STA 2023

Module 8

Testing Hypothesis

Learning Objectives

Upon completing this module, you should be able to

1. Define the terms associated with hypothesis testing.
2. Choose the null and alternative hypotheses for a hypothesis test.
3. Explain the logic behind hypothesis testing.
4. Identify the test statistic, rejection region, non-rejection region, and critical value(s) for a hypothesis test.
5. State and interpret the possible conclusions for a hypothesis test.

Learning Objectives

6. Obtain the critical value(s) for a specified significance level.
7. Perform a large-sample hypothesis test for a population proportion.
8. Obtain the P-value of a hypothesis test.
9. State and apply the steps for performing a hypothesis test, using the critical-value approach to hypothesis testing.
10. State and apply the steps for performing a hypothesis test, using the P-value approach to hypothesis testing.

How a Statistic is Used?

In the previous module, we looked at methods for obtaining the confidence intervals for one population proportion.

In this module, we are going to look at how a statistic (sample proportion) is used to make decisions about hypothesized values of a parameter (population proportion.)

Decision and Hypothesis Test

One of the commonly used methods for making decision is to perform a hypothesis test

A hypothesis test involves two hypotheses: the null hypothesis and the alternative hypothesis.

"Testing a hypothesis" is like "testing a claim."

What are Hypotheses?

- Hypotheses are working models that we adopt temporarily.
- Our starting hypothesis is called the null hypothesis.
- The null hypothesis, that we denote by H_0, specifies a population model parameter of interest and proposes a value for that parameter.
- We usually write down the null hypothesis in the form H_0: *parameter* = *hypothesized* value.
- The alternative hypothesis, which we denote by H_A, contains the value of the parameter that we consider plausible when we reject the null hypothesis.

Testing Hypotheses

- The first step in defining the null and alternative hypotheses is to determine which parameter is being tested. A parameter describes a population. Examples are the population proportion, the population standard deviation and a population mean.
- The next step is to define the null hypothesis, specifies a population model parameter of interest and proposes a value for that parameter.
  - We might have, for example, H_0: *parameter* = 0.20
  - We want to compare our data to what we would expect, given that H_0 is true.
- Note that the null hypothesis is a statement of equality (with the equal sign.)

A Trial as a Hypothesis Test

Think about the logic of jury trials:

- To prove someone is guilty, we start by *assuming* they are innocent.
- We retain that hypothesis until the facts make it unlikely beyond a reasonable doubt.
- Then, and only then, we reject the hypothesis of innocence and declare the person guilty.

A Trial as a Hypothesis Test (cont.)

The same logic used in jury trials is used in statistical tests of hypotheses:

- We begin by assuming that a hypothesis is true.
- Next we consider whether the data are consistent with the hypothesis.
- If they are, all we can do is retain the hypothesis we started with. If they are not, then like a jury, we ask whether they are unlikely beyond a reasonable doubt.

What are Null Hypothesis, Alternative Hypothesis and Hypothesis Test, again?

Null and Alternative Hypotheses; Hypothesis Test

Null hypothesis: A hypothesis to be tested. We use the symbol H_0 to represent the null hypothesis.

Alternative hypothesis: A hypothesis to be considered as an alternative to the null hypothesis. We use the symbol H_a to represent the alternative hypothesis.

Hypothesis test: The problem in a hypothesis test is to decide whether the null hypothesis should be rejected in favor of the alternative hypothesis.

What to Do with an "Innocent" Defendant?

If the evidence is not strong enough to reject the presumption of innocent, the jury returns with a verdict of "not guilty."

- The jury does not say that the defendant is innocent.
- All it says is that there is not enough evidence to convict, to reject innocence.
- The defendant may, in fact, be innocent, but the jury has no way to be sure.

What to Do with an "Innocent" Defendant ?(cont.)

- Said statistically, we will *fail to reject* the null hypothesis.
  - We never declare the null hypothesis to be true, because we simply do not know whether it's true or not.
  - Sometimes in this case we say that the *null hypothesis has been retained*.

What to Do with an "Innocent" Defendant ?(cont.)

- In a trial, the burden of proof is on the prosecution.
- In a hypothesis test, the burden of proof is on the unusual claim.
- The null hypothesis is the ordinary state of affairs, so it's the alternative to the null hypothesis that we consider unusual (and for which we must marshal evidence).

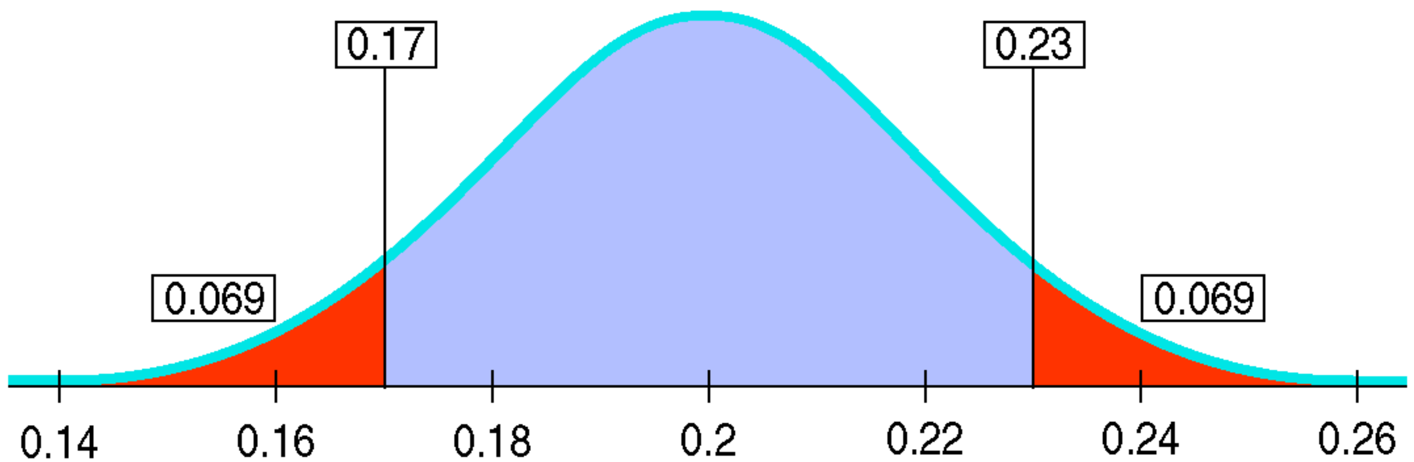
What are the Three Possible Alternative Hypotheses?

There are three possible alternative hypotheses:

- H_A: *parameter* < *hypothesized value*
- H_A: *parameter* ≠ *hypothesized value*
- H_A: *parameter* > *hypothesized value*

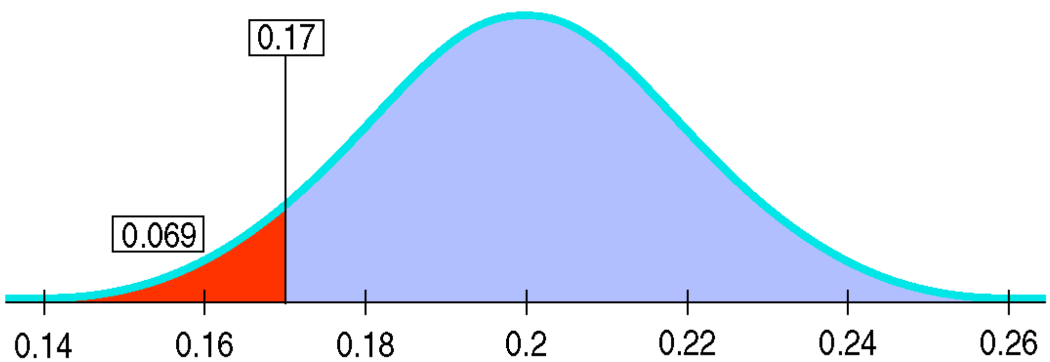
What is a Two-Sided Alternative Hypothesis?

- H_A: *parameter ≠ value* is known as a two-sided alternative or two-tailed alternative because we are equally interested in deviations on either side of the null hypothesis value.
- For two-sided alternatives, the P-value is the probability of deviating in *either* direction from the null hypothesis value.

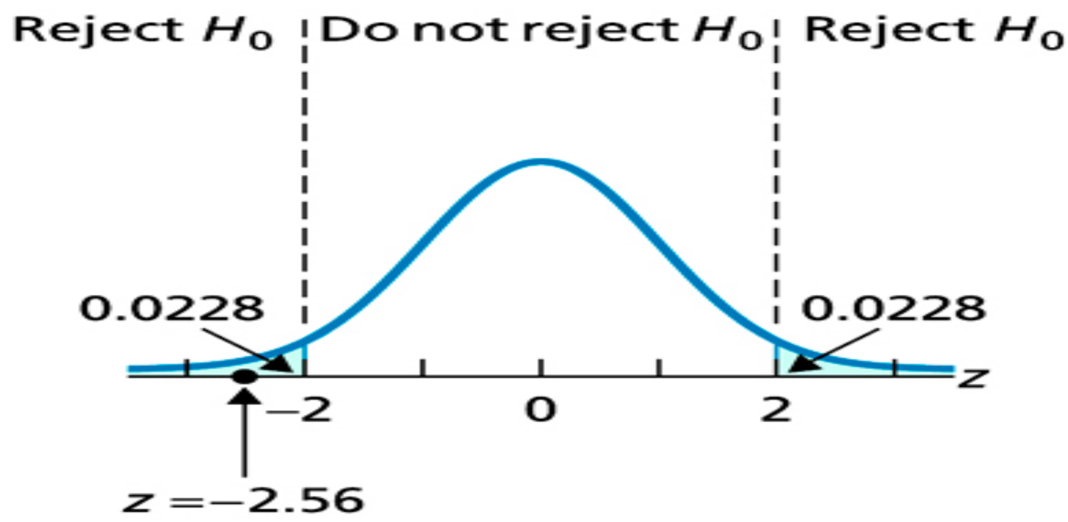


What is One-Sided Alternative Hypothesis?

- The other two alternative hypotheses are called one-sided alternatives or one-tailed alternative
- A one-sided alternative focuses on deviations from the null hypothesis value in only one direction.
- Thus, the P-value for one-sided alternatives is the probability of deviating *only in the direction of the alternative* away from the null hypothesis value.



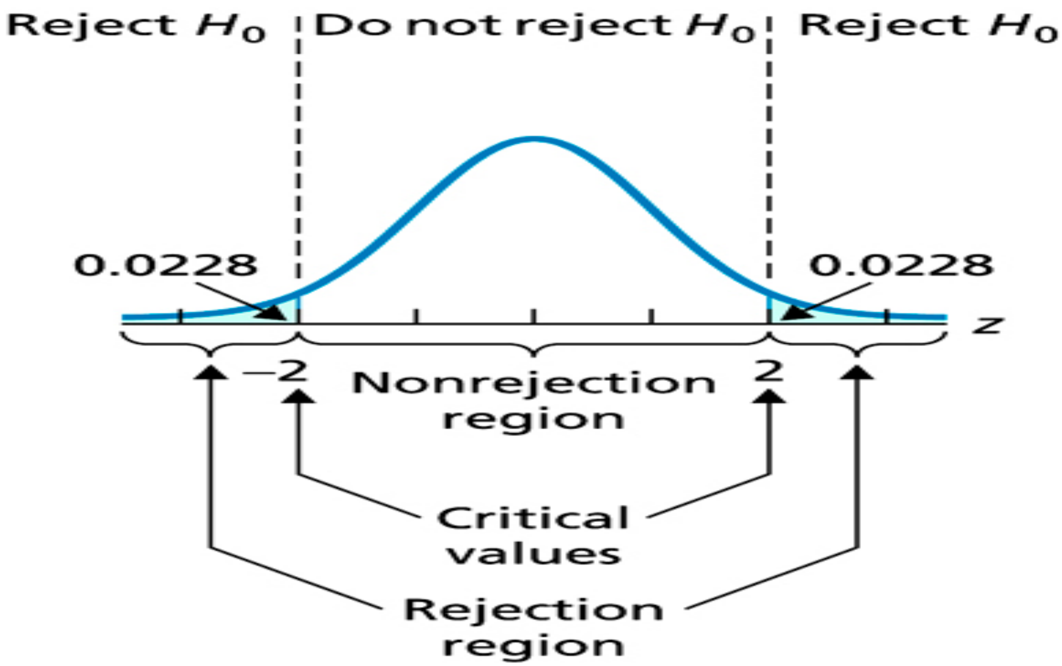
Critical Value and Decision



The rejection region (shaded area) is based on the chosen significance level, which determines the critical value -2 and 2.

When the value of the test-statistic (z = -2.56) falls in this rejection region, we reject the null hypothesis.

Critical Value and Rejection Region (Cont.)



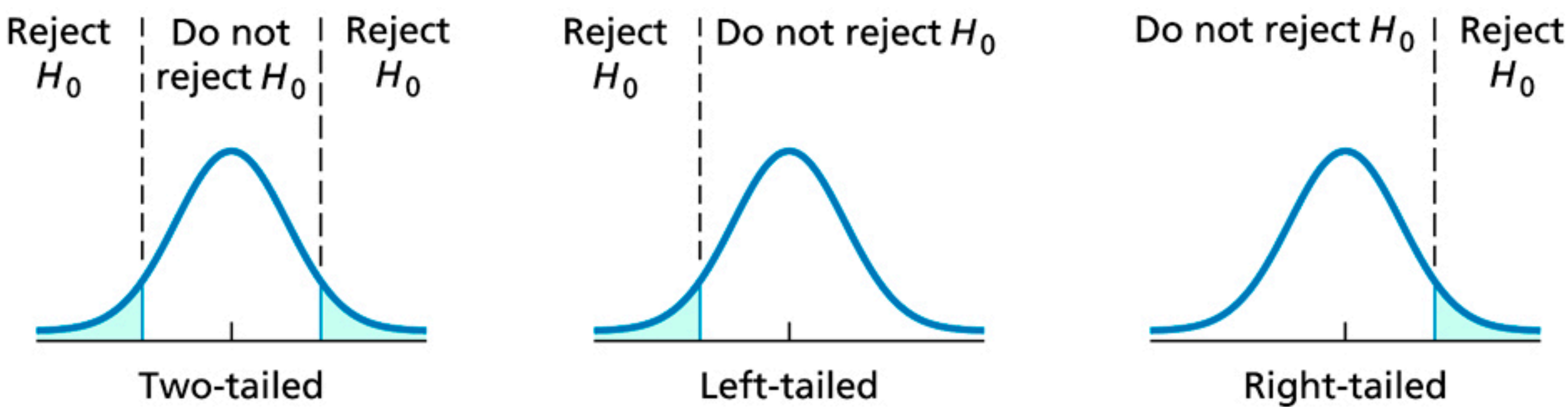
How to obtain the critical value -2 in this case (by TI-84+)?

Press the following keys: [2nd] [DISTR]

Then perform: invNorm(0.0228)

The critical value(s) can be obtained from the standard-normal table or technology, based on the chosen significance level. In this case, the significance level is 4.56% (by adding 0.0228 from both tails will give us 0.0456 or 4.56%). Since this is a two-tailed alternative/two-tailed test, we divide the significance level by two for each tail.

Where is the Rejection Region?



For a two-tailed or two-sided test, the rejection region is on both the left and right.

For a left-tailed test, the rejection region is on the left. For a right-tailed test, the rejection region is on the right.

Note that the rejection region is always at the tail(s).

Four Basic Parts to a Hypothesis Test

- There are four basic parts to a hypothesis test:
  1. Hypotheses
  2. Model
  3. Mechanics
  4. Conclusion
- Let's look at these parts in detail…

Four Basic Parts (cont.)

1. Hypotheses
   - The null hypothesis: To perform a hypothesis test, we must first translate our question of interest into a statement about model parameters.
     - In general, we have

       H_0: *parameter* = *hypothesized value*.
   - The alternative hypothesis: The alternative hypothesis, H_A, contains the values of the parameter we consider plausible if we reject the null.

Four Basic Parts (cont.)

2. Model

- To plan a statistical hypothesis test, specify the *model* you will use to test the null hypothesis and the parameter of interest.
- All models require assumptions, so state the assumptions and check any corresponding conditions.
- Your plan should end with a statement like
    - *Because the conditions are satisfied, I can model the sampling distribution of the proportion with a Normal model.*
    - Watch out, though. It might be the case that your model step ends with "*Because the conditions are not satisfied, I can't proceed with the text.*" If that's the case, stop and reconsider.

Four Basic Parts (cont.)

2. Model
   - Each test we discuss has a name that you should include in your report.
   - The test about proportions is called a one-proportion z-test.

One-Proportion *z*-Test

- The conditions for the one-proportion *z*-test are the same as for the one proportion *z*-interval. We test the hypothesis H$_0$: $p = p_0$

  using the statistic $$z = \frac{(\hat{p} - p_0)}{SD(\hat{p})}$$

  where $$SD(\hat{p}) = \sqrt{\frac{p_0 q_0}{n}}$$

- When the conditions are met and the null hypothesis is true, this statistic follows the standard Normal model, so we can use that model to obtain a P-value.

Four Basic Parts (cont.)

3. Mechanics
   - Under "mechanics" we place the actual calculation of our test statistic from the data.
   - Different tests will have different formulas and different test statistics.
   - Usually, the mechanics are handled by a statistics program or calculator, but it's good to know the formulas.

Four Basic Parts (cont.)

3. Mechanic

   The ultimate goal of the calculation is to obtain a P-value.

   – The P-value is the probability that the observed statistic value (or an even more extreme value) could occur if the null model were correct.
   – If the P-value is small enough, we'll reject the null hypothesis.

Note: The P-value is a conditional probability—it's the probability that the observed results could have happened *if the null hypothesis is true*.

Four Basic Parts (cont.)

4. Conclusion
   - The conclusion in a hypothesis test is always a statement about the null hypothesis.
   - The conclusion must state either that we reject or that we fail to reject the null hypothesis.
   - And, as always, the conclusion should be stated in context.
   - Your conclusion about the null hypothesis should never be the end of a testing procedure.
   - Often there are actions to take or policies to change.

How to Perform a One Proportion *z*-Test?

One-Proportion *z*-Test

Purpose To perform a hypothesis test for a population proportion, p

Assumptions

1. Simple random sample

2. Both np_0 and $n(1 - p_0)$ are 5 or greater

STEP 1 **The null hypothesis is H_0: $p = p_0$, and the alternative hypothesis is**

H_a: $p \neq p_0$	or	H_a: $p < p_0$	or	H_a: $p > p_0$
(Two tailed)		**(Left tailed)**		**(Right tailed)**

STEP 2 **Decide on the significance level, α.**

STEP 3 **Compute the value of the test statistic**

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}$$

and denote that value z_0.

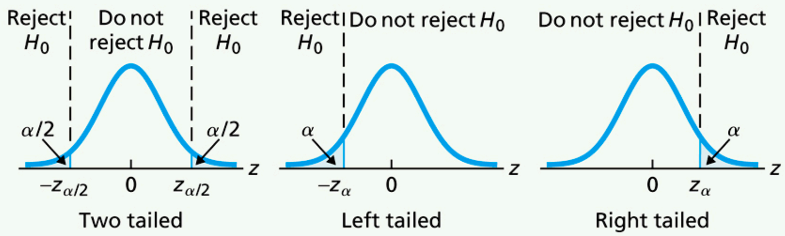
How to Perform a One Proportion *z*-Test ? (Cont.)

CRITICAL-VALUE APPROACH	or	*P*-VALUE APPROACH

STEP 4 The critical value(s) are

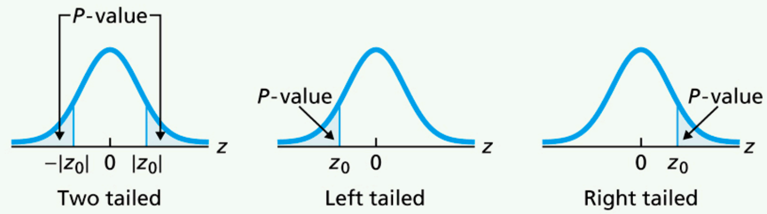
$\pm z_{\alpha/2}$ (Two tailed) or $-z_{\alpha}$ (Left tailed) or z_{α} (Right tailed)

Use Table II to find the critical value(s).



STEP 5 If the value of the test statistic falls in the rejection region, reject H_0; otherwise, do not reject H_0.

STEP 4 Use Table II to obtain the *P*-value.



STEP 5 If $P \leq \alpha$, reject H_0; otherwise, do not reject H_0.

STEP 6 Interpret the results of the hypothesis test.

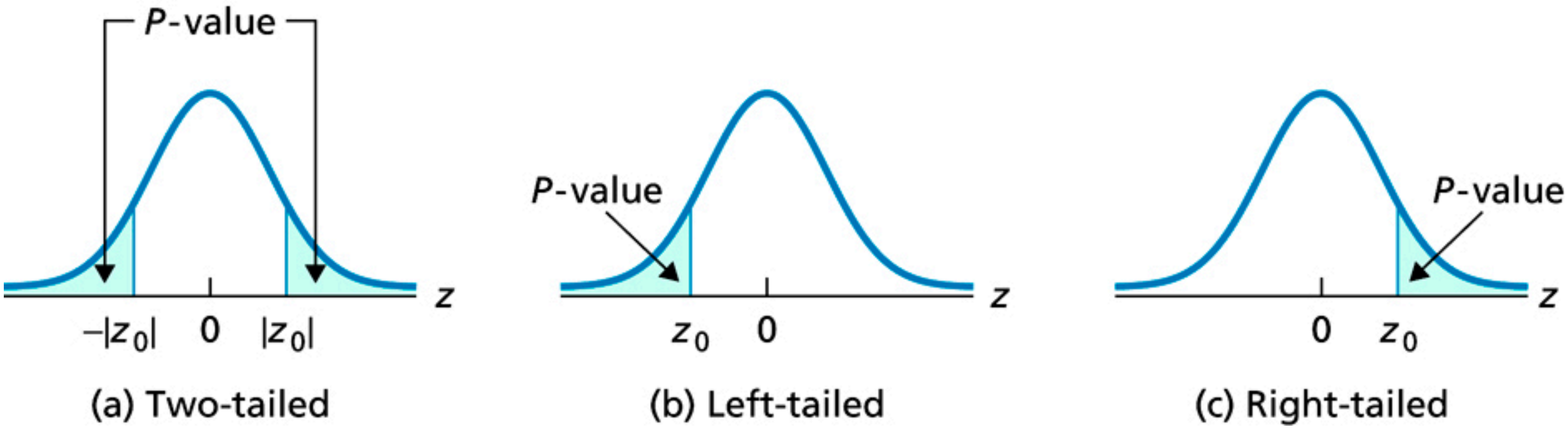
What is P-Value?

- The statistical twist is that we can quantify our level of doubt.
  - We can use the model proposed by our hypothesis to calculate the probability that the event we've witnessed could happen.
  - That's just the probability we're looking for—it quantifies exactly how surprised we are to see our results.
  - This probability is called a P-value.

What is P-Value? (cont.)

- When the data are consistent with the model from the null hypothesis, the P-value is high and we are unable to reject the null hypothesis.
  - In that case, we have to "retain" the null hypothesis we started with.
  - We can't claim to have proved it; instead we "*fail to reject the null hypothesis*" when the data are consistent with the null hypothesis model and in line with what we would expect from natural sampling variability.
- If the P-value is low enough, we will "*reject the null hypothesis*," since what we observed would be very unlikely were the null model true.

Where is the P-value?



Note that P-value is the probability at the tail(s). If the P-value is small, we reject the null hypothesis.

P-Values and Decisions: What to Tell About a Hypothesis Test

- How small should the P-value be in order for you to reject the null hypothesis?
- It turns out that our decision criterion is context-dependent.
  - When we're screening for a disease and want to be sure we treat all those who are sick, we may be willing to reject the null hypothesis of no disease with a fairly large P-value.
  - A longstanding hypothesis, believed by many to be true, needs stronger evidence (and a correspondingly small P-value) to reject it.
- Another factor in choosing a P-value is the importance of the issue being tested.

P-Values and Decisions (cont.)

- Your conclusion about any null hypothesis should be accompanied by the P-value of the test.
  - If possible, it should also include a confidence interval for the parameter of interest.
- Don't just declare the null hypothesis rejected or not rejected.
  - Report the P-value to show the strength of the evidence against the hypothesis.
  - This will let each reader decide whether or not to reject the null hypothesis.

Any Guidelines on How to Use P-value?

P-value	Evidence against H_0
$P > 0.10$	Weak or none
$0.05 < P \leq 0.10$	Moderate
$0.01 < P \leq 0.05$	Strong
$P \leq 0.01$	Very strong

Here is a Comparison of Critical-Value and P-Value Approach

CRITICAL-VALUE APPROACH	or	*P*-VALUE APPROACH
STEP 1 State the null and alternative hypotheses.		**STEP 1** State the null and alternative hypotheses.
STEP 2 Decide on the significance level, α.		**STEP 2** Decide on the significance level, α.
STEP 3 Compute the value of the test statistic.		**STEP 3** Compute the value of the test statistic.
STEP 4 Determine the critical value(s).		**STEP 4** Determine the P-value.
STEP 5 If the value of the test statistic falls in the rejection region, reject H_0; otherwise, do not reject H_0.		**STEP 5** If $P \leq \alpha$, reject H_0; otherwise, do not reject H_0.
STEP 6 Interpret the result of the hypothesis test.		**STEP 6** Interpret the result of the hypothesis test.

What Can Go Wrong?

- Don't base your null hypothesis on what you see in the data.
  - *Think* about the situation you are investigating and develop your null hypothesis appropriately.
- Don't base your alternative hypothesis on the data, either.
  - Again, you need to *Think* about the situation.

What Can Go Wrong? (cont.)

- Don't make your null hypothesis what you want to show to be true.
  - You can reject the null hypothesis, but you can never "accept" or "prove" the null.
- Don't forget to check the conditions.
  - We need randomization, independence, and a sample that is large enough to justify the use of the Normal model.

What have we learned?

We have learned to:

1. Define the terms associated with hypothesis testing.
2. Choose the null and alternative hypotheses for a hypothesis test.
3. Explain the logic behind hypothesis testing.
4. Identify the test statistic, rejection region, non-rejection region, and critical value(s) for a hypothesis test.
5. State and interpret the possible conclusions for a hypothesis test.

What have we learned? (Cont.)

6. Obtain the critical value(s) for a specified significance level.
7. Perform a large-sample hypothesis test for a population proportion
8. Obtain the P-value of a hypothesis test.
9. State and apply the steps for performing a hypothesis test, using the critical-value approach to hypothesis testing.
10. State and apply the steps for performing a hypothesis test, using the P-value approach to hypothesis testing.

Credit

Some of these slides have been adapted/modified in part/whole from the slides of the following textbooks.

- Weiss, Neil A., Introductory Statistics, 8th Edition
- Bock, David E., Stats: Data and Models, 3rd Edition