Testing Hypotheses

EDP 613

Week 10

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.

Thin Margins

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

> - Formally - the range of values above and below a sample statistic within a confidence interval.
> - In Better Terms - how many percentage points your results will differ from the real population value.

Note: We're NOT talking about the confidence interval!

Interpretation

- 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** states

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

The **alternative hypothesis** states

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

More about Hypothesis Testing

Say a null hypothesis is $H_0\colon \mu = 50$. 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.

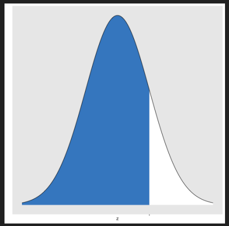
  > - You know to test for everything to the *right* of $H_1 = 50$.
  > - Called a **left-tailed test**

- $H_1 \neq 50$: alternative hypothesis states that the parameter is *note* the value of the null.

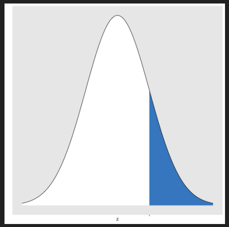
  > - You know to test for everything to the *left* and *right* of $H_1 = 50$.
  > - Called a **two-tailed test**

Visual Hypothesis Testing

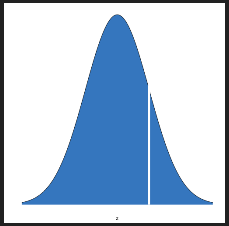
- $H_1 < H_0$:



- $H_1 > H_0$:



- $H_1 \neq H_0$:



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**.
    - f the value of the test statistic is in the critical regions, we *reject* H_0.

Transitioning (Again!)

> - Going from a z-distribution (known population variance using z-scores) to a t-distribution (known sample variance using t-tests)

- We 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

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. Uses n-1 **degrees of freedom** (*df*).

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 3: ...

   - Day 6: ... from 2

   - Day 7: Only one choice left!

4. You had $7 - 1 = 6$ days of hat freedom!

This is basically what a *df* is.

Formula

- $t = \frac{\overline{Y} - \mu}{s/\sqrt{N}}$
- $df = N - 1$

Example

We want to determine if the mean per capita income of West Virginia counties is different than the national average, and we suspect based on *a priori* (fancy term for *before*) knowledge that it is lower.

- H_0: The per capita income of West Virginia counties is not significantly less than the national average.
- H_1: The per capita income of West Virginia counties is significantly less than the national average.

Step 1

Determine if West Virginia county per capita income is normally distributed.

- If *yes*, proceed with a one-sample *t*-test
- If *no*, you may still be able to use a *t&-test but it requires additional assumptions

NOTE: We do not need to know the normality of US per capita income. All we need to know is the mean.

Region	Per Capita Income	Median Household Income	Median Family Income	Population	Households
United States	$27,334	$51,914	$62,982	308,745,538	116,716,292
West Virginia	$19,443	$38,380	$48,896	1,852,994	763,831

- From `Per Capita Income`, we suspect WV is lower but is the difference significant enough?
- Remember that H_0 is based on what we see in the data.

Assumptions

- Population mean is $27,334
- Take a sample of 61 people with standard deviation \$3075.28 and mean \$19253.2
- *df*: $61 - 1 = 60$
- CV: ± 2

Calulations

$$t = \frac{19253.2 - 27334.0}{3075.28/\sqrt{50}} \approx -18.58$$

So we reject H_0.

Interpretation: The per capita income is significantly less than the national average!

Two-sample t-test

Used to compare one sample mean to another.

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

Formulas

- $S_{\overline{Y_1}-\overline{Y_1}} = \sqrt{\frac{(N_1 - 1) \cdot s_1^2 + (N_2 - 1) \cdot s_2^2}{(N_1 + N_2) - 2}} \cdot \sqrt{\frac{N_1 + N_2}{N_1 \cdot N_2}}$
- $t = \frac{\overline{Y_1} - \overline{Y_2}}{S_{\overline{Y_1}-\overline{Y_1}}}$
- $df = (N_1 + N_2) - 2$

Example

At the Olympic level of competition, even the smallest factors can make the difference between winning and losing.

Pelton (1983) has shown that Olympic marksmen shoot much better if they fire between heartbeats, rather than squeezing the trigger during a heartbeat.

The small vibration caused by a heartbeat appears to be sufficient to affect the marksman's aim.

Samples

A sample of six (6) Olympic marksmen fires a series of rounds while a researcher records heartbeats. For each marksman, an accuracy score (out of 100) is recorded for shots fired during heartbeats and for shots fired between heartbeats. Do the data indicate a significant difference?

During Heartbeats	Between Heartbeats
93	98
90	94
95	96
92	91
95	97
91	97

Setup

- Research question: Is better accuracy achieved by marksmen when firing the trigger between heartbeats than during a heartbeat?
- Hypothesis:
  - $H_0 : \mu \geq 0$ – The average accuracy score is equal to or better when measured between heartbeats than during heartbeats.
  - $H_1 : \mu < 0$ – The average accuracy score is worse when measures between heartbeats than during heartbeats.
- Alpha: α= 0.05

Appendix C:

- $df = 6 - 1 = 5$
- $t_{critical} = -2.015$

During Heartbeats	Between Heartbeats	Difference	Variance of Difference
93	98	-5	25
90	94	-4	16
95	96	-1	1
92	91	1	1
95	97	-2	4
91	97	-6	36
NA	NA	-17	83

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