Testing Hypotheses

EDP 613

Week 9

A Note About The Slides

Currently the equations do not show up properly in Firefox. Other browsers such as Chrome and Safari do work

A Note About Previous Items

We're going to use some introduce some concepts from Chapter 7 here as well

Definitions

The **margin of error** (**MoE**) is

> - *Formally* : the range of values above and below a sample statistic within a confidence interval
> - *In a nutshell* : how many percentage points your results will differ from the real population value
> - NOT the same as a **confidence interval**

(Statistical) **significance** is

> - *Formally* : a measure of the probability of the null hypothesis being true compared to the acceptable level of uncertainty regarding the true answer
> - *In a nutshell* : a result is probably not due to chance so it is likely real
> - NOT the same as a **practical significance**

Statistical Methods I

Interpretation

Example result:

A **95%** confidence interval with a
3% margin of error

What it means:

Your statistic will be within
3 percentage points of the real population value
95% of the time

The **MoE** is a probability!

Back to Hypothesis Testing

Recall

The **null hypothesis** H_0 states

> - *Formally* : that a parameter is equal to a specific value
> - *Informally* : nothing probably happened

The **alternative hypothesis** H_1 states

> - *Formally* : that a parameter differs from the value specified by the null hypothesis
> - *Informally* : something probably happened

Statistical Methods I

More about Hypothesis Testing

Say a null hypothesis is H_0: $\mu = 50$. Then three things can occur from a frequentist perspective

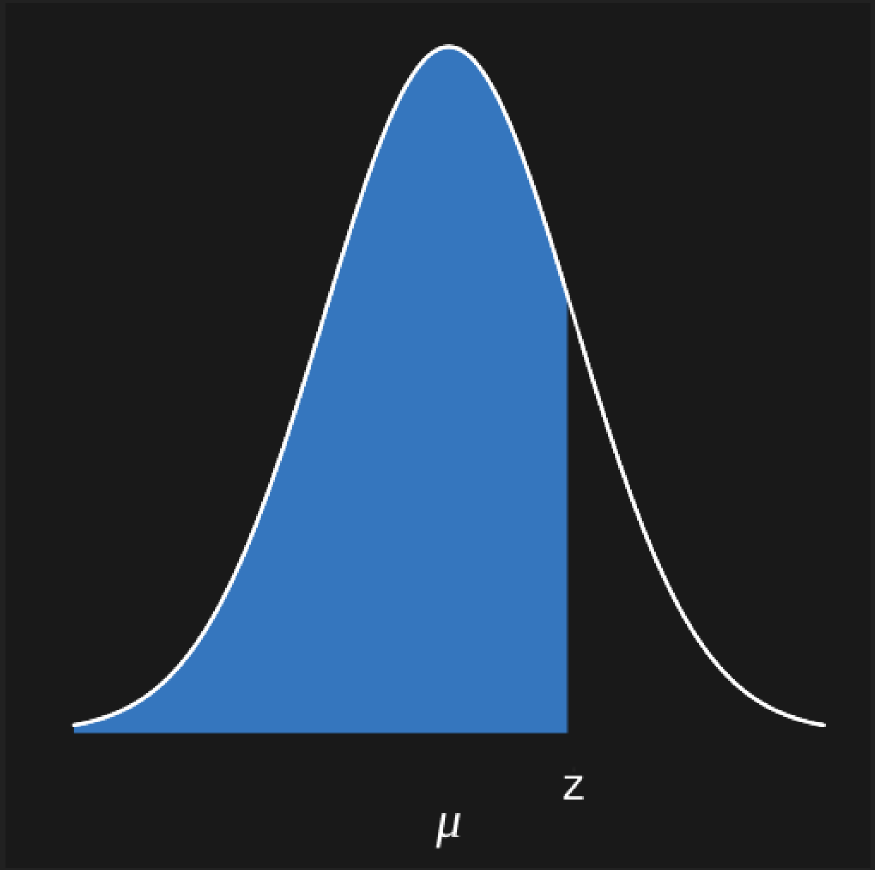
- $H_1 < 50$: alternative hypothesis states that the parameter is *less* than the value of the null.
    - You know to test for everything to the **left** of $H_1 = 50$.
    - Called a **left-tailed test**

- $H_1 > 50$: alternative hypothesis states that the parameter is *more* than the value of the null.
    - You know to test for everything to the **right** of $H_1 = 50$.
    - Called a **right-tailed test**

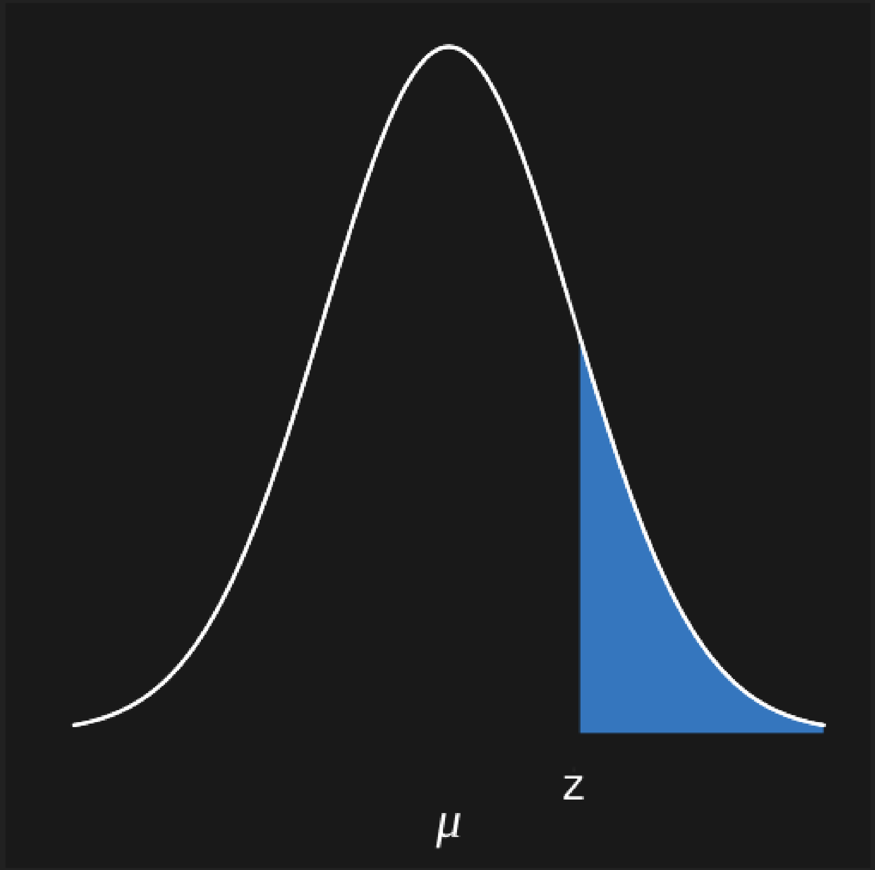
- $H_1 \neq 50$: alternative hypothesis states that the parameter is *note* the value of the null.
    - You know to test for everything to the **left** and **right** of $H_1 = 50$.
    - Called a **two-tailed test**

Statistical Methods I

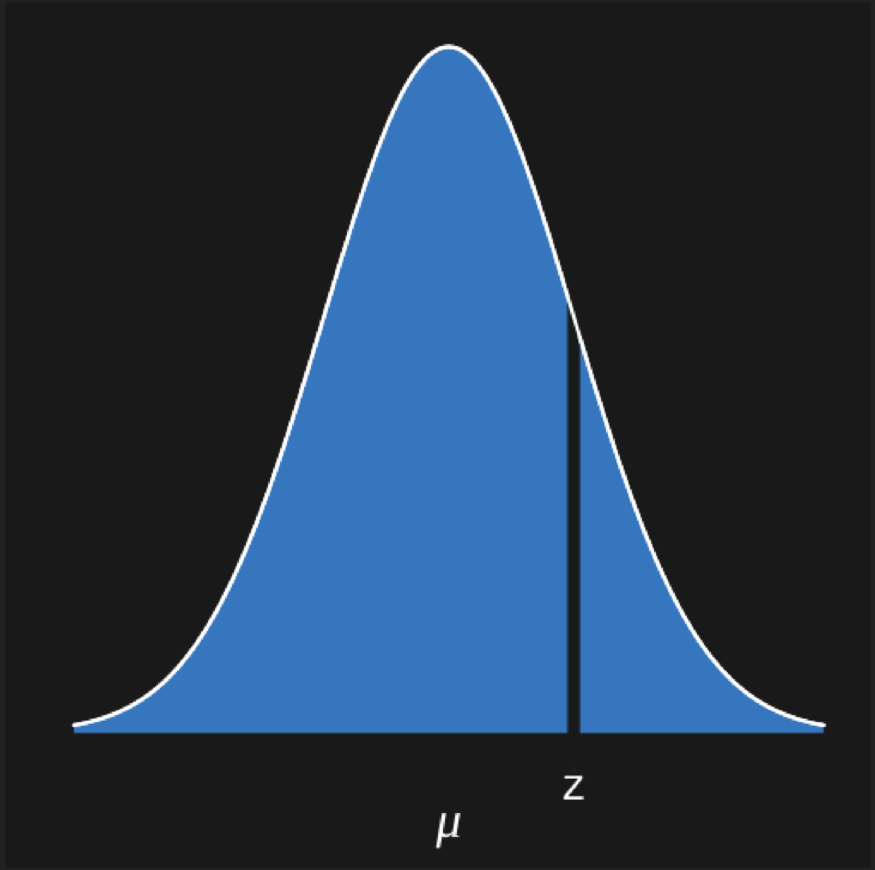
$H_1 < H_0$: Left-Tailed Test



$H_1 > H_0$: Right-Tailed Test



$H_1 \neq H_0$: Two-Tailed Test



Testing Methods

Critical Value (CV)

p - **value**

Critical Value

- Uses a **test statistic** - determines how strong the disagreement between a sample mean and a null hypothesis
- Idea: We should reject H_0 if the value of the test statistic is unusual when we assume H_0 to be true
- Process
  - We choose a *CV* which forms a boundary between values that are considered unusual and values that are not
  - The region containing the unusual values is called the **critical region**
  - If the value of the test statistic is in the critical regions, we *reject* H_0

Statistical Methods I

Transitioning (Again!)

Going from a z-distribution

known population variance using z-scores

to

a t-distribution

known sample variance using t-tests

Note

We have to assume that for a large enough sample size, the t-distribution will closely match, or estimate, the a z-distribution

Things to note

- t-distribution table can be located in Appendix C
- Assumptions
  - *Normality* - Samples are drawn from a population that fits a bell curve
  - *Independence* - Samples do not share values
  - *Random Sampling* - Samples are randomized
  - ***Homogeneity*** (for more than one sample) - Samples have a similar makeup

Statistical Methods I

Steps to Solving

1. *Interpret the Question into Layman's Terms*
2. *Set Acceptable Threshold of Committing a Type I Error*
3. *State the Research Hypothesis*
4. *Calculate the Test Statistic*
5. *Determine the Critical Value*
6. *State the Decision Rule*
7. *Interpret the Results*

Statistical Methods I

One-sample t-test

- allows us to determine whether the mean of a sample data set is different than a known value
- used when the population variance is not known
- can be used when the sample size is small ~ typically $N < 30$.

What the heck is a *degree of freedom*?

1. Forget statistics
2. Say you only own seven hats and want to wear a different one each day of the week.
3. Process
    - Day 1: Choose from 7
    - Day 2: Choose from 6

      .
      .

      .
      .
    - Day 6: Choose from 2
    - Day 7: Choose from 1

u had 7 - 1 = **6** *days of hat freedom!*

That's essentially a degree of freedom (written *df*)

One Sample Mean

t-distribution	*with degrees of freedom*
$t = \dfrac{\overline{Y} - \mu}{s/\sqrt{N}}$	$df = N - 1$

Example

Is the the median household income of West Virginia counties different than the national average?

Region	Median Household Income	Median Male Salary	Median Female Salary	Population	Households
United States	$68,703	$57,511	$43,820	331,449,281	139,684,244
West Virginia	$48,850	$57,456	$47,299	1,793,716	763,831

Statistical Methods I

Interpret the Question into Layman's Terms

From **Median Household Income**, we think WV (48,850 USD) is *practically lower* than the US (68,850 USD), but is it *significantly lower*?

Set Acceptable Threshold of Committing a Type I Error

$\alpha = 0.5$

State the Research Hypothesis

H_0

The **Median Household Income** of West Virginia counties is NOT significantly less than the national average

H_1

The **Median Household Income** of West Virginia counties is significantly less than the national average

Statistical Methods I

Calculate the Test Statistic (1/2)

Population

$$\mu = \$68,703$$

Sample

$$N = 61$$

$$s = \$5075.28$$

$$\overline{Y} = \$45732.20$$

df

$$61 - 1 = 60$$

Statistical Methods I

Calculate the Test Statistic (2/2)

$$t = \frac{45732.20 - 68703.00}{5075.28/\sqrt{61}}$$

$$\approx -0.579$$

Determine the Critical Value

In Appendix C, we see that for $df = 60$ at $\alpha = 0.05$, we have $t_{crit} = 1.671$

Statistical Methods I

State the Decision Rule

Since $-0.579 < 1.671$ we reject H_0

Interpret the Results

*The **Median Household Income** is significantly less than the national average!*

Statistical Methods I

Two-sample Mean

Used to compare one sample mean to another.

- We use two different test:
  - Equal variances
  - Unequal variances (assumed)
- **Homoscedasticity** – the assumption of equal variances.

Difference Between Two Independent Means

- Observations in each sample are not related
- Need to compare differences between the sample means

Estimated Standard Error

$$S_{\overline{Y_1}-\overline{Y_1}} = \sqrt{\frac{(N_1-1)\cdot s_1^2+(N_2-1)\cdot s_2^2}{(N_1+N_2)-2}} \cdot \sqrt{\frac{N_1+N_2}{N_1\cdot N_2}}$$

difference between means *t*-statistic

$$t = \frac{\overline{Y_1}-\overline{Y_2}}{S_{\overline{Y_1}-\overline{Y_1}}}$$

df

$$df = (N_1+N_2)-2$$

Difference Between Proportions

.pull-left[

Estimated Standard Error

$$S_{p_1-p_2} = \sqrt{\frac{p_1(1-p_1)}{N_1} + \frac{p_2(1-p_2)}{N_2}}$$

difference between means *z*-statistic

$$z = \frac{p_1 - p_2}{S_{p_1-p_2}}$$

Example

As part of the Pew Internet and American Life Project, researchers conducted two surveys in late 2009. The first survey asked a random sample of 800 U.S. teens about their use of social media and the Internet. A second survey posed similar questions to a random sample of 2253 U.S. adults. In these two studies, 73% of teens and 47% of adults said that they use social-networking sites. Use these results to construct and interpret a 95% confidence interval for the difference between the proportion of all U.S. teens and adults who use social-networking sites.

Interpret the Question into Layman's Terms

Is there a difference between the percent of teens and adults who use social networking sites?

Set Acceptable Threshold of Committing a Type I Error

$\alpha = 0.05$

State the Research Hypothesis

H_0

There is no difference between the proportion of teens and adults who use social media

H_1

There is a difference between the proportion of teens and adults who use social media

Statistical Methods I

Calculate the Test Statistic

$$S_{p_1-p_2} = \sqrt{\frac{0.73(1-0.73)}{800} + \frac{0.47(1-0.47)}{2253}}$$

$$\approx 0.0189$$

$$z = \frac{0.73 - 0.47}{0.0189}$$

$$\approx 13.76$$

State the Decision Rule

$z = 13.76$ is greater than p value implying that we reject H_0

Statistical Methods I

Interpret the Results

We are 95% confident that in late 2009 more teens than adults in the United States engaged in social media

That's it. Take a break before our R session!