Gov 51: What Is A Hypothesis Test?

Matthew Blackwell

Harvard University

The lady tasting tea

Your friend asks you to grab a tea with milk for her before meeting up and she says that she prefers tea poured before the milk. You stop by a local tea shop and ask for a tea with milk. When you bring it to her, she complains that it was prepared milk-first.

The lady tasting tea

Your friend asks you to grab a tea with milk for her before meeting up and she says that she prefers tea poured before the milk. You stop by a local tea shop and ask for a tea with milk. When you bring it to her, she complains that it was prepared milk-first.

- You're skeptical that she can tell the difference, so you devise a test:

The lady tasting tea

Your friend asks you to grab a tea with milk for her before meeting up and she says that she prefers tea poured before the milk. You stop by a local tea shop and ask for a tea with milk. When you bring it to her, she complains that it was prepared milk-first.

- You're skeptical that she can tell the difference, so you devise a test:
    - Prepare 8 cups of tea, 4 milk-first, 4 tea-first

The lady tasting tea

Your friend asks you to grab a tea with milk for her before meeting up and she says that she prefers tea poured before the milk. You stop by a local tea shop and ask for a tea with milk. When you bring it to her, she complains that it was prepared milk-first.

- You're skeptical that she can tell the difference, so you devise a test:
  - Prepare 8 cups of tea, 4 milk-first, 4 tea-first
  - Present cups to friend in a **random** order

The lady tasting tea

Your friend asks you to grab a tea with milk for her before meeting up and she says that she prefers tea poured before the milk. You stop by a local tea shop and ask for a tea with milk. When you bring it to her, she complains that it was prepared milk-first.

- You're skeptical that she can tell the difference, so you devise a test:
    - Prepare 8 cups of tea, 4 milk-first, 4 tea-first
    - Present cups to friend in a **random** order
    - Ask friend to pick which 4 of the 8 were milk-first.

Assuming we know the truth

- Friend picks out all 4 milk-first cups correctly!

Assuming we know the truth

- Friend picks out all 4 milk-first cups correctly!
- Statistical thought experiment: how often would she get all 4 correct **if she were guessing randomly**?

Assuming we know the truth

- Friend picks out all 4 milk-first cups correctly!
- Statistical thought experiment: how often would she get all 4 correct **if she were guessing randomly**?
    - Only one way to choose all 4 correct cups.

Assuming we know the truth

- Friend picks out all 4 milk-first cups correctly!
- Statistical thought experiment: how often would she get all 4 correct **if she were guessing randomly**?
    - Only one way to choose all 4 correct cups.
    - But 70 ways of choosing 4 cups among 8.

Assuming we know the truth

- Friend picks out all 4 milk-first cups correctly!
- Statistical thought experiment: how often would she get all 4 correct **if she were guessing randomly**?
    - Only one way to choose all 4 correct cups.
    - But 70 ways of choosing 4 cups among 8.
    - Choosing at random $\approx$ picking each of these 70 with equal probability.

Assuming we know the truth

- Friend picks out all 4 milk-first cups correctly!
- Statistical thought experiment: how often would she get all 4 correct **if she were guessing randomly**?
    - Only one way to choose all 4 correct cups.
    - But 70 ways of choosing 4 cups among 8.
    - Choosing at random $\approx$ picking each of these 70 with equal probability.
- Chances of guessing all 4 correct is $\frac{1}{70} \approx 0.014$ or 1.4%.

Assuming we know the truth

- Friend picks out all 4 milk-first cups correctly!
- Statistical thought experiment: how often would she get all 4 correct **if she were guessing randomly**?
    - Only one way to choose all 4 correct cups.
    - But 70 ways of choosing 4 cups among 8.
    - Choosing at random $\approx$ picking each of these 70 with equal probability.
- Chances of guessing all 4 correct is $\frac{1}{70} \approx 0.014$ or 1.4%.
- $\rightsquigarrow$ the guessing hypothesis might be implausible.

Statistical hypothesis testing

- Statistical hypothesis testing is a **thought experiment**.

Statistical hypothesis testing

- Statistical hypothesis testing is a **thought experiment**.
    - Could our results just be due to random chance?

Statistical hypothesis testing

- Statistical hypothesis testing is a **thought experiment**.
    - Could our results just be due to random chance?
- What would the world look like **if we knew the truth**?

Statistical hypothesis testing

- Statistical hypothesis testing is a **thought experiment**.
    - Could our results just be due to random chance?
- What would the world look like **if we knew the truth**?
- Example 1:

Statistical hypothesis testing

- Statistical hypothesis testing is a **thought experiment**.
    - Could our results just be due to random chance?
- What would the world look like **if we knew the truth**?
- Example 1:
    - An analyst claims that 20% of Boston households are in poverty.

Statistical hypothesis testing

- Statistical hypothesis testing is a **thought experiment**.
    - Could our results just be due to random chance?
- What would the world look like **if we knew the truth**?
- Example 1:
    - An analyst claims that 20% of Boston households are in poverty.
    - You take a sample of 900 households and find that 23% of the sample is under the poverty line.

Statistical hypothesis testing

- Statistical hypothesis testing is a **thought experiment**.
    - Could our results just be due to random chance?
- What would the world look like **if we knew the truth**?
- Example 1:
    - An analyst claims that 20% of Boston households are in poverty.
    - You take a sample of 900 households and find that 23% of the sample is under the poverty line.
    - Should you conclude that the analyst is wrong?

Statistical hypothesis testing

- Statistical hypothesis testing is a **thought experiment**.
    - Could our results just be due to random chance?
- What would the world look like **if we knew the truth**?
- Example 1:
    - An analyst claims that 20% of Boston households are in poverty.
    - You take a sample of 900 households and find that 23% of the sample is under the poverty line.
    - Should you conclude that the analyst is wrong?
- Example 2:

Statistical hypothesis testing

- Statistical hypothesis testing is a **thought experiment**.
    - Could our results just be due to random chance?
- What would the world look like **if we knew the truth**?
- Example 1:
    - An analyst claims that 20% of Boston households are in poverty.
    - You take a sample of 900 households and find that 23% of the sample is under the poverty line.
    - Should you conclude that the analyst is wrong?
- Example 2:
    - Trump won 47.5% of the vote in the 2020 election.

Statistical hypothesis testing

- Statistical hypothesis testing is a **thought experiment**.
  - Could our results just be due to random chance?
- What would the world look like **if we knew the truth**?
- Example 1:
  - An analyst claims that 20% of Boston households are in poverty.
  - You take a sample of 900 households and find that 23% of the sample is under the poverty line.
  - Should you conclude that the analyst is wrong?
- Example 2:
  - Trump won 47.5% of the vote in the 2020 election.
  - Last YouGov poll of 1,363 likely voters said 44% planned to vote for Trump.

Statistical hypothesis testing

- Statistical hypothesis testing is a **thought experiment**.
  - Could our results just be due to random chance?
- What would the world look like **if we knew the truth**?
- Example 1:
  - An analyst claims that 20% of Boston households are in poverty.
  - You take a sample of 900 households and find that 23% of the sample is under the poverty line.
  - Should you conclude that the analyst is wrong?
- Example 2:
  - Trump won 47.5% of the vote in the 2020 election.
  - Last YouGov poll of 1,363 likely voters said 44% planned to vote for Trump.
  - Could the difference between the poll and the outcome be just due to random chance?

Null and alternative hypothesis

- **Null hypothesis**: Some statement about the population parameters.

Null and alternative hypothesis

- **Null hypothesis**: Some statement about the population parameters.
  - "Devil's advocate" position $\rightsquigarrow$ assumes what you seek to prove wrong.

Null and alternative hypothesis

- **Null hypothesis**: Some statement about the population parameters.
    - “Devil’s advocate” position $\rightsquigarrow$ assumes what you seek to prove wrong.
    - Usually that an observed difference is due to chance.

Null and alternative hypothesis

- **Null hypothesis**: Some statement about the population parameters.
  - “Devil’s advocate” position $\rightsquigarrow$ assumes what you seek to prove wrong.
  - Usually that an observed difference is due to chance.
  - Ex: poll drawn from the same population as all voters.

Null and alternative hypothesis

- **Null hypothesis**: Some statement about the population parameters.
    - "Devil's advocate" position $\rightsquigarrow$ assumes what you seek to prove wrong.
    - Usually that an observed difference is due to chance.
    - Ex: poll drawn from the same population as all voters.
    - Denoted H_0

Null and alternative hypothesis

- **Null hypothesis**: Some statement about the population parameters.
  - "Devil's advocate" position $\rightsquigarrow$ assumes what you seek to prove wrong.
  - Usually that an observed difference is due to chance.
  - Ex: poll drawn from the same population as all voters.
  - Denoted H_0
- **Alternative hypothesis**: The statement we hope or suspect is true instead of H_0.

Null and alternative hypothesis

- **Null hypothesis**: Some statement about the population parameters.
  - “Devil’s advocate” position $\rightsquigarrow$ assumes what you seek to prove wrong.
  - Usually that an observed difference is due to chance.
  - Ex: poll drawn from the same population as all voters.
  - Denoted H_0
- **Alternative hypothesis**: The statement we hope or suspect is true instead of H_0.
  - It is the opposite of the null hypothesis.

Null and alternative hypothesis

- **Null hypothesis**: Some statement about the population parameters.
    - "Devil's advocate" position $\rightsquigarrow$ assumes what you seek to prove wrong.
    - Usually that an observed difference is due to chance.
    - Ex: poll drawn from the same population as all voters.
    - Denoted H_0
- **Alternative hypothesis**: The statement we hope or suspect is true instead of H_0.
    - It is the opposite of the null hypothesis.
    - An observed difference is real, not just due to chance.

Null and alternative hypothesis

- **Null hypothesis**: Some statement about the population parameters.
  - "Devil's advocate" position ⇝ assumes what you seek to prove wrong.
  - Usually that an observed difference is due to chance.
  - Ex: poll drawn from the same population as all voters.
  - Denoted H_0
- **Alternative hypothesis**: The statement we hope or suspect is true instead of H_0.
  - It is the opposite of the null hypothesis.
  - An observed difference is real, not just due to chance.
  - Ex: polling for Trump is systematically wrong.

Null and alternative hypothesis

- **Null hypothesis**: Some statement about the population parameters.
  - "Devil's advocate" position $\rightsquigarrow$ assumes what you seek to prove wrong.
  - Usually that an observed difference is due to chance.
  - Ex: poll drawn from the same population as all voters.
  - Denoted H_0
- **Alternative hypothesis**: The statement we hope or suspect is true instead of H_0.
  - It is the opposite of the null hypothesis.
  - An observed difference is real, not just due to chance.
  - Ex: polling for Trump is systematically wrong.
  - Denoted H_1 or H_a

Null and alternative hypothesis

- **Null hypothesis**: Some statement about the population parameters.
    - "Devil's advocate" position $\rightsquigarrow$ assumes what you seek to prove wrong.
    - Usually that an observed difference is due to chance.
    - Ex: poll drawn from the same population as all voters.
    - Denoted H_0
- **Alternative hypothesis**: The statement we hope or suspect is true instead of H_0.
    - It is the opposite of the null hypothesis.
    - An observed difference is real, not just due to chance.
    - Ex: polling for Trump is systematically wrong.
    - Denoted H_1 or H_a
- **Probabilistic** proof by contradiction: try to "disprove" the null.

Hypothesis testing example

- Are we polling the same population as the actual voters?

Hypothesis testing example

- Are we polling the same population as the actual voters?
    - If so, how likely are we to see polling error this big by chance?

Hypothesis testing example

- Are we polling the same population as the actual voters?
    - If so, how likely are we to see polling error this big by chance?
- What is the parameter we want to learn about?

Hypothesis testing example

- Are we polling the same population as the actual voters?
    - If so, how likely are we to see polling error this big by chance?
- What is the parameter we want to learn about?
    - True population mean of the surveys, p.

Hypothesis testing example

- Are we polling the same population as the actual voters?
    - If so, how likely are we to see polling error this big by chance?
- What is the parameter we want to learn about?
    - True population mean of the surveys, p.
    - Null hypothesis: $H_0 : p = 0.475$ (surveys drawing from same population)

Hypothesis testing example

- Are we polling the same population as the actual voters?
  - If so, how likely are we to see polling error this big by chance?
- What is the parameter we want to learn about?
  - True population mean of the surveys, p.
  - Null hypothesis: $H_0 : p = 0.475$ (surveys drawing from same population)
  - Alternative hypothesis: $H_1 : p \neq 0.475$

Hypothesis testing example

- Are we polling the same population as the actual voters?
    - If so, how likely are we to see polling error this big by chance?
- What is the parameter we want to learn about?
    - True population mean of the surveys, p.
    - Null hypothesis: $H_0 : p = 0.475$ (surveys drawing from same population)
    - Alternative hypothesis: $H_1 : p \neq 0.475$
- Data: poll has $\overline{X} = 0.44$ with $n = 1363$.

Statistical thought experiment

- If the null were true, what should the distribution of the data be?

Statistical thought experiment

- If the null were true, what should the distribution of the data be?
    - X_i is 1 if respondents i will vote for Trump.

Statistical thought experiment

- If the null were true, what should the distribution of the data be?
  - X_i is 1 if respondents i will vote for Trump.
  - Under null, X_i is Bernoulli with $p = 0.475$.

Statistical thought experiment

- If the null were true, what should the distribution of the data be?
  - X_i is 1 if respondents i will vote for Trump.
  - Under null, X_i is Bernoulli with $p = 0.475$.
  - $\sum_{i=1}^{N} X_i$ is the number in sample that will vote for Trump.

Statistical thought experiment

- If the null were true, what should the distribution of the data be?
    - X_i is 1 if respondents i will vote for Trump.
    - Under null, X_i is Bernoulli with $p = 0.475$.
    - $\sum_{i=1}^{N} X_i$ is the number in sample that will vote for Trump.
    - This sum will be Binomial with $n = 1363$ and $p = 0.475$

Statistical thought experiment

- If the null were true, what should the distribution of the data be?
  - X_i is 1 if respondents i will vote for Trump.
  - Under null, X_i is Bernoulli with $p = 0.475$.
  - $\sum_{i=1}^{N} X_i$ is the number in sample that will vote for Trump.
  - This sum will be Binomial with $n = 1363$ and $p = 0.475$
- We can simulate draws from this distribution!

Statistical thought experiment

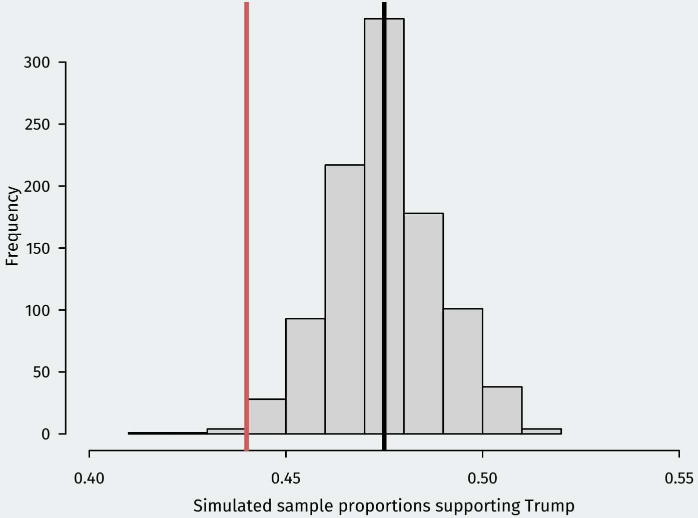
- If the null were true, what should the distribution of the data be?
    - X_i is 1 if respondents i will vote for Trump.
    - Under null, X_i is Bernoulli with $p = 0.475$.
    - $\sum_{i=1}^{N} X_i$ is the number in sample that will vote for Trump.
    - This sum will be Binomial with $n = 1363$ and $p = 0.475$
- We can simulate draws from this distribution!
- Compare the distribution of proportions under the null to the observed proportion.

Statistical thought experiment

- If the null were true, what should the distribution of the data be?
  - X_i is 1 if respondents i will vote for Trump.
  - Under null, X_i is Bernoulli with $p = 0.475$.
  - $\sum_{i=1}^{N} X_i$ is the number in sample that will vote for Trump.
  - This sum will be Binomial with $n = 1363$ and $p = 0.475$
- We can simulate draws from this distribution!
- Compare the distribution of proportions under the null to the observed proportion.

```
trump_voters <- rbinom(n = 1000, size = 1363, prob = 0.475)
trump_shares <- trump_voters / 1363
hist(trump_shares, xlim = c(0.4, 0.55),
     xlab = "Simulated sample proportions supporting Trump",
     main = "")
abline(v = 0.44, col = "indianred", lwd = 3)
abline(v = 0.475, lwd = 3)
```

Simulations of the null distribution



p-value

p-value

The **p-value** is the probability of observing data as or more extreme as our data under the null.

- If the null is true, how often would we expect polling errors this big?

p-value

p-value

The **p-value** is the probability of observing data as or more extreme as our data under the null.

- If the null is true, how often would we expect polling errors this big?
  - Smaller p-value $\rightsquigarrow$ stronger evidence against the null

p-value

p-value

The **p-value** is the probability of observing data as or more extreme as our data under the null.

- If the null is true, how often would we expect polling errors this big?
  - Smaller p-value $\rightsquigarrow$ stronger evidence against the null
  - **NOT** the probability that the null is true!

p-value

p-value

The **p-value** is the probability of observing data as or more extreme as our data under the null.

- If the null is true, how often would we expect polling errors this big?
  - Smaller p-value $\rightsquigarrow$ stronger evidence against the null
  - **NOT** the probability that the null is true!
- p-values are usually **two-sided**:

p-value

p-value

The **p-value** is the probability of observing data as or more extreme as our data under the null.

- If the null is true, how often would we expect polling errors this big?
  - Smaller p-value $\rightsquigarrow$ stronger evidence against the null
  - **NOT** the probability that the null is true!
- p-values are usually **two-sided**:
  - Observed error of 0.44 - 0.475 = -0.035 under the null.

p-value

p-value

The **p-value** is the probability of observing data as or more extreme as our data under the null.

- If the null is true, how often would we expect polling errors this big?
  - Smaller p-value $\rightsquigarrow$ stronger evidence against the null
  - **NOT** the probability that the null is true!
- p-values are usually **two-sided**:
  - Observed error of 0.44 - 0.475 = -0.035 under the null.
  - p-value is probability of sample proportions being less than 0.44 **plus**

p-value

p-value

The **p-value** is the probability of observing data as or more extreme as our data under the null.

- If the null is true, how often would we expect polling errors this big?
  - Smaller p-value $\rightsquigarrow$ stronger evidence against the null
  - **NOT** the probability that the null is true!
- p-values are usually **two-sided**:
  - Observed error of 0.44 - 0.475 = -0.035 under the null.
  - p-value is probability of sample proportions being less than 0.44 **plus**
  - Probability of sample proportions being greater than 0.475 + 0.035 = 0.51.

p-value

p-value

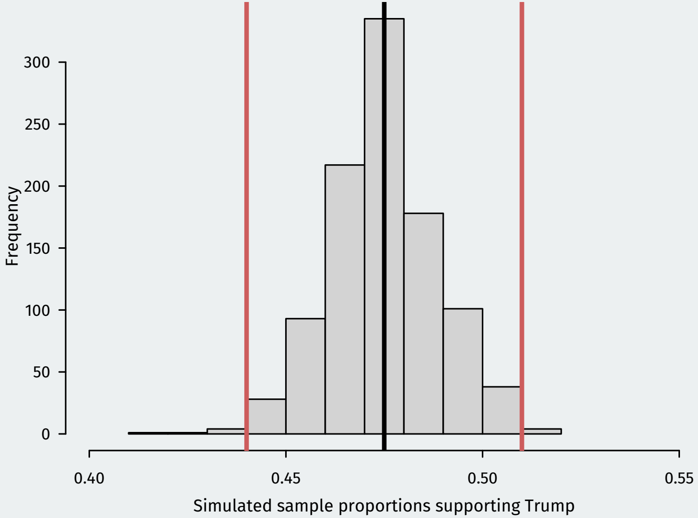
The **p-value** is the probability of observing data as or more extreme as our data under the null.

- If the null is true, how often would we expect polling errors this big?
  - Smaller p-value $\rightsquigarrow$ stronger evidence against the null
  - **NOT** the probability that the null is true!
- p-values are usually **two-sided**:
  - Observed error of 0.44 - 0.475 = -0.035 under the null.
  - p-value is probability of sample proportions being less than 0.44 **plus**
  - Probability of sample proportions being greater than 0.475 + 0.035 = 0.51.

```
mean(trump_shares < 0.44) + mean(trump_shares > 0.51)
```

```
## [1] 0.01
```

Two-sided p-value



One-sided tests

- Sometimes our hypothesis is directional.

One-sided tests

- Sometimes our hypothesis is directional.
  - We only consider evidence against the null from one direction.

One-sided tests

- Sometimes our hypothesis is directional.
  - We only consider evidence against the null from one direction.
- Null: our polls are from the same population as actual voters

One-sided tests

- Sometimes our hypothesis is directional.
    - We only consider evidence against the null from one direction.
- Null: our polls are from the same population as actual voters
    - $H_0 : p = 0.475$

One-sided tests

- Sometimes our hypothesis is directional.
    - We only consider evidence against the null from one direction.
- Null: our polls are from the same population as actual voters
    - $H_0 : p = 0.475$
- **One-sided alternative**: polls underestimate Trump support.

One-sided tests

- Sometimes our hypothesis is directional.
    - We only consider evidence against the null from one direction.
- Null: our polls are from the same population as actual voters
    - $H_0 : p = 0.475$
- **One-sided alternative**: polls underestimate Trump support.
    - $H_1 : p < 0.475$

One-sided tests

- Sometimes our hypothesis is directional.
  - We only consider evidence against the null from one direction.
- Null: our polls are from the same population as actual voters
  - $H_0 : p = 0.475$
- **One-sided alternative**: polls underestimate Trump support.
  - $H_1 : p < 0.475$
- Makes the p-value one-sided:

One-sided tests

- Sometimes our hypothesis is directional.
  - We only consider evidence against the null from one direction.
- Null: our polls are from the same population as actual voters
  - $H_0 : p = 0.475$
- **One-sided alternative**: polls underestimate Trump support.
  - $H_1 : p < 0.475$
- Makes the p-value one-sided:
  - What's the probability of a random sample underestimating Trump support by as much as we see in the sample?

One-sided tests

- Sometimes our hypothesis is directional.
    - We only consider evidence against the null from one direction.
- Null: our polls are from the same population as actual voters
    - $H_0 : p = 0.475$
- **One-sided alternative**: polls underestimate Trump support.
    - $H_1 : p < 0.475$
- Makes the p-value one-sided:
    - What's the probability of a random sample underestimating Trump support by as much as we see in the sample?
    - Always smaller than a two-sided p-value.

One-sided tests

- Sometimes our hypothesis is directional.
  - We only consider evidence against the null from one direction.
- Null: our polls are from the same population as actual voters
  - $H_0 : p = 0.475$
- **One-sided alternative**: polls underestimate Trump support.
  - $H_1 : p < 0.475$
- Makes the p-value one-sided:
  - What's the probability of a random sample underestimating Trump support by as much as we see in the sample?
  - Always smaller than a two-sided p-value.

```
mean(trump_shares < 0.44)
```

```
## [1] 0.006
```

Rejecting the null

- Tests usually end with a decision to reject the null or not.

Rejecting the null

- Tests usually end with a decision to reject the null or not.
- Choose a threshold below which you'll reject the null.

Rejecting the null

- Tests usually end with a decision to reject the null or not.
- Choose a threshold below which you'll reject the null.
    - **Test level** α: the threshold for a test.

Rejecting the null

- Tests usually end with a decision to reject the null or not.
- Choose a threshold below which you'll reject the null.
    - **Test level** α: the threshold for a test.
    - Decision rule: "reject the null if the p-value is below α"

Rejecting the null

- Tests usually end with a decision to reject the null or not.
- Choose a threshold below which you'll reject the null.
    - **Test level** α: the threshold for a test.
    - Decision rule: "reject the null if the p-value is below α"
    - Otherwise "fail to reject" or "retain", not "accept the null"

Rejecting the null

- Tests usually end with a decision to reject the null or not.
- Choose a threshold below which you'll reject the null.
    - **Test level** α: the threshold for a test.
    - Decision rule: "reject the null if the p-value is below α"
    - Otherwise "fail to reject" or "retain", not "accept the null"
- Common (arbitrary) thresholds:

Rejecting the null

- Tests usually end with a decision to reject the null or not.
- Choose a threshold below which you'll reject the null.
  - **Test level** α: the threshold for a test.
  - Decision rule: "reject the null if the p-value is below α"
  - Otherwise "fail to reject" or "retain", not "accept the null"
- Common (arbitrary) thresholds:
  - $p \geq 0.1$ "not statistically significant"

Rejecting the null

- Tests usually end with a decision to reject the null or not.
- Choose a threshold below which you'll reject the null.
    - **Test level** α: the threshold for a test.
    - Decision rule: "reject the null if the p-value is below α"
    - Otherwise "fail to reject" or "retain", not "accept the null"
- Common (arbitrary) thresholds:
    - $p \geq 0.1$ "not statistically significant"
    - $p < 0.05$ "statistically significant"

Rejecting the null

- Tests usually end with a decision to reject the null or not.
- Choose a threshold below which you'll reject the null.
    - **Test level** α: the threshold for a test.
    - Decision rule: "reject the null if the p-value is below α"
    - Otherwise "fail to reject" or "retain", not "accept the null"
- Common (arbitrary) thresholds:
    - $p \geq 0.1$ "not statistically significant"
    - $p < 0.05$ "statistically significant"
    - $p < 0.01$ "highly significant"

Testing errors

- A p-value of 0.05 says that data this extreme would only happen in 5% of repeated samples if the null were true.

Testing errors

- A p-value of 0.05 says that data this extreme would only happen in 5% of repeated samples if the null were true.
    - $\rightsquigarrow$ 5% of the time we'll reject the null when it is actually true.

Testing errors

- A p-value of 0.05 says that data this extreme would only happen in 5% of repeated samples if the null were true.
    - $\rightsquigarrow$ 5% of the time we'll reject the null when it is actually true.
- Test errors:

Testing errors

- A p-value of 0.05 says that data this extreme would only happen in 5% of repeated samples if the null were true.
  - $\rightsquigarrow$ 5% of the time we'll reject the null when it is actually true.
- Test errors:

	H_0 True	H_0 False
Retain H_0	Awesome!	Type II error
Reject H_0	Type I error	Good stuff!

Testing errors

- A p-value of 0.05 says that data this extreme would only happen in 5% of repeated samples if the null were true.
  - $\rightsquigarrow$ 5% of the time we'll reject the null when it is actually true.
- Test errors:

	H_0 True	H_0 False
Retain H_0	Awesome!	Type II error
Reject H_0	Type I error	Good stuff!

- Type I error because it's the worst

Testing errors

- A p-value of 0.05 says that data this extreme would only happen in 5% of repeated samples if the null were true.
  - $\rightsquigarrow$ 5% of the time we'll reject the null when it is actually true.
- Test errors:

	H_0 True	H_0 False
Retain H_0	Awesome!	Type II error
Reject H_0	Type I error	Good stuff!

- Type I error because it's the worst
  - "Convicting" an innocent null hypothesis

Testing errors

- A p-value of 0.05 says that data this extreme would only happen in 5% of repeated samples if the null were true.
    - $\rightsquigarrow$ 5% of the time we'll reject the null when it is actually true.
- Test errors:

	H_0 True	H_0 False
Retain H_0	Awesome!	Type II error
Reject H_0	Type I error	Good stuff!

- Type I error because it's the worst
    - "Convicting" an innocent null hypothesis
- Type II error less serious

Testing errors

- A p-value of 0.05 says that data this extreme would only happen in 5% of repeated samples if the null were true.
    - $\rightsquigarrow$ 5% of the time we'll reject the null when it is actually true.
- Test errors:

	H_0 True	H_0 False
Retain H_0	Awesome!	Type II error
Reject H_0	Type I error	Good stuff!

- Type I error because it's the worst
    - "Convicting" an innocent null hypothesis
- Type II error less serious
    - Missed out on an awesome finding