

## Unit 3: Foundations for inference

### 3. Decision errors, significance levels, sample size & power

Sta 101 - Spring 2015

Duke University, Department of Statistical Science

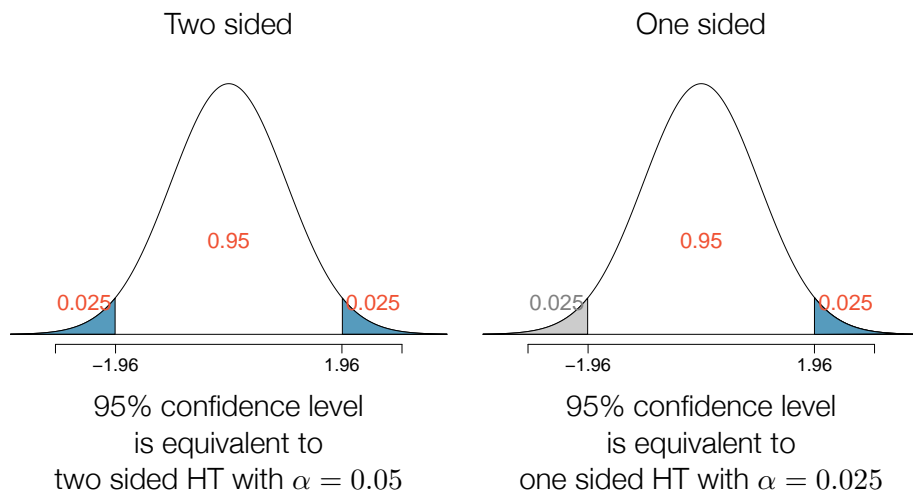
February 16, 2015

Dr. Çetinkaya-Rundel

Slides posted at <http://bitly.com/sta101sp15>

- ▶ PA3 due tonight
- ▶ Midterm review: ?

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



#### Clicker question

What is the significance level of a two-sided hypothesis test that is equivalent to a 90% confidence interval? *Hint: Draw a picture and mark the confidence level in the center.*

- (a) 0.001
- (b) 0.01
- (c) 0.025
- (d) 0.05
- (e) 0.10

Clicker question

What is the significance level of a one-sided hypothesis test that is equivalent to a 90% confidence interval? *Hint: Draw a picture and mark the confidence level in the center.*

- (a) 0.001
- (b) 0.01
- (c) 0.025
- (d) 0.05
- (e) 0.10

4

Clicker question

What is the confidence level of a confidence interval that is equivalent to a two-sided hypothesis test with  $\alpha = 0.01$ . *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

5

Clicker question

What is the confidence level of a confidence interval that is equivalent to a one-sided hypothesis test with  $\alpha = 0.01$ . *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

6

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  $H_A : \mu = 98.2$  would be rejected at  $\alpha = 0.05$  in favor of  $H_A : \mu \neq 98.2$ .
- (b) The hypothesis  $H_A : \mu = 98.2$  would be rejected at  $\alpha = 0.025$  in favor of  $H_A : \mu > 98.2$ .
- (c) The hypothesis  $H_A : \mu = 98$  would be rejected using a 90% confidence interval.
- (d) The hypothesis  $H_A : \mu = 98.2$  would not be rejected using a 99% confidence interval.

7

2. 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)  $n = 10,000$

8

3. Calculate the sample size *a priori* to achieve desired margin of error

Application exercise: 3.3 Sample size

See course website for details.

9

4. Hypothesis tests are prone to decision errors

		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:  $1 - \beta$

10

5. Power depends on the  $n$ ,  $a$ ,  $\alpha$ , effect size

Power can be increased (and hence Type 2 error rate can be decreased) by

- ▶ increasing the sample size
- ▶ decreasing the standard deviation of the sample (difficult to ensure but cautious measurement process and limiting the population so that it is more homogenous may help)
- ▶ increasing  $\alpha$
- ▶ increasing the **effect size**

11

1. Hypothesis tests and confidence intervals at equivalent significance/confidence levels should agree
2. Results that are statistically significant are not necessarily practically significant
3. Calculate the sample size a priori to achieve desired margin of error
4. Hypothesis tests are prone to decision errors
5. Power depends on the effect size,  $\alpha$ ,  $n$ , and  $s$