Exam #2

Math 112-R

Thursday, March 31, 2005

For full credit show all work. When in doubt, explain your reasoning. Show two digits after the decimal point when rounding your answers.

1. Give an example that we have not discussed in class of two variables that are strongly associated but where neither variable determines the other.
2. Describe the connection between the computation of the correlation coefficient and z-scores.
3. Give two reasons why samples are used more often than populations.
4. Given the following information, construct a linear regression model for use in predicting a person's height (in inches) given their weight (in pounds).

   For a sample of 100 students, their average weight was 140 pounds, their average height was 66 inches, the standard deviation of weights was 7 pounds, the standard deviation of heights was 3 inches, and the correlation coefficient between weights and heights was 0.7.
5. Why can we test for causality in a controlled experiment when we can't in an observational study?
6. Use the following sample data to find the conditional distribution for sophomore students with respect to commuting status.

	Freshmen	Sophomores	Juniors	Seniors
Commuters	35	100	300	520
Non-commuters	690	550	400	200

7. Use the data from the previous problem to find the marginal distribution of classes, i.e., freshmen, sophomores, juniors, and seniors.
8. Describe how simulations are used to estimate probabilities, and give an example that we have used in class.
9. Describe a control group in an experiment and why it is used.
10. What does the proportion of variability measure, and how is it computed?
11. What is the difference between a blind and a double-blind experiment?
12. Construct by hand a scatter-plot for the following data:

x	2	4	6	8	10	12	14	16
y	9	12	37	20	35	28	33	36

13. Describe the biases involved in the *Literary Digest* survey that predicted that Alf Landon would win the 1936 presidential election.
14. Explain the difference between explanatory and response variables.
15. When working with regression models, explain the difference between y and $\hat{y}$.