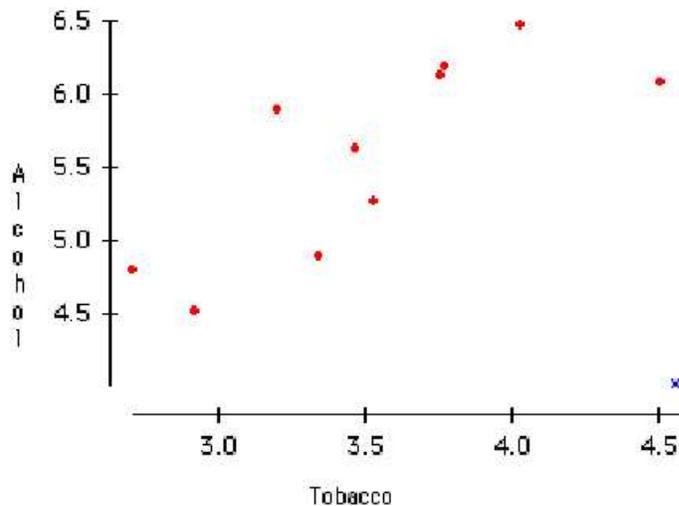


Stat 145: Exam 2 Review Questions

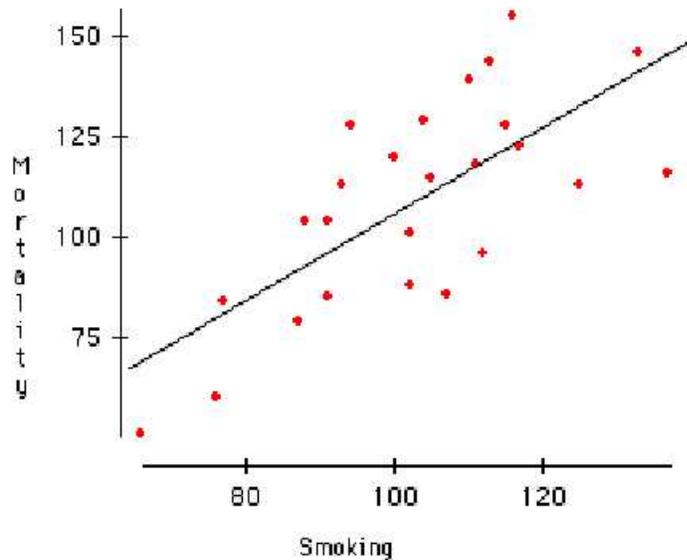
Note: This is not intended to be a preview of the actual exam. Rather, it is meant to give you an idea of the types of questions that will be asked. There are concepts in these review problems that will not appear on the actual exam, just as there will be concepts on the actual exam that are not covered in these review problems.

1. A scatterplot of average weekly household spending (in British pounds) on alcohol vs. average weekly household spending (in British pounds) on tobacco in the 11 regions of Great Britain appears below.



- (a) Will the correlation r between the two variables move closer to 0 or 1 if the the observation in the lower right-hand corner of the plot (marked with an “x”) is removed? Explain your answer.
 - (b) If average weekly household spending on alcohol and average weekly household spending on tobacco were converted from British pounds to U.S. dollars, would the value of the correlation r change? Why or why not?
2. In the early part of the 20th century it was noticed that, when viewed over time, the number of crimes increased with membership in the Church of England. What is the most likely explanation for this?

3. Government statisticians in England conducted a study of the relationship between smoking and lung cancer in 25 occupational groups. “Smoking” is a ratio that is 100 if men in an occupation are exactly average in their smoking, below 100 if they smoke less than average, and above 100 if they smoke more than average. “Mortality” is also measured relative to the entire population of men of the same ages as those studied, and is greater or less than 100 when there are more or fewer deaths from lung cancer, respectively, than would be expected based on the experience of all English men. A scatterplot of the two variables appears below.



- (a) Describe the direction and strength of the linear association between relative mortality rate and relative smoking rate.
- (b) Use the summary statistics below to find the equation of the least-squares regression line for predicting relative mortality rate from relative smoking rate.

Variable	Mean	Std Dev
Smoking	103	17.2
Mortality	109	26.1

The correlation is $r = 0.716$.

- (c) Use the equation above to predict the relative mortality rate for an occupational group with a relative smoking rate of 110.

4. Specify whether the following statements are true or false. If a statement is false, modify it to make it a true statement.
- (a) Points that lie above the least-squares regression line have negative residuals.
 - (b) The sum of the least-squares residuals is always zero.
 - (c) A regression line should not be used to make predictions for x values that lie far outside the range of the data you collected.
 - (d) R^2 can take on any value between -1 and 1 .
5. A publishing company is interested in the percentage of U.S. adults that have read at least one novel in the last month. A random sample of size 800 is taken from the subscription list of a literary magazine published by the company. It turns out that 74% of the 800 people in the sample have read a novel in the past month.
- (a) Why are the sample results biased?
 - (b) As an estimate of the percentage of U.S. adults who have read at least one novel in the last year, is 74% likely too high or too low? Explain your answer.
6. A report in *Circulation*, the journal of the American Heart Association, examined whether taking aspirin helps with the recovery from heart attacks. The subjects were the 539 patients who were admitted with heart attack or stroke symptoms to a University Hospital in Barcelona, Spain, between October 1991 and March 1993. The patients were asked if they had taken aspirin during the prior week: 214 said yes and 325 said no. It was found that the group who had taken aspirin had a much lower amount of permanent damage from their heart attacks than the group who had not taken aspirin. Is this an observational study or an experiment? Explain your answer.
7. Sickle cell disease is an inherited disorder of the red blood cells that can cause severe pain and many complications. Fifty subjects with sickle cell disease volunteer to participate in a study of the drug hydroxurea.
- (a) Can this drug reduce the severe pain caused by the disease? Design a study to answer this question.
 - (b) Use the excerpt from the table of random digits below to choose the first three subjects for one of your groups.

Line								
140	12975	13258	13048	45144	72321	81940	00360	02428

8. A group of children was shown an episode of Sesame Street at 8:00 AM, while another group of children was shown an episode of Blue's Clues at 11:00 AM. During both shows, crackers were available in a bowl, and the investigators compared the number of crackers eaten by the two groups. It was found that more crackers were eaten during Blue's Clues than during Sesame Street. The investigators concluded that the different shows had an effect on appetite.
- (a) What is the explanatory variable?
 - (b) What is the response variable?
 - (c) What is a lurking variable in this experiment?
9. The progress of a type of cancer differs in men and women. A clinical experiment to compare three therapies for this cancer is performed by first grouping the subjects by gender and then randomly assigning the subjects within each group to one of the three therapies. What type of design is this?
10. The following question was asked of 100 people outside Dane Smith Hall:
- Given Howard Dean's penchant for losing his temper, would you vote for him for President?

Of the 48 respondents, 67% said no. The person conducting the survey concluded "In a random sample of UNM students, 67% said they would not vote for Howard Dean for President." List three reasons why the conclusion drawn is misleading.