STA 291
Lecture 29

• Review
Final Exam, Thursday, May 6

- When: 6:00pm-8:00pm
- Where: CB 106

Make-up exam: Friday 10:00am-12:00noon
  - Only by prior arrangement
  - Room still unknown, watch the web for update, or come to 8th floor POT on Friday
    - Update: Makeup room: CB 303
Final Exam, Thursday, May 6

- It will be approx. one and half length long compared to the two midterms. (i.e. if midterm have 20 questions, final will have approx. 30 questions).

- Similar mixture of open answer questions and multiple choice questions, compared to the midterms.

- Covers all the topics (comprehensive). But more on the later (testing hypothesis, confidence interval) materials.
• Formula sheet and tables will be provided.
Some topics we covered

• Testing hypothesis.
• Confidence intervals. (even though it had been covered in midterm exam II)
• Connection between the above 2 topic.

• Use of Z (Normal) table to find probability
• When to use t-table instead?
• Setting up the correct hypothesis:
  -- it is always about a population parameter(s)
• Find the correct formula for the hypothesis

• Computation of the test statistic, and the P-value (Need to use table)

• What to do if falls outside the range of table?
• Reach a conclusion by comparing the P-value to the alpha level. (report the P-value)

• Potential error (which type?)
Connection between testing hypothesis and confidence interval

• Given a confidence interval, you can tell if the P-value is above or below alpha

• Given a P-value you can tell if the confidence interval will contain \( \mu_0 \)
• Similar question on Exam II might re-appear on final.
Comparing paired Samples: Example

<table>
<thead>
<tr>
<th>Student</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>60</td>
<td>73</td>
<td>42</td>
<td>88</td>
<td>66</td>
<td>77</td>
<td>90</td>
<td>63</td>
<td>55</td>
<td>96</td>
</tr>
<tr>
<td>After</td>
<td>70</td>
<td>80</td>
<td>40</td>
<td>94</td>
<td>79</td>
<td>86</td>
<td>93</td>
<td>71</td>
<td>70</td>
<td>97</td>
</tr>
</tbody>
</table>

a) Compare the mean weights after and before the drug by
   i. finding the difference of the sample means
   ii. finding the mean of the difference scores. Compare. --- (same)
   iii. But SD is different, One SD or two SD’s?

b) Calculate and interpret the P-value for testing whether the mean change equals 0

c) Compare the mean weights after and before taking the drug by constructing and interpreting a 90% confidence interval for the population mean difference
Comparing Dependent Samples: Example (contd.)

Output from Statistical Software Package SAS

<table>
<thead>
<tr>
<th>Test</th>
<th>-Statistic-</th>
<th>-----p Value------</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student's t</td>
<td>t 4.216901</td>
<td>Pr &gt;</td>
</tr>
<tr>
<td>Sign</td>
<td>M 4</td>
<td>Pr &gt;=</td>
</tr>
<tr>
<td>Signed Rank</td>
<td>S 25.5</td>
<td>Pr &gt;=</td>
</tr>
</tbody>
</table>
Which method to chose?

- Two-year Italian study on the effect of condoms and the spread of AIDS
- Heterosexual couples where one partner was infected with HIV virus
- 171 couples who always used condoms: 3 partners became infected with HIV
- 55 couples who did not always use condoms: 8 partners became infected with HIV
- *Test whether the rates are significantly different. Report the P-value and interpret.*
Which Method to Choose?

- A study compares the mean level of contributions to political campaigns in Pennsylvania by registered Democrats, and registered Republicans.
Which Method to Choose?

- Example: Compare new drug to placebo in a double-blind clinical trial
  - 24 patients
  - Randomly pick 12 assign to placebo
  - The other 12 receive the new drug
  - Research question: Is there a different effect of placebo and new drug on a “response” on, for example, cholesterol, blood parameter, health status, weight,…
Which Method to Choose?

• Example: Which of two suntan lotions (labeled X and Y) provides better protection against sunburn
  – 8 subjects expose their backs to the sun for a certain time, protected by suntan lotion
  – Possible design:
    • Randomly pick 4 subjects use lotion X
    • the other 4 subjects use lotion Y
Which Method to Choose?

- Example: Which of two suntan lotions (labeled X and Y) provides better protection against sunburn
  - 8 subjects expose their backs to the sun for a certain time, protected by suntan lotion
  - Different design:
    - Each of the 8 subjects uses both suntan lotions at the same time
    - one lotion on the left side of the back, the other on the right side (use a coin flip to decide which side for X)
Multiple Choice Question

• Which of the following statements are true?
• “95% confidence” means that
  – 95% of the true population parameters are in the confidence interval
  – If we were to repeat the procedure of sampling and calculating confidence intervals from the same population, then 95% of the population parameters are going to be in every calculated interval
  – If we were to repeat the procedure of sampling and calculating confidence intervals from the same population, then 95% of the times our confidence interval will contain the true population parameter
Multiple choice Q

• If a test turns out to be significant at alpha-level 0.01. (what exactly this mean for the p-value?)

• Will the same test also be significant at 0.05 level?
• P-value is *NOT* the probability that the H0 is true.

• A small p-value mean that we saw something happened that is *hard to explain* by H0.
• A large p-value do not automatically means H0 is true. (2 possibilities: either H0 is true or there is too few data/info)

• Another H0 could have even larger P-value
Test vs. Confidence Interval

Assume that the $p$-value is equal to 0.043 for a test of the null hypothesis $H_0: \mu = 2$, with two-sided alternative.

What conclusion can we make about a 95% confidence interval for $\mu$?
• Study hard and good luck!