

Here is the same video with interpreting: https://youtu.be/aG935FY7UpA