MAT212 Final (Spring 2004)

Including new material on sections 13.6 and Linear Regression

Name:

Directions:

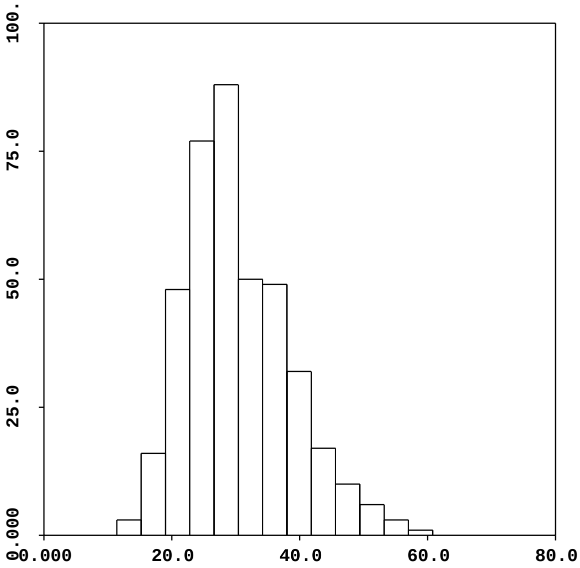
- Points for each problem are in parentheses. All answers to be graded must be on this test. Show all work to receive any credit.
- Table values (normal and t) are attached to your test, as are certain Minitab results relevant to particular problems. When using the tables, specify how you obtained a result.
- Remember to hand in your formula sheet with your exam. Good luck!

Problem 1. You are assured by your staff that individual phone bills are normally distributed with a mean of \$30 and a standard deviation of \$8. Supposing this is true:

1. What percentage of bills are between 20 and 45 dollars? (10 points)

2. You are considering sending a special offer for an additional service to those spending in the top 10% on their phone bills. At what dollar amount will you begin sending out the special offer? (10 points)

Problem 2. You ask your staff for a plot of the phone bill data. They provide you with the following histogram:



1. In the space above, characterize the distribution based on this histogram of 400 randomly chosen bills. (12)

2. What do you make of the claim that the phone bills are normally distributed? (4)

3. Which is larger, the mean or the median (explain!)? (4)

4. What evidence supports the claim that the standard deviation is indeed (or is not) 8? (4)

Problem 3. Suppose that the measured standard deviation of the sample of 400 bills is 7.84, and the mean is 30.13.

- Assume that the population standard deviation σ is indeed \$8. Find a 90% confidence interval for the mean bill amount. (10)

- Assume that σ is unknown. Find a 90% confidence interval for the mean bill. (10)

Problem 4. You decide to study the impact of your special offer on the phone customers by doing an experiment: you send the offer to 40 customers, and you don't extend the offer to 40 customers with similar bills in the top 8%. After several months you compare bills:

	N	Mean	StDev
offer	40	48.16	4.30
no offer	40	46.91	4.69

1. Test to determine whether the offer is increasing the phone bills. (18)

2. State when the procedure used in part A is valid. (5)

Problem 5. A golf league's mean score is $\mu = 84$. Your team of 25 golfers has a mean score of 78.2, with a sample standard deviation of 7.

1. Is there evidence that your team has a better mean than the population? Test using a 5% significance level. (16 points)

2. Under what conditions will your test be valid? (4 points)

Problem 6. Comparison of two groups of nurses (with different specialties) on the same 100 question skills test revealed apparent differences between their scores, as measured by percent correct. Each group had 30 members, and the average number correct was 64 for Group A and 86 for Group B. Is there evidence of a difference in the proportions of correct answers between the two groups? Test using a .01 level of significance. (20)

Problem 7. You consider adding long-term care insurance coverage to your employees, but note that it would require 40% participation in order to be worthwhile (any less and you would be subsidizing the program at a loss). A random sample of 50 employees shows that 24 are willing to participate.

1. Should you implement this program? Test using $\alpha = .05$. (15)

2. Describe a Type I Error and its consequences in terms of this problem. Would you want a large or small value of α? Explain. (6)

3. Describe a Type II Error and its consequences in terms of this problem. Would you want a large or small value of β? Explain. (6)

Problem 8. Short answers:

1. What is the difference between a point estimator and an interval estimator? (4)

2. You're doing a one-tailed test for a negative difference in means, and get a mean difference value of 4: what's your next step? (4)

3. Carefully draw a normal distribution, and then indicate roughly what an $\alpha = .05$ rejection region would look like for a two-tailed test. (4)

4. Describe when we need to use a $t-$distribution rather than a normal distribution. (4)

5. How often will you encounter processes with known parameter values? (4)

Problem 9. (Refer to the Minitab output) The heat input in MMBTU (trillions of BTUs) and the CO_2 output (in tons) for most electricity generation plants in California for 1997 is obtained. By comparing heat input to CO_2 output, one can get a measure of the "pollution efficiency" of each plant, as seen in the Minitab output.

1. Give the Least Squares prediction equation and interpret the coefficients in terms of this problem. (4)

2. Is there evidence that the CO_2 output is linearly related to the heat input? Use $\alpha = .05$. (8)

3. Fully describe the strength of the linear relationship. (4)

4. What do you find striking about this data, and how might you proceed next? (4)

5. If appropriate, use 95% confidence to predict the pollution level for a heat input of 4487554. Comments? Why do we focus on that level? (4)