

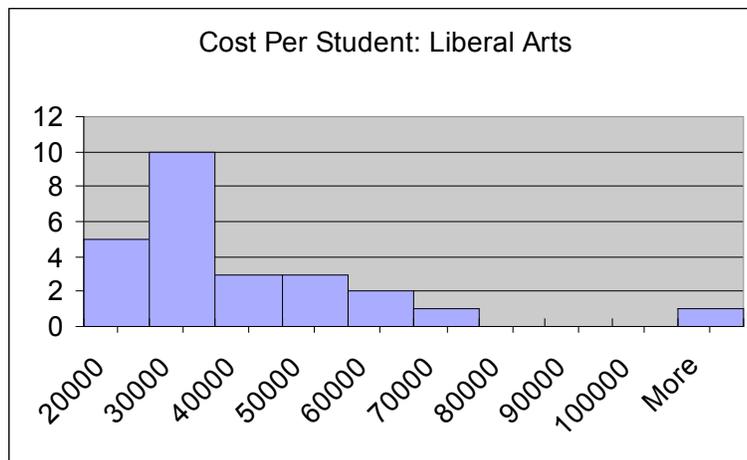
## Midterm 03

### FREE RESPONSE QUESTIONS

1. Explain why it is so difficult to measure the amount of stress in someone's life from a statistical standpoint.
2. Give two major advantages to using a (well designed and conducted) sample survey rather than a census.
3. What is wrong with using a list of registered voters to predict election outcomes?
4. What is the main purpose of using random assignment in experiments?
5. Name two ways in which an experimenter can bias the results in a poorly designed experiment.

Narrative: School costs

Suppose a random sample of liberal arts schools was taken, and the average cost per student was measured for each school. The data are pictured in the histogram below:



6. {School cost narrative} Describe two important characteristics of this data set (for example, the shape) using words that a parent exploring the cost of liberal arts colleges would find useful.

Narrative: Entrance exam

Suppose a certain college has its own entrance exam, and scores on this exam follow a normal distribution with mean 150 and standard deviation 20.

7. {Entrance exam narrative} Bob's score on this exam was 130. What is Bob's standard score, and what does this mean in terms of where he stands in this population?
8. Determine whether or not the following statement could be statistically correct. If not, explain why not. "The correlation between tree diameter and weight of fruit harvested was found to be 2.3."

9. Suppose the correlation between height and weight for all the students in your class is found to be .20. You think that you must be overlooking something, because the relationship should be stronger than that. What additional variable may be masking an underlying relationship here?
10. Name three (of the seven) reasons two variables could be related.

### MULTIPLE-CHOICE QUESTIONS

1. Why is it important to know about any differences in two groups being compared, in addition to the factor of interest?
- If the groups are too different, it becomes difficult to see the effects that are due only to the factor of interest.
  - If the groups are too similar, then any difference found in the factor must be simply due to chance.
  - Unless there is a big difference in the results, it doesn't matter whether or not the groups themselves are different.
  - None of the above.
2. Which of the following is not an example of a discrete variable?
- The number of coins in a person's pocket right now.
  - Your exact age.
  - The number of heads that could appear on two flips of a fair coin.
  - All of the above are examples of a discrete variable.
3. Which of the following describes an experiment?
- Create differences in the explanatory variable and then examine the results.
  - Observe differences in the explanatory variable and then notice whether these are related to differences in the response variable.
  - Both a) and b) are experiments.
  - Neither a) nor b) are experiments
4. In order to establish a connection between an explanatory and response variable in an observational study, what is needed?
- Statistical methods that examine the connection.
  - A comparison to determine if changes in the explanatory variable are related to changes in the response variable.
  - Either a) or b)
  - You can never establish a connection between an explanatory and a response variable without doing a randomized experiment.
5. Suppose you look at two boxplots comparing the weights of male cats vs. female cats, and you find that the box for the males is much wider than the box for the females. What does this mean about the data sets?
- Male cats weigh more than female cats overall.
  - Male cats have more variability in their weights than female cats.
  - Weights of male cats are more skewed than for female cats.
  - None of the above.
6. What type of statistical error is being made in the following statement? "If this uphill linear trend continues, 50 years from now, one out of every three of us will be an Elvis impersonator."
- Extrapolation
  - Exaggeration

- c. Overprediction
- d. Expectation

Narrative: Genetic defect

Suppose a population contains 100,000 individuals, of which 500 are carriers of a certain genetic defect.

7. Assuming there is a statistical relationship between height and weight for adult females, which of the following statements is true?
- a. If we knew a woman's height, we could predict her weight.
  - b. If we knew a woman's height, we could determine the exact weight for all women with that same height.
  - c. If we knew a woman's height, we could predict the average weight for all women with that same height.
  - d. All of the above are true.
8. {Genetic defect narrative} Which of the following is an appropriate way to express this result in terms of odds?
- a. The odds of carrying the genetic defect are 5 to 995.
  - b. The odds of carrying the genetic defect are 500 out of 100,000.
  - c. The odds of carrying the genetic defect are 500 to 1.
  - d. None of the above