STA 291

Lecture 26

- Two types of errors in testing hypothesis.
- Connection between testing hypothesis and confidence intervals

STA 291 - Lecture 26

1

2

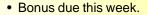
3

• A P-value that is smaller than 0.01 must also be smaller than 0.05

• A P-value that is smaller than 0.05 may or may not be smaller than 0.01

STA 291 - Lecture 26

- Read the wiki page on P-value:
- http://en.wikipedia.org/wiki/P-value



- Get familiar with the final exam formula sheet.
- 70% of Final exam will cover contents after 2nd midterm.

STA 291 - Lecture 26

4

5

6

In testing a hypothesis

• If our conclusion is to "reject the null hypothesis".....then we either made a correct decision or we made a type I error.

STA 291 - Lecture 26

 If our conclusion is "do not reject the null hypothesis" then we either made a correct decision or made a type <u>II</u> error.

Type I and Type II Errors

- Type I Error: The null hypothesis is rejected, even though it is true.
- Type II Error: The null hypothesis is not rejected, even though it is false.
- Setting the alpha-level (significance level) low protect us from type I Error. (the probability of making a type I error is less than alpha)

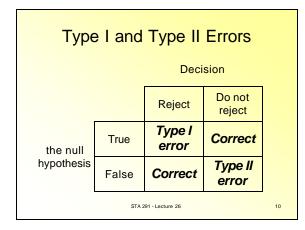
• The chance of making a Type II error can be made small by increasing the sample size. (assume you use the correct testing procedure)

STA 291 - Lecture 26

Decisions and Types of Errors in Tests of Hypotheses

• Terminology:

- The alpha-level (significance level) is a threshold number such that one rejects the null hypothesis if the *p*-value falls below it. The most common alpha-levels are 0.05 and 0.01
- The choice of the alpha-level reflects how cautious the researcher wants to be (when it comes to reject null hypothesis)





 If after all the calculations and review the evidences you decide to reject the null hypothesis H₀, you may be right or you may made a type 1 error.

 If after all the calculations you decide not to reject the null hypothesis H₀, you may be right or you may made a type II error.

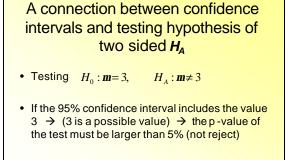
STA 291 - Lecture 26

• When sample size(s) increases, both error probabilities could be made to decrease.

- Strategy:
- keep type I error probability small by pick a small alpha.
- Increase sample size to force the probability of making type II error, Beta, small. (or increase Power)

STA 291 - Lecture 26

12



STA 291 - Lecture 26

• 🗲 also true.

 If the 95% confidence interval do not includes the value 3 → the p-value of the test must be smaller than 5% (reject).

- 🗲 also holds
- True for other parameters. True for other confidence levels.
- Only works for two sided *H*_A hypothesis.

STA 291 - Lecture 26

• Confidence interval for a parameter consist of those values that are plausible, not rejectable, in a testing setting (of two sided H_A hypothesis)

13

- So, the confidence interval consists of those values of parameters that are compatible with the observed data data.
- 95% confidence → 5% error ← → p-value of 5%
- 90% confidence → 10% error ← → p-value of 10% etc.

STA 291 - Lecture 26

• suppose the 95% confidence interval computed from data for **m** is [2.2, 4.1] any test of

 $H_0: \mathbf{m} = 3, \quad H_A: \mathbf{m} \neq 3$ $H: \mathbf{m} = 2.8, \quad H: \mathbf{m} \neq 2.8$

$$H_0: \mathbf{m} = 4, \qquad H_A: \mathbf{m} \neq 4.$$

• Would result a p-value larger than 5% (not reject) i.e. any value inside [2.2, 4.1] are plausible.

STA 291 - Lecture 26

$$H_0: m = 1.9, \quad H_A: m \neq 1.9$$

 $H_0: m = 4.8, \qquad H_A: m \neq 4.8$

Would result a p-value smaller than 5% (reject null, using alpha=0.05)

26

18

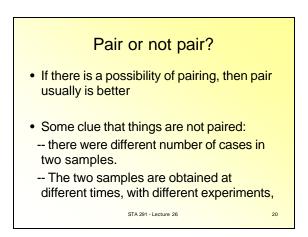
16

Why not always confidence interval?

- In some cases, confidence interval is hard to obtain,
- Yet testing a specific hypothesis is easier.
- Confidence interval amounts to obtain all values m_0 that you cannot reject as H_0

STA 291 - Lecture 26

19



Paired Experiment: focus on the differences

- One subject contribute two results, we often can focus on the difference of the two from the same subject.
- Sometimes not possible..... how long a mice live before cancer kill. Same mice cannot be used twice. Strength of the shipping packaging test of strength would destroy the package

STA 291 - Lecture 26

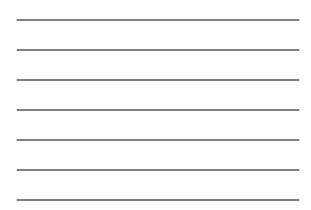
65 randomly chosen subjects are given two bottles of shampoos: A and B. After a week, each subject state which one they prefer

	Subject 1	Subject 2	Subject 3			
A	prefer		prefer	•		
В		prefer		•		
STA 291 - Lecture 26 2						



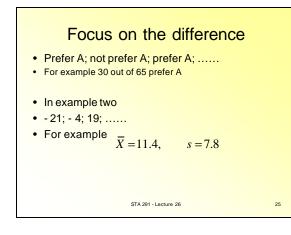
125 randomly chosen subjects are given two bottles of pills: A and B. After a month on each pill, report LDL cholesterol level

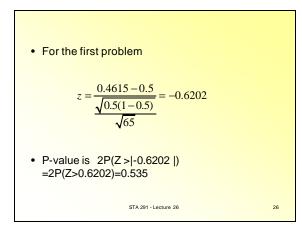
	Subject 1	Subject 2	Subject 3	•
A	167	155	233	•
В	188	159	214	· · · ·
	23			



• After we take the difference for each subject, the problem becomes a one sample problem:

- If the preference has 50-50% chance?
- If the difference has mean zero?





- If you use the computer to do the problem, the p-value will be slightly different. Due to the fact that our calculation is only an approximation (use CLT).
- Computer is more accurate.
- For sample size very large the difference goes away. (For example 300 subjects out of 650 prefer A)

```
STA 291 - Lecture 26
```

For the cholesterol problem

$$H_{0}: \mathbf{m} = 0, \quad H_{1}: \mathbf{m} \neq 0$$

$$c = \frac{11.4 - 0}{\frac{7.8}{\sqrt{125}}} = 16.34$$

$$r = 2P(Z > |16.34|) = 0.0000000000$$

• What would be a 95% confidence interval for the mu – the population mean of the difference of cholesterol, when using pill A/B?

STA 291 - Lecture 26

29

• Actually I should be looking up the t-table with degrees of freedom 125-1 = 124

(since I used s in place of sigma)

 Using t-table applet I get also a tiny pvalue (with more than 30 zero's after decimal point)

The formal conclusion: reject (overwhelmingly) the null hypothesis of difference = 0, imply the difference is not zero. Apparently the difference is positive – the average difference is 11.4.

This imply pill B has lower LDL values compared to pill A.

STA 291 - Lecture 26

Multiple Choice Question II

- The P-value for testing the null hypothesis mu=100 (two-sided) is P=0.0001. This indicates
 - a) There is strong evidence that mu = 100
 - b) There is strong evidence that mu > 100
 - c) There is strong evidence that mu < 100
 - d) If mu were equal to 100, it would be unusual (probability 0.0001) to obtain data such as those observed

STA 291 - Lecture 26

32

31

Multiple Choice Question

- A 95% confidence interval for mu is (96,110). Which of the following statements about significance tests for the same data are correct?
 - a) In testing the null hypothesis mu=100 (two-sided), P>0.05
 - b) In testing the null hypothesis mu=100 (two-sided), P<0.05
 - c) In testing the null hypothesis mu=x (two-sided), P>0.05 if x is any of the numbers inside the confidence interval
 - d) In testing the null hypothesis mu=x (two-sided), P<0.05 if x is any of the numbers outside the confidence interval

STA 291 - Lecture 26

Attendance Survey Question 26

-your name and section number

-Today's Question:

The online homework

A. is helpful, since I know if I aced it right away B. not helped me, because of computer glitches

STA 291 - Lecture 26

C. no opinion

This p is not that p-value

• P-value of a test procedure. Any hypothesis testing should result a p -value. It summarizes the strength of the evidence in the sample against H_0

 Proportion p or rate p or percentage p of success in the population: value specified in the null hypothesis to be tested. Only in the testing of the proportion, Bernoulli type populations

Test H₀: p=0.5 etc.

STA 291 - Lecture 26

35