Hypothesis Testing

or: How I Learned to Stop Worrying and Love Inferential Statistics

Objectives

1. State the null and alternative hypothesis
2. Understand errors and interpret p-value

What is Hypothesis Testing?

Hypothesis testing is the process of testing a claim about a (population) parameter by obtaining a sample and analyzing the results against that parameter.

What is Hypothesis Testing?

Hypothesis testing is the process of testing a claim about a (population) parameter by obtaining a sample and analyzing the results against that parameter.

Null Hypothesis

The **null hypothesis** is the original claim about the parameter, denoted H_0, and usually is stated in terms of equality.

What is Hypothesis Testing?

Hypothesis testing is the process of testing a claim about a (population) parameter by obtaining a sample and analyzing the results against that parameter.

Null Hypothesis

The **null hypothesis** is the original claim about the parameter, denoted H_0, and usually is stated in terms of equality.

Alternative Hypothesis

The **alternative hypothesis**, denoted H_A, is the new claim that is made against the null hypothesis.

Types of Hypothesis Testing

- Left-tailed

Types of Hypothesis Testing

- Left-tailed
  - H_A: parameter $< H_0$

Types of Hypothesis Testing

- Left-tailed
  - H_A: parameter $< H_0$
- Right-tailed

Types of Hypothesis Testing

- Left-tailed
  - H_A: parameter $< H_0$
- Right-tailed
  - H_A: parameter $> H_0$

Types of Hypothesis Testing

- Left-tailed
  - H_A: parameter $< H_0$
- Right-tailed
  - H_A: parameter $> H_0$
- Two-tailed

Types of Hypothesis Testing

- Left-tailed
  - H_A: parameter $< H_0$
- Right-tailed
  - H_A: parameter $> H_0$
- Two-tailed
  - H_A: parameter $\neq H_0$

Example 1

Determine the null and alternative hypotheses for each.

(a) A new car's mpg is listed as 33. You want to know if the mpg is not 33.

Example 1

Determine the null and alternative hypotheses for each.

(a) A new car's mpg is listed as 33. You want to know if the mpg is not 33.

$H_0: \mu = 33$ mpg

Example 1

Determine the null and alternative hypotheses for each.

(a) A new car's mpg is listed as 33. You want to know if the mpg is not 33.

$H_0 : \mu = 33$ mpg

$H_A : \mu \neq 33$mpg

Example 1

(b) A hospital says the mean wait time for patients to see a doctor is 3.5 minutes. You want to know if the mean wait time is more than 3.5 minutes.

Example 1

(b) A hospital says the mean wait time for patients to see a doctor is 3.5 minutes. You want to know if the mean wait time is more than 3.5 minutes.

$H_0:\ \mu = 3.5$ minutes

Example 1

(b) A hospital says the mean wait time for patients to see a doctor is 3.5 minutes. You want to know if the mean wait time is more than 3.5 minutes.

$H_0: \mu = 3.5$ minutes

$H_A: \mu > 3.5$ minutes

Example 1

(c) A company claims their program will increase your grade in statistics class by 10%. You think it might not be that much.

Example 1

(c) A company claims their program will increase your grade in statistics class by 10%. You think it might not be that much.

$H_0 : \mu = 0.10$

Example 1

(c) A company claims their program will increase your grade in statistics class by 10%. You think it might not be that much.

$H_0 : \mu = 0.10$

$H_A : \mu < 0.10$

Objectives

1. State the null and alternative hypothesis
2. Understand errors and interpret p-value

What Do Our Sample Results Mean?

Our sample results will allow us to do one of two things:

What Do Our Sample Results Mean?

Our sample results will allow us to do one of two things:

- Reject the null hypothesis

What Do Our Sample Results Mean?

Our sample results will allow us to do one of two things:

- Reject the null hypothesis
- Fail to reject the null hypothesis

Rejecting the Null Hypothesis

If our sample statistics give us reason to believe that the null hypothesis may not in fact be true, then we will state our rejection of it.

Rejecting the Null Hypothesis

If our sample statistics give us reason to believe that the null hypothesis may not in fact be true, then we will state our rejection of it.

> *... we have sufficient evidence to reject [the null hypothesis].*

Rejecting the Null Hypothesis

If our sample statistics give us reason to believe that the null hypothesis may not in fact be true, then we will state our rejection of it.

> *... we have sufficient evidence to reject [the null hypothesis].*

Rejecting the null hypothesis is like a jury declaring a defendant guilty.

Rejecting the Null Hypothesis

If our sample statistics give us reason to believe that the null hypothesis may not in fact be true, then we will state our rejection of it.

> *... we have sufficient evidence to reject [the null hypothesis].*

Rejecting the null hypothesis is like a jury declaring a defendant guilty.

Note: There is still a chance that the defendant is innocent, but the evidence is strong enough to bring a guilty verdict.

Failing to Reject the Null Hypothesis

However, our sample statistics might not give us reason to believe the null hypothesis is false.

Failing to Reject the Null Hypothesis

However, our sample statistics might not give us reason to believe the null hypothesis is false.

> *... we do not have sufficient evidence to reject [the null hypothesis].*

Failing to Reject the Null Hypothesis

However, our sample statistics might not give us reason to believe the null hypothesis is false.

> *... we do not have sufficient evidence to reject [the null hypothesis].*

Failing to reject the null hypothesis is like a jury declaring a defendant not guilty.

Failing to Reject the Null Hypothesis

However, our sample statistics might not give us reason to believe the null hypothesis is false.

> *... we do not have sufficient evidence to reject [the null hypothesis].*

Failing to reject the null hypothesis is like a jury declaring a defendant not guilty.

Note: A declaration of not guilty is not the same as a declaration of innocence. There just is not sufficient evidence to declare guilt, and the defendant *could still actually be guilty*.

Errors

Just like in courts of law, sometimes innocent people are declared guilty and guilty people are declared not guilty. The same can happen in statistics.

Errors

Just like in courts of law, sometimes innocent people are declared guilty and guilty people are declared not guilty. The same can happen in statistics.

Type I Error

A **Type I error** in statistics occurs when we reject a null hypothesis that is actually true.

Errors

Just like in courts of law, sometimes innocent people are declared guilty and guilty people are declared not guilty. The same can happen in statistics.

Type I Error

A **Type I error** in statistics occurs when we reject a null hypothesis that is actually true.

The probability of making a Type I error is α, which we saw earlier with our confidence intervals in the form $1 - \alpha$.

Errors

Just like in courts of law, sometimes innocent people are declared guilty and guilty people are declared not guilty. The same can happen in statistics.

Type I Error

A **Type I error** in statistics occurs when we reject a null hypothesis that is actually true.

The probability of making a Type I error is α, which we saw earlier with our confidence intervals in the form $1 - \alpha$.

In hypothesis testing, α is called the **level of significance**.

Errors

Type II Error

A **Type II error** in statistics occurs when we fail to reject a null hypothesis that is actually false.

Errors

Type II Error

A **Type II error** in statistics occurs when we fail to reject a null hypothesis that is actually false.

The probability of making a Type II error is β.

Errors

Type II Error

A **Type II error** in statistics occurs when we fail to reject a null hypothesis that is actually false.

The probability of making a Type II error is β.

Note: As we decrease the chances of making a Type I error, we increase the chances of making a Type II error.

Errors

Type II Error

A **Type II error** in statistics occurs when we fail to reject a null hypothesis that is actually false.

The probability of making a Type II error is β.

Note: As we decrease the chances of making a Type I error, we increase the chances of making a Type II error.

If we make it harder to put an innocent person in jail, we make it tougher to return a guilty verdict. This will also have the affect of increasing the number of actual guilty defendants who are let go.

Errors

Type II Error

A **Type II error** in statistics occurs when we fail to reject a null hypothesis that is actually false.

The probability of making a Type II error is β.

Note: As we decrease the chances of making a Type I error, we increase the chances of making a Type II error.

If we make it harder to put an innocent person in jail, we make it tougher to return a guilty verdict. This will also have the affect of increasing the number of actual guilty defendants who are let go.

The **power** of a test is given as $1 - P(\beta)$

Errors Summary

$\boldsymbol{H_0}$	Reject H_0	Fail to reject H_0
H_0 True	Type I error	Correct decision
H_0 False	Correct decision	Type II error

Defendant	Declare Guilty	Declare Not Guilty
Actually Innocent	Type I error	Correct decision
Actually Guilty	Correct decision	Type II error

p-Value

Recall that **statistically significant** means that something is not likely to occur by chance.

p-Value

Recall that **statistically significant** means that something is not likely to occur by chance.

p-Value

The **p-value** is the probability of obtaining a sample as extreme as the one obtained **under the assumption that the null hypothesis is true.**

p-Value

Recall that **statistically significant** means that something is not likely to occur by chance.

p-Value

The **p-value** is the probability of obtaining a sample as extreme as the one obtained **under the assumption that the null hypothesis is true.**

If our *p*-value is less than a given acceptable value (α), then our sample was not likely to occur by chance *assuming the null hypothesis is true*,

p-Value

Recall that **statistically significant** means that something is not likely to occur by chance.

p-Value

The **p-value** is the probability of obtaining a sample as extreme as the one obtained **under the assumption that the null hypothesis is true.**

If our p-value is less than a given acceptable value (α), then our sample was not likely to occur by chance *assuming the null hypothesis is true*, so we have sufficient evidence to reject the null hypothesis.

Example 2