MATH 210 - Test #3 - 4/23/98 _____

When conducting hypothesis tests clearly state the *null* and *alternative hypotheses* and *p-value* for the test.

1. You are a marketer of fine clothing and wish to study the household incomes of people who shop in high-end fashion stores. You randomly survey 51 people shopping in such stores located in New York city and 61 people in such stores in Los Angeles. The results are as follows. (Incomes are measured in 1000s of dollars.)

[20]

	Sample Size (n)	Mean Income ($\bar{x}$)	Standard Deviation (s)
New York	51	95	10
Los Angeles	61	105	15

You wish to estimate how much greater the mean income is in Los Angeles compared to New York. Construct a *90% confidence interval* for this difference.

2. *Does the above data provide evidence* that the mean household income for the customers of the Los Angeles stores is above $100,000? Conduct a *formal hypothesis test* at the .01 level. **(You should ignore the data relating to New York!)**

[15]

3. You wish to estimate the percent of people in the U.S. who have a favorable view of Bill Gates (Mr. Microsoft). You randomly choose 500 people in a nationwide poll. 200 of those surveyed say they have a favorable view.

a. Construct a *99% confidence interval* for the percent of the entire U.S. population that view Gates favorably.

[15]

b. Suppose you still want to have 99% confidence but desire a margin of error of *at most 3%*. How many people do you need to survey?

[5]

4. You are the manager of an advertising firm. You wish to determine whether your employees are more alert at 9 AM or at 2 PM, as these are the times when you commonly have meetings with clients.

[20]

Describe *in detail* how you might set up and analyze a study to assess this. Assume you have 16 employees who are available to participate in your study and that you have available some tool to assess mental alertness. Your description must include *at least* the following. **(5 points extra credit for extremely thorough answers!)**

> how the alertness assessment tool will be utilized
> how the data is to be collected
> what values will be calculated from the sample data
> the statistical technique you will use to reach your conclusion

5. For Question #1 (first page) explain whether the following observations would significantly affect your conclusions. **Justify your answers!**

[8]

a. Suppose the Los Angeles data has one extremely high outlier.

b. Suppose both sets of data are extremely skewed to the right.

6. In assessing cause-and-effect (e.g., effectiveness of a medical treatment) it would be desirable to use a matched-pairs experiment because it is more powerful.

[7]

a. Explain what "more powerful" means in this context.

b. However, we would generally want to use a randomized-comparative experiment to eliminate the effects of _____ _____.

c. This is particularly important if the experiment is conducted over a very _____ _____.

7. The probability that you correctly conclude the alternative hypothesis (that is, when it is actually true) is called the _____ of the test.

[6]

If concluding that the alternative hypothesis is true will lead you to make a "costly" decision (e.g., you will implement an expensive program) then you should use a [smaller / larger] significance level.

Incorrectly concluding the alternative hypothesis is called a _____ _____ .

8. Suppose in Question #2 (first page) someone "magically" told you that the standard deviation for *all* Los Angeles customers was 12. Would the resulting confidence interval be more accurate or less accurate? **Explain your answer fully.**

[5]