

# Statistics 101 L – Homework 11

**Due Wednesday, April 29, 2009**

**Reading:** April 17 – April 24                      Chapters 24, 25  
                  April 27 – May 1                        Chapter 26

**Assignment:**

1. For each of the following situations, indicate whether it is a two-sample problem or a paired sample problem. Explain briefly your choice. Also indicate whether you would construct a confidence interval or perform a test of hypothesis.
  - a) A study is conducted to determine whether taking an ACT preparation course improves the overall score on the ACT. A random sample of 40 high school seniors from a large school district is selected. Half are assigned at random to go through a special six week ACT preparation course. The other half does nothing special for those six weeks. At the end of the six weeks, the students are given the ACT. The school district would like to determine if students completing the preparation course have significantly higher ACT scores, on average, than those who did nothing.
  - b) The effectiveness of a weight loss program that combines exercise and diet is to be evaluated. Thirty volunteers are weighed before beginning the program and re-weighed after following the program for three months. One wishes to estimate, with 95% confidence, the mean difference between the post-program and pre-program weights.
  - c) Over the years, the game show Jeopardy has had more male winners than female winners. One explanation for this phenomenon could be a difference in the reaction times between men and women (men buzz in faster than women). A random sample of 15 previous male contestants and a random sample of 15 previous female contestants are selected and a test of reaction times is given. The show’s producers would like to determine if the mean reaction time of women is slower than the mean reaction time of men, on average.
  - d) In order to compare computer speeds, a set of benchmark programs are run and the CPU time required to complete each program is recorded. Six benchmark programs are selected. Each program is run on two computers and the CPU times are recorded. One wishes to estimate, with 98% confidence, the mean difference in CPU times for the two computers.
  
2. A major medical center in the Northeastern U.S. conducted a study looking at blood cholesterol levels and incidence of heart attack. Below are data from 16 people who had a heart attack and 20 people who did not have a heart attack.

<b>(1) Heart Attack</b>				<b>(2) No Heart Attack</b>				
242	186	266	206	182	222	198	192	238
318	294	282	234	198	188	166	204	182
224	276	262	360	178	202	164	230	186
310	280	248	258	162	182	218	170	200

- a) Is this an experiment or an observational study? Explain briefly.
- b) Compute 5-number summaries for each group.
- c) Construct side-by-side box plots. Compare the two groups in terms of center and spread.
- d) Describe how the individuals in the study need to be selected in order for the randomization condition to be met.

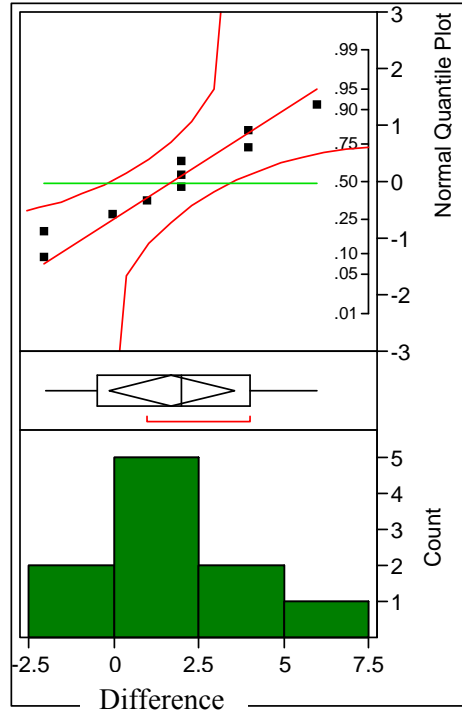
On the next page are summary statistics for the two groups.

(1) Heart Attack	(2) No Heart Attack
$\bar{y}_1 = 265.375$	$\bar{y}_2 = 193.1$
$s_1 = 43.645$	$s_2 = 21.623$
$n_1 = 16$	$n_2 = 20$

- e) Is there sufficient evidence to indicate that the mean cholesterol for people who have had a heart attack is greater than that for people who have not had a heart attack? Perform the appropriate test of hypothesis. Use  $df = 20$  and  $\alpha=0.05$ . Note: the nearly normal condition is satisfied.
- f) Construct a 95% confidence interval for the difference in population mean cholesterol levels.
3. As part of a study on pulse rates some participants (randomly assigned as controls) had their pulse rate taken and asked to sit quietly for 5 minutes. At the end of 5 minutes, their pulse rates were taken again. Below are the data.

Participant	1	2	3	4	5	6	7	8	9	10
Pulse rate 1	86	71	90	68	71	68	74	70	78	69
Pulse rate 2	88	73	88	72	77	68	76	71	82	67

- a) Why is this a paired sample problem?
- b) Based on the JMP output below, is the population mean difference zero or is it something different? You must take the information from the JMP output and use it to provide supporting evidence for your answer to this question.



Mean	1.7
Std Dev	2.5841397
Std Err Mean	0.8171767
upper 95% Mean	3.5485822
lower 95% Mean	-0.148582
N	10
Hypothesized Value	0
Actual Estimate	1.7
df	9
Std Dev	2.58414
	t Test
Test Statistic	2.0803
Prob >  t	0.0672
Prob > t	0.0336
Prob < t	0.9664