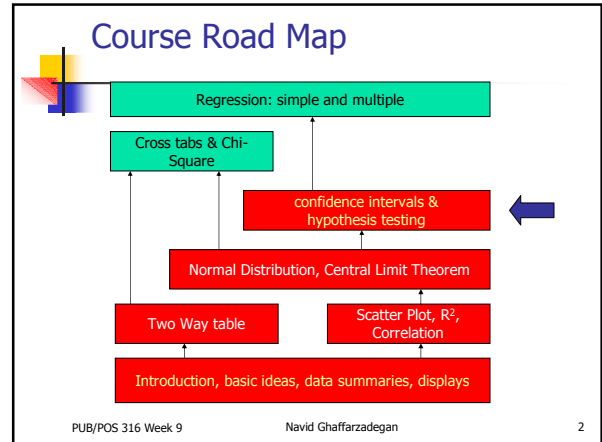


PUB – POS 316
Week 9

Hypothesis testing

Navid Ghaffarzadegan
 navidg@gmail.com
 Last updated – March 27, 10



Introduction

- Hypothesis
- P-Values
- Hypothesis testing
 - Z-test

PUB/POS 316 Week 9 Navid Ghaffarzadegan 3

Introduction

- Last time we talked about margin of errors and confidence intervals.
- Example: Average price of coffee in Albany:
 - \$ 1.85 ±\$0.35 (with 95% confidence)
 - Or: interval: (\$1.5,\$2.2)
- That was “a” systematic way of stating inferences.

PUB/POS 316 Week 9 Navid Ghaffarzadegan 4

Introduction

- Another way is:
 1. to state hypothesis about a phenomenon,
 2. report possibility that the hypothesis is correct, and
 3. accept or reject it.
- Our class is about this way of investigation.

PUB/POS 316 Week 9 Navid Ghaffarzadegan 5

1. Hypothesis

PUB/POS 316 Week 9 Navid Ghaffarzadegan 6

1. Hypothesis

Hypothesis about how something works

Design a study

Test the hypothesis (reject/accept)

reject/accept hypothesis

PUB/POS 316 Week 9 Navid Ghaffarzadegan 7

1. Hypothesis

Hypothesis about how something works

Design a study

Test the hypothesis (reject/accept)

reject/accept hypothesis

PUB/POS 316 Week 9 Navid Ghaffarzadegan 8

1. Hypothesis

- Null hypothesis
 - The statement being tested in a test of significance called the null hypothesis. – The status quo.
 - Usually the null hypothesis is a statement of “no effect” or “no difference.”
 - We love to reject it! →
- Alternative hypothesis

PUB/POS 316 Week 9 Navid Ghaffarzadegan 9

1. Hypothesis

- H_0 = Null hypothesis
- H_a = Alternative hypothesis
- Hypotheses are about the population, we study them through our sample.

PUB/POS 316 Week 9 Navid Ghaffarzadegan 10

1. Hypothesis

- Example: We want to study if a health care plan (like free vaccination for H1-N1) had any effect.
 - H_0 = Null hypothesis:
 - The policy did not have any effect on health
 - H_a = Alternative hypothesis
 - The policy had a positive effect
- OR:
 - The policy had an effect (positive or negative)
- One sided vs. two sided tests.

PUB/POS 316 Week 9 Navid Ghaffarzadegan 11

1. Hypothesis

- Example: We want to study if a free milk plan in schools helps students health
 - H_0 = Null hypothesis:
 - The plan does not have any effect on health
 - H_a = Alternative hypothesis
 - The plan has a positive effect
- OR:
 - The plan has an effect (positive or negative)
- One sided vs. two sided tests.

PUB/POS 316 Week 9 Navid Ghaffarzadegan 12

1. Hypothesis

Example: We want to study if job satisfaction is more in non-profit organizations than for-profit organizations:

- H_0 = Null hypothesis:
 - Job satisfaction is equal
- H_a = Alternative hypothesis
 - Job satisfaction in non-profit is more

OR:

- Job satisfaction is different between non-profit and for profit

1. Hypothesis

■ *A tobacco company claims that a new technology results in the mean nicotine content of a cigarette being less than 1.5 milligrams.*

- What are the hypotheses?
- State the question statistically:

■ $H_0: \mu \geq 1.5$

■ $H_a: \mu < 1.5$

1. Hypothesis

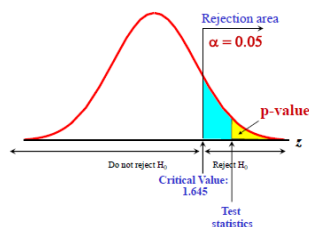
- After stating the hypotheses we should investigate them.
- Find the probability that the null hypothesis is correct (and see if we can reject it!)
- P-values.

2- P-Values

- The probability that H_0 is true, based on our data.
- The smaller P-Value, the stronger the evidence against H_0 !
- So, we go and find p-value for our hypotheses, compare it with a critical number (α), and see if we can reject the null-hypotheses. [we love to have small p-values!]

2- P-Values

■ Find p-value for our hypothesis, compare it with a critical number (α), and see if we can reject the null-hypothesis.



2- P-Values

■ Example: A nutritionist is interested in determining whether the average caloric intake of patients on the standard hospital diet is different from 2000 calories/day. Based on a sample of 32 patients, she found that the average caloric intake was 2124 calories/day. Past research indicated that the population standard deviation was 200 calories/day.

- At the 5% level of significance, can the study conclude that the average caloric intake of patients on the standard hospital diet is different from 2000 calories/day?

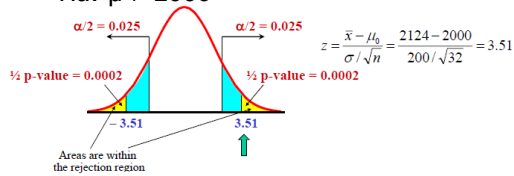
(source: Carnegie Mellon University, 90-711, Empirical Methods)

2- P-Values

- State the hypotheses:

- $H_0: \mu = 2000$

- $H_a: \mu \neq 2000$



p-value is $P(z \leq -3.51 \text{ or } z \geq 3.51) = 0.0002 + 0.0002 = 0.0004$

PUB/POS 316 Week 9

Navid Ghaffarzadegan

19

3- Z-test

- Then compare your p-value with the critical value

- (both of these numbers came from the z-distribution table)

PUB/POS 316 Week 9

Navid Ghaffarzadegan

20

Example

- Example: You're an analyst for Ford. You want to make sure if the average miles per gallon of Escorts is less than 32 mpg. Similar models have a standard deviation of 3.8 mpg. You take a sample of 60 Escorts & compute a sample mean of 30.7 mpg.

- At the 0.05 level, is there evidence that the miles per gallon is less than 32?

(source: Carnegie Mellon University, 90-711, Empirical Methods)

PUB/POS 316 Week 9

Navid Ghaffarzadegan

21

Example

- State the hypotheses:

- $H_0: \mu \geq 32$

- $H_a: \mu < 32$

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{30.7 - 32}{3.8 / \sqrt{60}} = -2.65$$

$$P(z < -2.65) = ?$$

PUB/POS 316 Week 9

Navid Ghaffarzadegan

22

Example

- Example: You're an analyst for Ford. You want to find out if the average miles per gallon of Escorts is less than 32 mpg. Similar models have a standard deviation of 3.8 mpg. You take a sample of 60 Escorts & compute a sample mean of 30.7 mpg.

- At the 0.01 level, is there evidence that the miles per gallon is less than 32?

(source: Carnegie Mellon University, 90-711, Empirical Methods)

PUB/POS 316 Week 9

Navid Ghaffarzadegan

23

Example

- Example: You performed a significance test of $H_0: \mu = 25$ based on an SRS of $n=100$. Assume the standard deviation in your population is equal to 20.

- If you came up with 28 in your sample, what is the p-value for $H_0: \mu = 25$? Can you reject that in 0.05 level of significance?

(source: Carnegie Mellon University, 90-711, Empirical Methods)

PUB/POS 316 Week 9

Navid Ghaffarzadegan

24

Review

- We have learned two systematic ways of making systematic inferences from data.
 - Margin of error
 - Hypothesis testing

Review: Margin of error

Remember: Only remembering the following equation is enough

$$x = \mu \pm z \cdot s \quad \Leftrightarrow \quad x = \mu \pm z \cdot \frac{\sigma}{\sqrt{n}}$$

$$x = \mu \pm z \cdot \sqrt{\frac{p(1-p)}{n}}$$

The procedure:

1. Write $x = \mu \pm z \cdot s$
2. Write the proper extension
3. Find z (you usually know it!), σ or p , and n

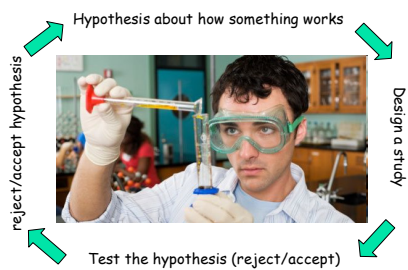
Review: Margin of error

- A recent pole compares the chances of candidate A with candidate B for winning a presidential election. 1000 people participated in this pole. 480 people said that they support candidate A, 436 people said they support candidate B, the rest not decided.
 - What is the 95% confidence interval for the result of each candidate?
 - Which candidate has a better position?
 - What about margin of error in 80% confidence?

Excel

- Standard Normal distribution
 - Remember our z-table? Excel has it! (read help)
 - NORMSDIST(z)
 - NORMSINV(area)→z
- Normal Distribution
 - {Not necessarily "standard" normal dis., but all kind of normal distribution}
 - NORMDIST
- Confidence intervals
 - CONFIDENCE(alpha,standard_dev,size)

Review: Hypothesis testing



Review: Hypothesis testing

- Important:
 - State your hypotheses correctly.
 - What is the status quo? H_0
 - What do we like to prove? $\rightarrow H_a$
 - Hypothesis of two states being equal? $\rightarrow H_0$
 - Still it will be difficult! Experience....

Review: Hypothesis testing

- Important:
 - Clarify if it is a one-tail test or two tail test? (one sided, or two sided?)
 - Based on that you can find
 - P-value
 - the critical value, α
 - Compare them. See if you can reject the null hypothesis.

PUB/POS 316 Week 9

Navid Ghaffarzadegan

31

Review: Hypothesis testing

- The **university believes** that the average hours of study among under grad students is **3.5 hours per day**. **You think students are studying harder!** And decide to conduct a study to test it. Based on a sample of **40 students** you see that the result is **4.0 hours per day**. Assuming the **standard deviation** of hours of study in the **population is 2 hours**, can you reject the belief?! What is the p-value?

PUB/POS 316 Week 9

Navid Ghaffarzadegan

32

Review: Hypothesis testing

- It is significant, so what?
- P- Value is more important.
- Don't ignore lack of significance.
- Plot the data. (scatter plots – bar charts)

PUB/POS 316 Week 9

Navid Ghaffarzadegan

33

Excel

- Normal distribution
- Read help for z-test.
- Commands: Give you p-values:
 - One tail test:
 - ZTEST(Array, μ)
 - e.g.: =ZTEST(A2:A11,4)
 - Two tail tests:
 - $2 * \text{MIN}(\text{ZTEST}(A2:A11,4),$
 - $1 - \text{ZTEST}(A2:A11,4))$

PUB/POS 316 Week 9

Navid Ghaffarzadegan

34

Two cases

- How to read and interpret confidence intervals in papers and professional reports.
 - The health policy paper
- How to read and interpret hypotheses and technical tables in papers and professional reports.
 - The organizational behavior paper

PUB/POS 316 Week 9

Navid Ghaffarzadegan

35