

## Unit 3: Foundations for inference

### 3. Hypothesis tests

Sta 101 - Spring 2016

Duke University, Department of Statistical Science

Dr. Çetinkaya-Rundel

Slides posted at [http://bit.ly/sta101\\_s16](http://bit.ly/sta101_s16)

Midterm 1: Feb 24, Wed

#### ► Preparation

- Come to class with questions on Monday
- Sample MT posted on course website

#### ► Rules

- Bring a calculator + cheat sheet (one sheet, both sides, typed or handwritten, must be prepared by you) + writing utensil
- We'll provide tables

1

## 1. Use hypothesis tests to make decisions about population parameters

Hypothesis testing framework:

1. Set the hypotheses.
2. Check assumptions and conditions.
3. Calculate a *test statistic* and a p-value.
4. Make a decision, and interpret it in context of the research question.

2

## Hypothesis testing for a population mean

### 1. Set the hypotheses

- $H_0 : \mu = \text{null value}$
- $H_A : \mu < \text{or } > \text{ or } \neq \text{null value}$

### 2. Check assumptions and conditions

- Independence: random sample/assignment, 10% condition when sampling without replacement
- Sample size / skew:  $n \geq 30$  (or larger if sample is skewed), no extreme skew

### 3. Calculate a *test statistic* and a p-value (draw a picture!)

$$Z = \frac{\bar{x} - \mu}{SE}, \text{ where } SE = \frac{s}{\sqrt{n}}$$

### 4. Make a decision, and interpret it in context of the research question

- If p-value  $< \alpha$ , reject  $H_0$ , data provide evidence for  $H_A$
- If p-value  $> \alpha$ , do not reject  $H_0$ , data do not provide evidence for  $H_A$

3

### Application exercise: 3.2 Hypothesis testing for a single mean

See course website for details.

4

### Clicker question

Which of the following is the correct interpretation of the p-value from App Ex 3.2?

- (a) The probability that average GPA of Duke students has changed since 2001.
- (b) The probability that average GPA of Duke students has not changed since 2001.
- (c) The probability that average GPA of Duke students has not changed since 2001, if in fact a random sample of 63 Duke students this year have an average GPA of 3.58 or higher.
- (d) The probability that a random sample of 63 Duke students have an average GPA of 3.58 or higher, if in fact the average GPA has not changed since 2001.
- (e) The probability that a random sample of 63 Duke students have an average GPA of 3.58 or higher or 3.16 or lower, if in fact the average GPA has not changed since 2001.

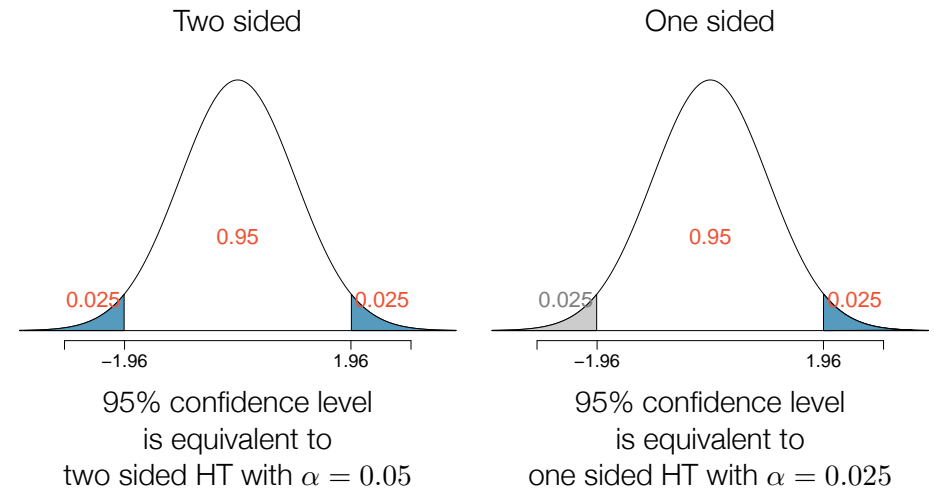
5

### Common misconceptions about hypothesis testing

1. P-value is the probability that the null hypothesis is true  
*A p-value is the probability of getting a sample that results in a test statistic as or more extreme than what you actually observed (and in favor of the null hypothesis) if in fact the null hypothesis is correct. It is a conditional probability, conditioned on the null hypothesis being correct.*
2. A high p-value confirms the null hypothesis.  
*A high p-value means the data do not provide convincing evidence for the alternative hypothesis and hence that the null hypothesis can't be rejected.*
3. A low p-value confirms the alternative hypothesis.  
*A low p-value means the data provide convincing evidence for the alternative hypothesis, but not necessarily that it is confirmed.*

6

### 2. Hypothesis tests and confidence intervals at equivalent significance/confidence levels should agree



7

Clicker question

What is the confidence level for a confidence interval that is equivalent to a two-sided hypothesis test at the 1% significance level? *Hint: Draw a picture and mark the confidence level in the center.*

- (a) 0.80
- (b) 0.90
- (c) 0.95
- (d) 0.98
- (e) 0.99

8

Clicker question

What is the confidence level for a confidence interval that is equivalent to a one-sided hypothesis test at the 1% significance level? *Hint: Draw a picture and mark the confidence level in the center.*

- (a) 0.80
- (b) 0.90
- (c) 0.95
- (d) 0.98
- (e) 0.99

9

Clicker question

A 95% confidence interval for the average normal body temperature of humans is found to be (98.1 F, 98.4 F). Which of the following is true?

- (a) The hypothesis  $H_0 : \mu = 98.2$  would be rejected at  $\alpha = 0.05$  in favor of  $H_A : \mu \neq 98.2$ .
- (b) The hypothesis  $H_0 : \mu = 98.2$  would be rejected at  $\alpha = 0.025$  in favor of  $H_A : \mu > 98.2$ .
- (c) The hypothesis  $H_0 : \mu = 98$  would be rejected using a 90% confidence interval.
- (d) The hypothesis  $H_0 : \mu = 98.2$  would be rejected using a 99% confidence interval.

10

3. Results that are statistically significant are not necessarily practically significant

Clicker question

All else held equal, will p-value be lower if  $n = 100$  or  $n = 10,000$ ?

- (a)  $n = 100$
- (b)  $n = 10,000$

11

		Decision	
		fail to reject $H_0$	reject $H_0$
Truth	$H_0$ true	✓	Type 1 Error, $\alpha$
	$H_A$ true	Type 2 Error, $\beta$	Power, $1 - \beta$

- ▶ A *Type 1 Error* is rejecting the null hypothesis when  $H_0$  is true:  $\alpha$ 
  - For those cases where  $H_0$  is actually true, we do not want to incorrectly reject it more than 5% of those times
  - Increasing  $\alpha$  increases the Type 1 error rate, hence we prefer to small values of  $\alpha$
- ▶ A *Type 2 Error* is failing to reject the null hypothesis when  $H_A$  is true:  $\beta$
- ▶ *Power* is the probability of correctly rejecting  $H_0$ , and hence the complement of the probability of a Type 2 Error:  $1 - \beta$

1. Use hypothesis tests to make decisions about population parameters
2. Hypothesis tests and confidence intervals at equivalent significance/confidence levels should agree
3. Results that are statistically significant are not necessarily practically significant
4. Hypothesis tests are prone to decision errors