Lecture 9: Hypothesis Testing 2

Jacob M. Montgomery

Quantitative Political Methodology

Lecture 9: Hypothesis Testing 2

Roadmap

Last class:

- What is a hypothesis test?
- The five steps of hypothesis testing.

Roadmap

Last class:

- What is a hypothesis test?
- The five steps of hypothesis testing.

This class:

- Hypothesis tests with small samples
- Types of errors
- Discussion of one-sided/two-sided tests
- Relationship between CI and NHPT

Small sample significance testing for quantitative variables

Step 1: Assumptions

- Random sampling
- Quantitative data

Small sample significance testing for quantitative variables

Step 1: Assumptions

- Random sampling
- Quantitative data
- **Population is distributed normally**

Step 2: State hypotheses

- $H_0 : \mu = \mu_0$ (e.g., $\mu = 12$)
- $H_a : \mu \neq \mu_0$
- This is a "two-sided test," but it may be a "one-sided."

Step 3: Calculate a test statistic

- $t^* = \dfrac{\bar{Y} - \mu_0}{\sigma_{\bar{Y}}}, df = (n-1)$
- Just as before, this comes from the sampling distribution of $\bar{Y}$

Step 3: Calculate a test statistic

- $t^* = \dfrac{\bar{Y} - \mu_0}{\sigma_{\bar{Y}}}, df = (n - 1)$
- Just as before, this comes from the sampling distribution of $\bar{Y}$

Step 4: P-Value

- $p = Pr\Big(|t| \geq |\frac{\bar{Y}-\mu_0}{\sigma_{\bar{Y}}}|\Big) = 2 \times Pr(t \geq |\frac{\bar{Y}-\mu_0}{\sigma_{\bar{Y}}}|)$

Step 3: Calculate a test statistic

- $t^* = \dfrac{\bar{Y} - \mu_0}{\sigma_{\bar{Y}}}, df = (n-1)$
- Just as before, this comes from the sampling distribution of $\bar{Y}$

Step 4: P-Value

- $p = Pr\Big(|t| \geq |\frac{\bar{Y}-\mu_0}{\sigma_{\bar{Y}}}|\Big) = 2 \times Pr(t \geq |\frac{\bar{Y}-\mu_0}{\sigma_{\bar{Y}}}|)$
- Make sure you are using the right degrees of freedom.

Step 3: Calculate a test statistic

- $t^* = \dfrac{\bar{Y} - \mu_0}{\sigma_{\bar{Y}}}, df = (n-1)$
- Just as before, this comes from the sampling distribution of $\bar{Y}$

Step 4: P-Value

- $p = Pr\Big(|t| \geq |\frac{\bar{Y}-\mu_0}{\sigma_{\bar{Y}}}|\Big) = 2 \times Pr(t \geq |\frac{\bar{Y}-\mu_0}{\sigma_{\bar{Y}}}|)$
- Make sure you are using the right degrees of freedom.
- We use both tails, because we want to find the probability of error in both directions.
- `2*pt(abs(`t^*`), df = n-1, lower.tail=F)`

Step 5: Draw a conclusion

- If $p \leq \alpha$ we conclude that the evidence supports H_a
- But always report the p-value

Example: State spending on education

Assume that the theory is that states are spending less than 5% of their income on education. The data indicate that:

- $\bar{Y} = 4.7$, $S = 0.0922$
- $t^* = \frac{4.7-5}{0.09/\sqrt{50}} = -2.279$, $df = 49$
- P-value=`2*pt(2.279, df=49, lower.tail = F)` $= 0.027$

Type 1 and Type II Error

		Jury decision	
		Guilty	Innocent
Truth	Guilty	Correct	Type II
	Innocent	Type 1	Correct

Imagine we put the null hypothesis on trial:

		Researcher Conclusion	
		Reject Null	Don't reject
Truth	Null is False	Correct	Type II
	Null is True	Type 1	Correct

Imagine we put the null hypothesis on trial:

		Researcher Conclusion	
		Reject Null	Don't reject
Truth	Null is False	Correct	Type II
	Null is True	Type 1	Correct

- Type I error is when we reject a null hypothesis when the null it is actually true.

Imagine we put the null hypothesis on trial:

		Researcher Conclusion	
		Reject Null	Don't reject
Truth	Null is False	Correct	Type II
	Null is True	Type 1	Correct

- Type I error is when we reject a null hypothesis when the null it is actually true.
- Type II error is when we fail to reject a null hypothesis when the null is actually false.

Imagine we put the null hypothesis on trial:

		Researcher Conclusion	
		Reject Null	Don't reject
Truth	Null is False	Correct	Type II
	Null is True	Type 1	Correct

- Type I error is when we reject a null hypothesis when the null it is actually true.
- Type II error is when we fail to reject a null hypothesis when the null is actually false.
- We tend to prioritize reducing Type I error, although there are trade-offs.

Notes on Type I and Type II error

How to calculate each?

- Type I:
  - $Pr(Reject\ H_0 | H_0\, True)$

Notes on Type I and Type II error

How to calculate each?

- Type I:
  - $Pr(Reject\ H_0|H_0 True) = \alpha$
  - This is our primary focus

Notes on Type I and Type II error

How to calculate each?

- Type I:
  - $Pr(Reject\ H_0|H_0\,True) = \alpha$
  - This is our primary focus

- Type II (hard):
  - $Pr(Fail\ to\ reject|H_0\ False)$
  - H_0 can be false for many values

Notes on Type I and Type II error

How to calculate each?

- Type I:
  - $Pr(Reject\ H_0 | H_0\, True) = \alpha$
  - This is our primary focus
- Type II (hard):
  - $Pr(Fail\ to\ reject | H_0\ False)$
  - H_0 can be false for many values
  - "Power" of a test is $1 - Pr(Type\ II\ error)$

Notes on Type I and Type II error

How to calculate each?

- Type I:
  - $Pr(Reject\ H_0|H_0\,True) = \alpha$
  - This is our primary focus

- Type II (hard):
  - $Pr(Fail\ to\ reject|H_0\ False)$
  - H_0 can be false for many values
  - "Power" of a test is $1 - Pr(Type\ II\ error)$
  - Leave this for a more advance class
  - There is a trade-off between Type I and Type II error

Reprise: Large sample significance testing for proportions

Step 1: Assumptions

- Random sampling
- Qualitative data
- $n \geq \dfrac{10}{min(\pi_0, 1 - \pi_0)}$
- *min*() means the minimum of the two numbers
- If our n is bigger than this, we can use the calculations below.
- The 10 is sort of arbitrary (it is a good rule of thumb).

Step 2: State hypotheses

- $H_0 : \pi = \pi_0$ (e.g., $\pi_0 = 0.5$)
- $H_a : \pi \neq \pi_0$
- This is a "two-sided test," but it may be a "one-sided."

Step 3: Calculate a test statistic

- $Z^* = \dfrac{\hat{\pi} - \pi_0}{\sigma_{\pi_0}}$,

Step 3: Calculate a test statistic

- $Z^* = \dfrac{\hat{\pi} - \pi_0}{\sigma_{\pi_0}}$, $\sigma_{\pi_0} = \sqrt{\frac{\pi_0(1-\pi_0)}{n}}$
- Just as before, this comes from the sampling distribution $\hat{\pi}$ that would exist if H_0 were true.
- Note we are assuming that $\pi = \pi_0$ to calculate the SE.

Step 4: P-Value

- $p = Pr(Z \geq |\frac{\hat{\pi}-\pi_0}{\sigma_{\pi_0}}|) + Pr(Z \leq -|\frac{\hat{\pi}-\pi_0}{\sigma_{\pi_0}}|) = Pr(|Z| \geq |\frac{\hat{\pi}-\pi_0}{\sigma_{\pi_0}}|)$
- $= 2 \times Pr(Z \geq |\frac{\hat{\pi}-\pi_0}{\sigma_{\pi_0}}|)$
- We use both tails.

Step 5: Draw a conclusion

- If $p \leq \alpha$ we conclude that the evidence supports H_a
- You should **still** report the p-value.

Example: Women in Congress

Women are (roughly) 50% of the U.S. population. Are they represented (descriptively) in Congress? Using data from the 107th House we see that there are 60 women in Congress.

Example: Women in Congress

Women are (roughly) 50% of the U.S. population. Are they represented (descriptively) in Congress? Using data from the 107th House we see that there are 60 women in Congress.

- $H_0 : \pi = 0.5$, $H_a : \pi \neq 0.5$
- $n = 435, \hat{\pi} = \frac{60}{435}$

Example: Women in Congress

Women are (roughly) 50% of the U.S. population. Are they represented (descriptively) in Congress? Using data from the 107th House we see that there are 60 women in Congress.

- $H_0 : \pi = 0.5$, $H_a : \pi \neq 0.5$
- $n = 435, \hat{\pi} = \frac{60}{435}$
- Are assumptions met?

Example: Women in Congress

Women are (roughly) 50% of the U.S. population. Are they represented (descriptively) in Congress? Using data from the 107th House we see that there are 60 women in Congress.

- $H_0 : \pi = 0.5,\ H_a : \pi \neq 0.5$
- $n = 435, \hat{\pi} = \frac{60}{435}$
- Are assumptions met?
- $n \geq \dfrac{10}{min(.5, 1 - .5)} = 20$

Example: Women in Congress

Women are (roughly) 50% of the U.S. population. Are they represented (descriptively) in Congress? Using data from the 107th House we see that there are 60 women in Congress.

- $H_0 : \pi = 0.5,\ H_a : \pi \neq 0.5$
- $n = 435, \hat{\pi} = \frac{60}{435}$
- Are assumptions met?
- $n \geq \dfrac{10}{min(.5, 1 - .5)} = 20$
- Since our n is bigger, we can proceed
- $Z = \frac{\hat{\pi} - 0.5}{\sigma_{\pi_0}} = \frac{0.1379 - 0.5}{\sqrt{\frac{.5(1-.5)}{435}}} = \frac{-0.3621}{0.02397} = -15.11$

Example: Women in Congress

Women are (roughly) 50% of the U.S. population. Are they represented (descriptively) in Congress? Using data from the 107th House we see that there are 60 women in Congress.

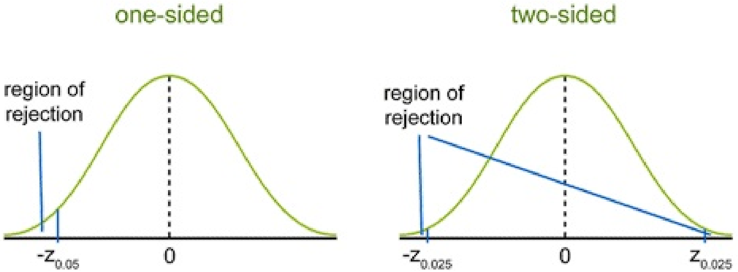
- $H_0 : \pi = 0.5$, $H_a : \pi \neq 0.5$
- $n = 435, \hat{\pi} = \frac{60}{435}$
- Are assumptions met?
- $n \geq \dfrac{10}{min(.5, 1 - .5)} = 20$
- Since our n is bigger, we can proceed
- $Z = \frac{\hat{\pi} - 0.5}{\sigma_{\pi_0}} = \frac{0.1379 - 0.5}{\sqrt{\frac{.5(1-.5)}{435}}} = \frac{-0.3621}{0.02397} = -15.11$
- `2*pnorm(15.11, lower.tail=F)` $\approx 0.0001 \Rightarrow$ reject null.

Example: Women in Congress

Women are (roughly) 50% of the U.S. population. Are they represented (descriptively) in Congress? Using data from the 107th House we see that there are 60 women in Congress.

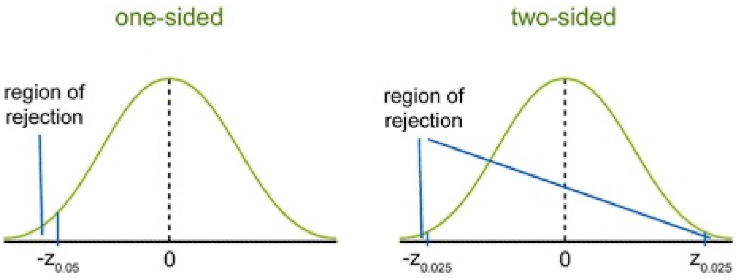
- $H_0 : \pi = 0.5,\ H_a : \pi \neq 0.5$
- $n = 435, \hat{\pi} = \frac{60}{435}$
- Are assumptions met?
- $n \geq \dfrac{10}{min(.5, 1 - .5)} = 20$
- Since our n is bigger, we can proceed
- $Z = \frac{\hat{\pi} - 0.5}{\sigma_{\pi_0}} = \frac{0.1379 - 0.5}{\sqrt{\frac{.5(1-.5)}{435}}} = \frac{-0.3621}{0.02397} = -15.11$
- `2*pnorm(15.11, lower.tail=F)` $\approx 0.0001 \Rightarrow$ reject null.
- Notice that I used positive 15.11 and not -15.11. I used the absolute value.

Critical values



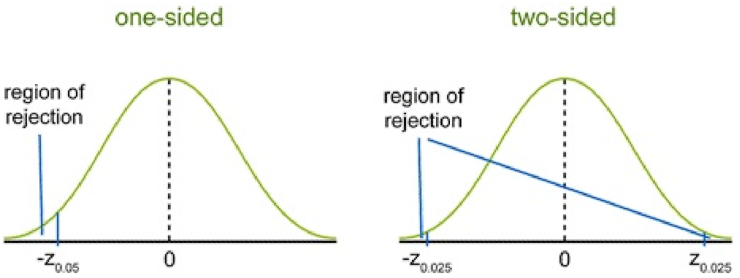
- You can think of hypothesis testing as, “Is the t-statistic so extreme that it is in the rejection region?”

Critical values



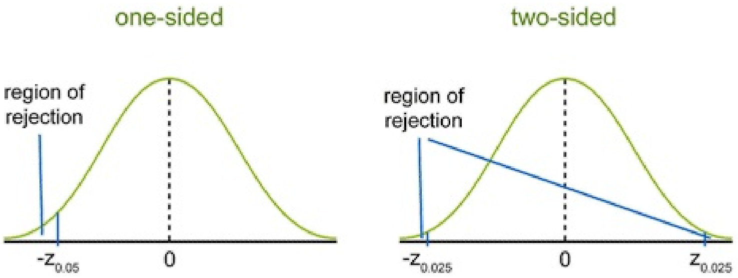
- You can think of hypothesis testing as, "Is the t-statistic so extreme that it is in the rejection region?"
- If $|Z^*|$ is bigger than some "critical value," Z_α, then p-value will be smaller than α.

Critical values



- You can think of hypothesis testing as, "Is the t-statistic so extreme that it is in the rejection region?"
- If $|Z^*|$ is bigger than some "critical value," Z_α, then p-value will be smaller than α.
- Note that the p-value calculations will be different if you are only calculating the area under one-tail.

Critical values



- You can think of hypothesis testing as, "Is the t-statistic so extreme that it is in the rejection region?"
- If $|Z^*|$ is bigger than some "critical value," Z_α, then p-value will be smaller than α.
- Note that the p-value calculations will be different if you are only calculating the area under one-tail.
- When you calculate the p-value for one-sided tests, you need to make sure you are calculating it for the correct tail.

One sided vs. two sided

Central message #1: If you are doing a one-sided test *and the estimator IS NOT in the region of the null hypothesis*, don't multiply by 2 when calculating the p-value.

- Go back to our example of women in the U.S. House.
- If $H_0 : \pi > 0.5, H_a : \pi < 0.5$
- The test statistic IS in the region covered by the alternative hypothesis and IS NOT in the region covered by the null.

One sided vs. two sided

Central message #1: If you are doing a one-sided test *and the estimator IS NOT in the region of the null hypothesis*, don't multiply by 2 when calculating the p-value.

- Go back to our example of women in the U.S. House.
- If $H_0 : \pi > 0.5, H_a : \pi < 0.5$
- The test statistic IS in the region covered by the alternative hypothesis and IS NOT in the region covered by the null.
- This might still be written as $H_0 : \pi = 0$, but it should still be $H_a : \pi < 0.5$.
- $Z = \frac{\hat{\pi} - 0.5}{\sigma_{\pi_0}} = \frac{0.1379 - 0.5}{\sqrt{\frac{.5(1-.5)}{435}}} = \frac{-0.3621}{0.02397} = -15.11$

One sided vs. two sided

Central message #1: If you are doing a one-sided test *and the estimator IS NOT in the region of the null hypothesis*, don't multiply by 2 when calculating the p-value.

- Go back to our example of women in the U.S. House.
- If $H_0 : \pi > 0.5, H_a : \pi < 0.5$
- The test statistic IS in the region covered by the alternative hypothesis and IS NOT in the region covered by the null.
- This might still be written as $H_0 : \pi = 0$, but it should still be $H_a : \pi < 0.5$.
- $Z = \frac{\hat{\pi}-0.5}{\sigma_{\pi_0}} = \frac{0.1379-0.5}{\sqrt{\frac{.5(1-.5)}{435}}} = \frac{-0.3621}{0.02397} = -15.11$
- Note that we are still using $\pi_0 = 0.5$. This is the smallest value possible for π in the region covered by the null hypothesis.

One sided vs. two sided

Central message #1: If you are doing a one-sided test *and the estimator IS NOT in the region of the null hypothesis*, don't multiply by 2 when calculating the p-value.

- Go back to our example of women in the U.S. House.
- If $H_0 : \pi > 0.5, H_a : \pi < 0.5$
- The test statistic IS in the region covered by the alternative hypothesis and IS NOT in the region covered by the null.
- This might still be written as $H_0 : \pi = 0$, but it should still be $H_a : \pi < 0.5$.
- $Z = \frac{\hat{\pi} - 0.5}{\sigma_{\pi_0}} = \frac{0.1379 - 0.5}{\sqrt{\frac{.5(1-.5)}{435}}} = \frac{-0.3621}{0.02397} = -15.11$
- Note that we are still using $\pi_0 = 0.5$. This is the smallest value possible for π in the region covered by the null hypothesis.
- We use $P(Z > |\frac{\hat{\pi} - 0.5}{\sigma_{\pi_0}}|)$ ... no absolute value signs around the Z.
- `pnorm(15.11, lower.tail=F)` $\approx 0.00005 \Rightarrow$ reject null.

One sided vs. two sided

Central message #2: If you are doing a one-sided test *and the estimator IS in the region of the null hypothesis*, don't multiply by 2 when calculating the p-value **and** choose the tail so that you get an answer larger than 0.5.

- Go back to our example of women in the U.S. House.
- If $H_0 : \pi < 0.5, H_a : \pi > 0.5$

One sided vs. two sided

Central message #2: If you are doing a one-sided test *and the estimator IS in the region of the null hypothesis*, don't multiply by 2 when calculating the p-value **and** choose the tail so that you get an answer larger than 0.5.

- Go back to our example of women in the U.S. House.
- If $H_0 : \pi < 0.5, H_a : \pi > 0.5$
- The test statistics IS NOT in the region covered by the alternative hypothesis.

One sided vs. two sided

Central message #2: If you are doing a one-sided test *and the estimator IS in the region of the null hypothesis*, don't multiply by 2 when calculating the p-value **and** choose the tail so that you get an answer larger than 0.5.

- Go back to our example of women in the U.S. House.
- If $H_0 : \pi < 0.5, H_a : \pi > 0.5$
- The test statistics IS NOT in the region covered by the alternative hypothesis.
- This might still be written as $H_0 : \pi = 0$, but it should still be $H_a : \pi > 0.5$.

One sided vs. two sided

Central message #2: If you are doing a one-sided test *and the estimator IS in the region of the null hypothesis*, don't multiply by 2 when calculating the p-value **and** choose the tail so that you get an answer larger than 0.5.

- Go back to our example of women in the U.S. House.
- If $H_0 : \pi < 0.5, H_a : \pi > 0.5$
- The test statistics IS NOT in the region covered by the alternative hypothesis.
- This might still be written as $H_0 : \pi = 0$, but it should still be $H_a : \pi > 0.5$.
- $Z = \frac{\hat{\pi} - 0.5}{\sigma_{\pi_0}} = \frac{0.1379 - 0.5}{\sqrt{\frac{.5(1-.5)}{435}}} = \frac{-0.3621}{0.02397} = -15.11$

One sided vs. two sided

Central message #2: If you are doing a one-sided test *and the estimator IS in the region of the null hypothesis*, don't multiply by 2 when calculating the p-value **and** choose the tail so that you get an answer larger than 0.5.

- Go back to our example of women in the U.S. House.
- If $H_0 : \pi < 0.5, H_a : \pi > 0.5$
- The test statistics IS NOT in the region covered by the alternative hypothesis.
- This might still be written as $H_0 : \pi = 0$, but it should still be $H_a : \pi > 0.5$.
- $Z = \frac{\hat{\pi} - 0.5}{\sigma_{\pi_0}} = \frac{0.1379 - 0.5}{\sqrt{\frac{.5(1-.5)}{435}}} = \frac{-0.3621}{0.02397} = -15.11$
- We want $P(Z \leq |\dfrac{\hat{\pi} - 0.5}{\sigma_{\pi_0}}|)$

One sided vs. two sided

Central message #2: If you are doing a one-sided test *and the estimator IS in the region of the null hypothesis*, don't multiply by 2 when calculating the p-value **and** choose the tail so that you get an answer larger than 0.5.

- Go back to our example of women in the U.S. House.
- If $H_0 : \pi < 0.5, H_a : \pi > 0.5$
- The test statistics IS NOT in the region covered by the alternative hypothesis.
- This might still be written as $H_0 : \pi = 0$, but it should still be $H_a : \pi > 0.5$.
- $Z = \frac{\hat{\pi}-0.5}{\sigma_{\pi_0}} = \frac{0.1379-0.5}{\sqrt{\frac{.5(1-.5)}{435}}} = \frac{-0.3621}{0.02397} = -15.11$
- We want $P(Z \leq |\frac{\hat{\pi} - 0.5}{\sigma_{\pi_0}}|)$
- `pnorm(15.11, lower.tail=T)` $\approx 0.99995 \Rightarrow$ do not reject the null.
- Notice that `lower.tail = T`