Basics of Hypothesis Testing

Colby Community College

Definition

In statistics, a **hypothesis** is a claim or statement about a property of a population.

Definition

In statistics, a **hypothesis** is a claim or statement about a property of a population.

Definition

A **hypothesis test** (or **test of significance**) is a procedure for testing a claim about a property of a population.

Definition

In statistics, a **hypothesis** is a claim or statement about a property of a population.

Definition

A **hypothesis test** (or **test of significance**) is a procedure for testing a claim about a property of a population.

Note

the "property of a population" is often the value of a population parameter.

Example 1

In a Pitney Bowes survey, 1009 consumers were asked if they are comfortable with having drones deliver their purchases, and 54% of them responded “no.”

Example 1

In a Pitney Bowes survey, 1009 consumers were asked if they are comfortable with having drones deliver their purchases, and 54% of them responded "no."

Consider the claim that "the majority of consumers are not comfortable with drone deliveries."

Example 1

In a Pitney Bowes survey, 1009 consumers were asked if they are comfortable with having drones deliver their purchases, and 54% of them responded "no."

Consider the claim that "the majority of consumers are not comfortable with drone deliveries."

Using p to denote the proportion of consumers not comfortable with drone deliveries, the "majority" claim is equivalent to the claim that the proportion is greater than half, or $p > 0.5$. The expression $p > 0.5$ is the symbolic form of the original claim.

Example 1

In a Pitney Bowes survey, 1009 consumers were asked if they are comfortable with having drones deliver their purchases, and 54% of them responded "no."

Consider the claim that "the majority of consumers are not comfortable with drone deliveries."

Using p to denote the proportion of consumers not comfortable with drone deliveries, the "majority" claim is equivalent to the claim that the proportion is greater than half, or $p > 0.5$. The expression $p > 0.5$ is the symbolic form of the original claim.

We need to know if 545 is significantly high. In other words, do we have $P\,(545 \text{ or more consumers}) < 0.05$?

Example 1

In a Pitney Bowes survey, 1009 consumers were asked if they are comfortable with having drones deliver their purchases, and 54% of them responded "no."

Consider the claim that "the majority of consumers are not comfortable with drone deliveries."

Using p to denote the proportion of consumers not comfortable with drone deliveries, the "majority" claim is equivalent to the claim that the proportion is greater than half, or $p > 0.5$. The expression $p > 0.5$ is the symbolic form of the original claim.

We need to know if 545 is significantly high. In other words, do we have $P\,(545 \text{ or more consumers}) < 0.05$?

Using technology we have

$$P\,(545 \text{ or more consumers}) \approx 0.005386$$

Definition

The **null hypothesis** is a statement that the value of the population parameter (such as proportion, mean, or standard deviation) is equal to some claimed value.

Definition

The **null hypothesis** is a statement that the value of the population parameter (such as proportion, mean, or standard deviation) is equal to some claimed value.

Definition

The **alternative hypotheses** is a statement that the parameter has a value that somehow differs from the null hypotheses.

Definition

The **null hypothesis** is a statement that the value of the population parameter (such as proportion, mean, or standard deviation) is equal to some claimed value.

Definition

The **alternative hypotheses** is a statement that the parameter has a value that somehow differs from the null hypotheses.

Note

For the methods of this chapter, the symbolic form of the alternative hypotheses must use one of these symbols: $<$, $>$, $\neq$.

Definition

The **null hypothesis** is a statement that the value of the population parameter (such as proportion, mean, or standard deviation) is equal to some claimed value.

Definition

The **alternative hypotheses** is a statement that the parameter has a value that somehow differs from the null hypotheses.

Note

For the methods of this chapter, the symbolic form of the alternative hypotheses must use one of these symbols: $<$, $>$, $\neq$.

Notation

The null hypotheses is denoted by $\boldsymbol{H_0}$.

The alternative hypotheses is denoted $\boldsymbol{H_1}$ or $\boldsymbol{H_a}$ or $\boldsymbol{H_A}$.

Example 2

Here is an example of a null hypothesis involving a proportion:

$$H_0: \; p = 0.5$$

Example 2

Here is an example of a null hypothesis involving a proportion:

$$H_0: \; p = 0.5$$

Example 3

Here are different examples of alternative hypotheses involving proportions:

$$H_A: \; p > 0.5, \quad H_A: \; p < 0.5, \quad H_A: \; p \neq 0.5$$

Example 2

Here is an example of a null hypothesis involving a proportion:

$$H_0: \; p = 0.5$$

Example 3

Here are different examples of alternative hypotheses involving proportions:

$$H_A: \; p > 0.5, \quad H_A: \; p < 0.5, \quad H_A: \; p \neq 0.5$$

Example 4

Returning to Example 1, for

"the majority of consumers are not comfortable with drone deliveries."

we have the hypotheses:

$$H_0: \; p = 0.5$$
$$H_A: \; p > 0.5$$

Note

If you are conducting a study and want to use a hypothesis test to *support* your claim, your claim must be worded such that it becomes the alternative hypothesis and can be expressed using only the symbols $>$, $<$, or $\neq$.

Note

If you are conducting a study and want to use a hypothesis test to *support* your claim, your claim must be worded such that it becomes the alternative hypothesis and can be expressed using only the symbols $>$, $<$, or $\neq$.

Caution

You *never* support a claim that a parameter is equal to a specified value.

Definition

The **significance level** α for a hypothesis test is the probability value used as the cutoff for determining when the sample evidence constitutes significant evidence against the null hypothesis.

Definition

The **significance level α** for a hypothesis test is the probability value used as the cutoff for determining when the sample evidence constitutes significant evidence against the null hypothesis.

Note

By its nature, the significance level α is the probability of mistakely rejecting the null hypothesis when it is true:

$$\text{Significance level } \alpha = P\,(\text{rejecting } H_0 \text{ when } H_0 \text{ is true})$$

Definition

The **significance level $\boldsymbol{\alpha}$** for a hypothesis test is the probability value used as the cutoff for determining when the sample evidence constitutes significant evidence against the null hypothesis.

Note

By its nature, the significance level α is the probability of mistakely rejecting the null hypothesis when it is true:

$$\text{Significance level } \alpha = P\left(\text{rejecting } H_0 \text{ when } H_0 \text{ is true}\right)$$

Note

The significance level α is the same α we talked about in Chapter 7, when discussing confidence intervals.

Definition

The **test statistic** is a value used in making a decision about the null hypotheses. It is found by converting the sample statistic to a score with the assumption that the null hypothesis is true.

Definition

The **test statistic** is a value used in making a decision about the null hypotheses. It is found by converting the sample statistic to a score with the assumption that the null hypothesis is true.

Test Statistic for Proportion p

Sampling Distribution: Normal (z)

Requirements: $np \geq 5$ and $nq \geq 5$

Test Statistic: $z = \dfrac{\hat{p} - p}{\sqrt{\dfrac{pq}{n}}}$

Example 5

In Example 1 we made a claim about the population proportion p, where we have $n = 1009$ and $x = 545$.

Example 5

In Example 1 we made a claim about the population proportion p, where we have $n = 1009$ and $x = 545$.

This means we have

$$\hat{p} = \frac{x}{n}$$

Example 5

In Example 1 we made a claim about the population proportion p, where we have $n = 1009$ and $x = 545$.

This means we have

$$\hat{p} = \frac{x}{n} = \frac{545}{1009}$$

Example 5

In Example 1 we made a claim about the population proportion p, where we have $n = 1009$ and $x = 545$.

This means we have

$$\hat{p} = \frac{x}{n} = \frac{545}{1009} = 0.540$$

Example 5

In Example 1 we made a claim about the population proportion p, where we have $n = 1009$ and $x = 545$.

This means we have

$$\hat{p} = \frac{x}{n} = \frac{545}{1009} = 0.540$$

The null hypotheses is $H_0: \ p = 0.5$.

Example 5

In Example 1 we made a claim about the population proportion p, where we have $n = 1009$ and $x = 545$.

This means we have

$$\hat{p} = \frac{x}{n} = \frac{545}{1009} = 0.540$$

The null hypotheses is $H_0: \ p = 0.5$.

Which means we are working form the assumption that $p = 0.5$ and so $q = 1 - p = 0.5$.

Example 5

In Example 1 we made a claim about the population proportion p, where we have $n = 1009$ and $x = 545$.

This means we have

$$\hat{p} = \frac{x}{n} = \frac{545}{1009} = 0.540$$

The null hypotheses is $H_0 : \; p = 0.5$.

Which means we are working form the assumption that $p = 0.5$ and so $q = 1 - p = 0.5$.

The test statistic is then

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

Example 5

In Example 1 we made a claim about the population proportion p, where we have $n = 1009$ and $x = 545$.

This means we have

$$\hat{p} = \frac{x}{n} = \frac{545}{1009} = 0.540$$

The null hypotheses is $H_0: \; p = 0.5$.

Which means we are working form the assumption that $p = 0.5$ and so $q = 1 - p = 0.5$.

The test statistic is then

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0.540 - 0.5}{\sqrt{\frac{0.5 \cdot 0.5}{1009}}}$$

Example 5

In Example 1 we made a claim about the population proportion p, where we have $n = 1009$ and $x = 545$.

This means we have

$$\hat{p} = \frac{x}{n} = \frac{545}{1009} = 0.540$$

The null hypotheses is $H_0 : \; p = 0.5$.

Which means we are working form the assumption that $p = 0.5$ and so $q = 1 - p = 0.5$.

The test statistic is then

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0.540 - 0.5}{\sqrt{\frac{0.5 \cdot 0.5}{1009}}} = 2.54$$

Test Statistic for Mean μ

Sampling Distribution: Student t

Requirements: Both of the following:

- σ not known.
- Normally distributed or $n > 30$.

Test Statistic: $t = \dfrac{\bar{x} - \mu}{\dfrac{s}{\sqrt{n}}}$

Test Statistic for Mean μ

Sampling Distribution: Student t

Requirements: Both of the following:

- σ not known.
- Normally distributed or $n > 30$.

Test Statistic: $t = \dfrac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$

Test Statistic for Mean μ

Sampling Distribution: Normal (z)

Requirements: Both of the following:

- σ known.
- Normally distributed or $n > 30$.

Test Statistic: $z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$

Definition

The **critical region** (or **rejection region**) is the area corresponding to all values of the test statistic that cause us to reject the null hypothesis.

Definition

The **critical region** (or **rejection region**) is the area corresponding to all values of the test statistic that cause us to reject the null hypothesis.

Definition

In a hypothesis test, the ***P*-value** is the probability of getting a value of the test statistic that is at least as extreme as the test statistic obtained from the sample data, assuming that the null hypothesis is true.

Definition

The **critical region** (or **rejection region**) is the area corresponding to all values of the test statistic that cause us to reject the null hypothesis.

Definition

In a hypothesis test, the ***P*-value** is the probability of getting a value of the test statistic that is at least as extreme as the test statistic obtained from the sample data, assuming that the null hypothesis is true.

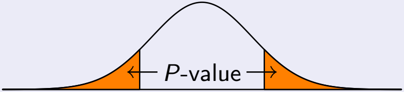
Caution

Be careful not to confuse the notation.

***P*-value**	The probability of a test statistic at least as extreme as the one obtained.
$\boldsymbol{p}$	The population proportion.
$\boldsymbol{\hat{p}}$	The sample proportion.

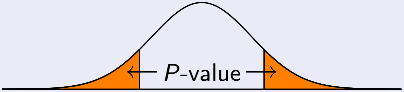
Two-tailed Test ($H_A : \neq$)

The critical region is in the two extreme regions under the curve.



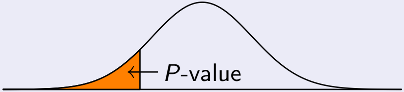
Two-tailed Test ($H_A : \neq$)

The critical region is in the two extreme regions under the curve.



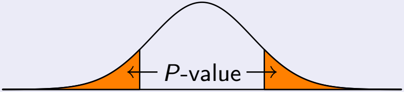
Left-tailed Test ($H_A : <$)

The critical region is in the extreme left region under the curve.



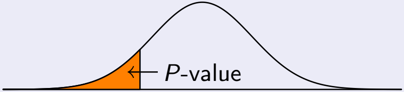
Two-tailed Test ($H_A: \neq$)

The critical region is in the two extreme regions under the curve.



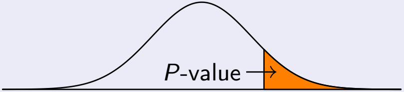
Left-tailed Test ($H_A: <$)

The critical region is in the extreme left region under the curve.



Right-tailed Test ($H_A: >$)

The critical region is in the extreme right region under the curve.



Decision Criteria

- If P-value $\leq \alpha$, reject H_0.
- If P-value $\geq \alpha$, fail to reject H_0.

Decision Criteria

- If P-value $\leq \alpha$, reject H_0.
- If P-value $\geq \alpha$, fail to reject H_0.

Example 6

In Example 1 we made a claim about the population proportion p, where we have $n = 1009$ and $x = 545$.

Decision Criteria

- If P-value $\leq \alpha$, reject H_0.
- If P-value $\geq \alpha$, fail to reject H_0.

Example 6

In Example 1 we made a claim about the population proportion p, where we have $n = 1009$ and $x = 545$.

The alternative hypothesis is $H_A: \ p > 0.5$, so this is a right-tailed test.

Decision Criteria

- If P-value $\leq \alpha$, reject H_0.
- If P-value $\geq \alpha$, fail to reject H_0.

Example 6

In Example 1 we made a claim about the population proportion p, where we have $n = 1009$ and $x = 545$.

The alternative hypothesis is $H_A: \ p > 0.5$, so this is a right-tailed test.

The test statistic is $z = 2.54$ and the area to the right of z is 0.0055.

Decision Criteria

- If P-value $\leq \alpha$, reject H_0.
- If P-value $\geq \alpha$, fail to reject H_0.

Example 6

In Example 1 we made a claim about the population proportion p, where we have $n = 1009$ and $x = 545$.

The alternative hypothesis is $H_A: \; p > 0.5$, so this is a right-tailed test.

The test statistic is $z = 2.54$ and the area to the right of z is 0.0055.

If we are working with a significance level $\alpha = 0.05$, so we reject the null hypothesis.

Decision Criteria

- If P-value $\leq \alpha$, reject H_0.
- If P-value $\geq \alpha$, fail to reject H_0.

Example 6

In Example 1 we made a claim about the population proportion p, where we have $n = 1009$ and $x = 545$.

The alternative hypothesis is $H_A: \; p > 0.5$, so this is a right-tailed test.

The test statistic is $z = 2.54$ and the area to the right of z is 0.0055.

If we are working with a significance level $\alpha = 0.05$, so we reject the null hypothesis.

Note

Technology will compute P-values for you.

Restate the Decision Using Nontechnical Terms

After you have decided to reject or not reject the null hypothesis, you need to restate the decision in terms that a layperson can understand.

Restate the Decision Using Nontechnical Terms

After you have decided to reject or not reject the null hypothesis, you need to restate the decision in terms that a layperson can understand.

Example 7

In Example 1 we restate the decision to reject the null hypothesis as:

"There is sufficient evidence to support the claim that the majority of consumers are uncomfortable with drone deliveries."

Helpful Wording

Original claim does not include equality and you reject H_0:

> "There is sufficient evidence to support the claim that . . . (claim)."

Helpful Wording

Original claim does not include equality and you reject H_0:

"There is sufficient evidence to support the claim that ... (claim)."

Original claim does not include equality and you fail to reject H_0:

"There is not sufficient evidence to support the claim that ... (claim)."

Helpful Wording

Original claim does not include equality and you reject H_0:

"There is sufficient evidence to support the claim that . . . (claim)."

Original claim does not include equality and you fail to reject H_0:

"There is not sufficient evidence to support the claim that . . . (claim)."

Original claim includes equality and you reject H_0:

"There is sufficient evidence to warrant rejection the claim that . . . (claim)."

Helpful Wording

Original claim does not include equality and you reject H_0:

"There is sufficient evidence to support the claim that ... (claim)."

Original claim does not include equality and you fail to reject H_0:

"There is not sufficient evidence to support the claim that ... (claim)."

Original claim includes equality and you reject H_0:

"There is sufficient evidence to warrant rejection the claim that ... (claim)."

Original claim includes equality and you fail to reject H_0:

"There is not sufficient evidence to warrant rejection the claim that ... (claim)."

Helpful Wording

Original claim does not include equality and you reject H_0:

"There is sufficient evidence to support the claim that . . . (claim)."

Original claim does not include equality and you fail to reject H_0:

"There is not sufficient evidence to support the claim that . . . (claim)."

Original claim includes equality and you reject H_0:

"There is sufficient evidence to warrant rejection the claim that . . . (claim)."

Original claim includes equality and you fail to reject H_0:

"There is not sufficient evidence to warrant rejection the claim that . . . (claim)."

Caution

We say "fail to reject the null hypothesis" instead of "accept the null hypothesis."

Procedure for Hypothesis Tests Flow Chart

Page 360 in your textbook contains a summary of all the steps.

Procedure for Hypothesis Tests Flow Chart

Page 360 in your textbook contains a summary of all the steps.

Note

A confidence interval estimate of a population parameter contains the likely values of that parameter.

We should therefore reject a claim that the population parameter has a value that is not included in the confidence interval.

Definition

A **type I error** is the mistake of rejecting the null hypothesis when it is actually true.

The symbol α is used to represent the probability of a type I error.

$$\alpha = P\,(\text{type I error}) = P\,(\text{rejecting } H_0 \text{ when } H_0 \text{ is true})$$

Definition

A **type I error** is the mistake of rejecting the null hypothesis when it is actually true.

The symbol α is used to represent the probability of a type I error.

$$\alpha = P\left(\text{type I error}\right) = P\left(\text{rejecting } H_0 \text{ when } H_0 \text{ is true}\right)$$

Definition

A **type II error** is the mistake of failing to reject the null hypothesis when it is actually false.

The symbol β is used to represent the probability of a type I error.

$$\beta = P\left(\text{type II error}\right) = P\left(\text{failing to reject } H_0 \text{ when } H_0 \text{ is false}\right)$$

Definition

A **type I error** is the mistake of rejecting the null hypothesis when it is actually true.

The symbol α is used to represent the probability of a type I error.

$$\alpha = P\,(\text{type I error}) = P\,(\text{rejecting } H_0 \text{ when } H_0 \text{ is true})$$

Definition

A **type II error** is the mistake of failing to reject the null hypothesis when it is actually false.

The symbol β is used to represent the probability of a type I error.

$$\beta = P\,(\text{type II error}) = P\,(\text{failing to reject } H_0 \text{ when } H_0 \text{ is false})$$

Describing Type I and Type II Errors

When wording a statement representing a type I / II error, be sure that the conclusion addresses the original claim, which may or may not be H_0.

Example 8

Consider the claim that a medical procedure designed to increase the likelihood of a baby girl is effective, which means the probability of a baby girl is $p > 0.5$.

Example 8

Consider the claim that a medical procedure designed to increase the likelihood of a baby girl is effective, which means the probability of a baby girl is $p > 0.5$.

Given the following null and alternative hypotheses

$$H_0: \ p = 0.5$$
$$H_A: \ p > 0.5$$

what is a statement that describes a type I error?

Example 8

Consider the claim that a medical procedure designed to increase the likelihood of a baby girl is effective, which means the probability of a baby girl is $p > 0.5$.

Given the following null and alternative hypotheses

$$H_0: \; p = 0.5$$
$$H_A: \; p > 0.5$$

what is a statement that describes a type I error?

In reality $p = 0.5$, but sample evidence leads us to conclude the $p > 0.5$. That is, we conclude that the medical procedure is effective when it reality it has no effect.

Example 8

Consider the claim that a medical procedure designed to increase the likelihood of a baby girl is effective, which means the probability of a baby girl is $p > 0.5$.

Given the following null and alternative hypotheses

$$H_0 : \ p = 0.5$$
$$H_A : \ p > 0.5$$

what is a statement that describes a type I error?

In reality $p = 0.5$, but sample evidence leads us to conclude the $p > 0.5$. That is, we conclude that the medical procedure is effective when it reality it has no effect.

What is a statement that describes a type II error?

Example 8

Consider the claim that a medical procedure designed to increase the likelihood of a baby girl is effective, which means the probability of a baby girl is $p > 0.5$.

Given the following null and alternative hypotheses

$$H_0 : \ p = 0.5$$
$$H_A : \ p > 0.5$$

what is a statement that describes a type I error?

In reality $p = 0.5$, but sample evidence leads us to conclude the $p > 0.5$. That is, we conclude that the medical procedure is effective when it reality it has no effect.

What is a statement that describes a type II error?

In reality $p > 0.5$, but we fail to support that conclusion. That is, we conclude that the medical procedure has no effect, when it really is effective in increasing the likelihood of a baby girl.