Exam 2 (100 pts.)
Stat 145 · Spring 2013

Name: ____________________________________ Section: __________________

For credit show work or explain all answers.

1. A researcher designs a one-year study to examine the effect of physical activity and nutrition classes on weight loss in obese adolescents aged 14-17. The study has three levels of physical activity class: none (0), 30 minutes daily (30), and 60 minutes daily (60); and two levels of nutrition class: none (N) and once a week (W). Sixteen boys are randomly assigned to each combination of physical activity and nutrition classes. Each study participant's body fat percentage is measured before and after receiving treatment.

(a) (5 pts.) How many factors are in this study? Specify each one.

(b) (5 pts.) How many treatments are in this study? Specify each one.

(c) (5 pts.) How many adolescents, aged 14-17, does this experiment require altogether?

(d) (4 pts.) What is the response variable?

2. (5 pts.) Researchers followed 1,523 young adults for 8-months; they recorded the amount of time each study participant spent playing video or computer games, and the participant's feelings of depression. They concluded that young adults who spend a lot of time playing video or computer games have a greater risk of depression. Did these researchers conduct an observational study or an experiment? **Explain** your answer.

3. *(Round final answers to two decimal places.)*

In 10 states, data was collected on each state's gas tax in cents per gallon and fuel use in gallons per registered vehicle. In parts (a) through (c), use the summary statistics below to find the equation of the least-squares regression line for predicting a state's fuel use from its gas tax.

Variable	**Mean**	**Std. Dev.**
fuel use (gallons per registered vehicle)	796.5	187.9
gas tax (cents per gallon)	27.02	3.77
The correlation is $r = -0.32$		

(a) (5 pts.) Calculate the slope.

(b) (5 pts.) Calculate the intercept.

(c) (4 pts.) Write the equation for the least squares regression line.

(d) (2 pts.) Calculate r^2.

(e) (5 pts.) Explain what r^2 means in terms of the explanatory and response variables.

.

4. A least squares regression line to predict a student's Stat145 test score (from 0-to-100) from the number of hours studied was determined from a class of 55 Stat145 students: $\hat{y} = 46.2 + 2.71x$. One student in the class studied for 16 hours and scored 87 on the exam.

(a) (5 pts.) What is the predicted value of this student's Stat145 exam score?

(b) (5 pts.) What is the residual for this student?

(c) (5 pts.) Explain what the **slope** of this least squares regression line tells us with regard to the explanatory and response variables. Be specific with regard to the value of the slope.

Multiple Choice (4 pts. each). Select the BEST answer:

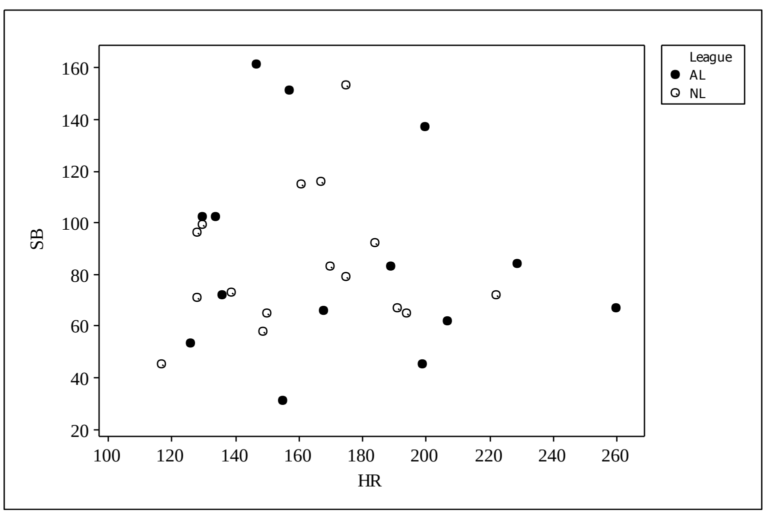
Use the following to answer questions 5-6:

The U.S. census bureau mails survey forms to 250,000 U.S. households asking questions about the people living in the household and about such things as motor vehicles and housing costs. Telephone calls are made to households that don't return the form. Responses were obtained from 240,000 of the households contacted.

5. The sample is
   A) the 240,000 households that respond.
   B) the 250,000 households initially contacted.
   C) the 10,000 households that did not respond.
   D) all U.S. households.

6. The population of interest is
   A) the 250,000 households initially contacted.
   B) the 240,000 households that respond.
   C) only U.S. households with phones.
   D) all U.S. households.

7. Below is a scatterplot of number of home runs versus number of stolen bases for major league teams in 2009. American League teams are represented by filled circles and National League teams by open circles.



We conclude that

A) there is a strong positive association for American League teams but a negative association for National League teams.
B) there is a strong negative association for American League teams but a positive association for National League teams.
C) all American League teams hit more home runs and stole more bases than did National League teams.
D) there is a weak association for both leagues.

8. The magazine High Times has a website that once asked visitors whether recreational marijuana use should be legal. This is an example of
   A) a survey with little bias since someone who responded would know his or her opinion.
   B) a survey with little bias because a large SRS was used.
   C) voluntary response sampling.
   D) All of the above

9. Will a fluoride mouthwash used after brushing reduce cavities? Twenty sets of twins were used to investigate this question. One member of each set of twins used the mouthwash after each brushing, the other did not. After six months, the difference in the number of cavities of those using the mouthwash was compared with the number of those who did not use the mouthwash. This experiment uses
   A) random placeboes.
   B) a matched pairs design.
   C) double replication.
   D) double-blinding.

Use the following to answer questions 10 and 11:

Import customs officials randomly select crates of cargo for close inspection. Suppose there are nine crates of cargo from the following companies, and customs officials randomly select four for close inspection.

1. Ravenburg	4. Dallhoise	7. Cherryport
2. Corsair	5. Baggate	8. Foxwood
3. Sapphire	6. Strommond	9. Bamboro

Use the numerical labels attached to the previous names and the following list of random digits to select an SRS of size four. Read the list of random digits from left to right, starting at the beginning of the list.

74880 32009 45287 71753 98236 66419 84533 11793 20495 05907 11384

10. The simple random sample is
    A) 7488.
    B) 7483.
    C) Cherryport, Dallhoise, Foxwood, and Foxwood.
    D) Cherryport, Dallhoise, Foxwood, and Sapphire.

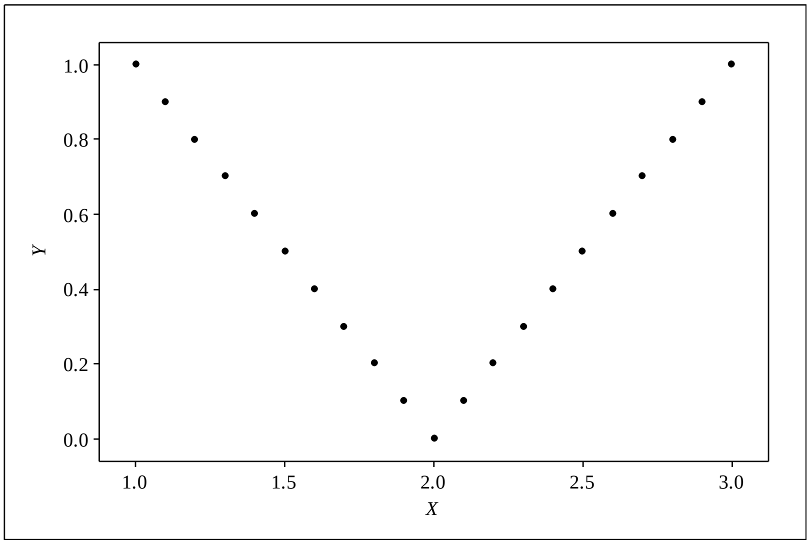
11. Which of the following statements is true?
    A) If we used another list of random digits to select the sample, we will certainly get a different sample.
    B) If we used another list of random digits to select the sample, the list we got above would be less likely to occur than any other particular list of crates.
    C) If we used another list of random digits to select the sample, the company Cherryport would be unlikely to be selected because it was selected above.
    D) If we used another list of random digits to select the sample, the result obtained with the list actually used would be just as likely to be selected as any other set of four names.

12. A researcher wants to determine whether the rate of water flow (in liters per second) over an experimental soil bed can be used to predict the amount of soil washed away (in kilograms). The researcher measures the amount of soil washed away for various flow rates, and from these data calculates the least-squares regression line to be amount of eroded soil = 0.4 + 1.3 × (flow rate)

    The correlation between amount of eroded soil and flow rate would be

    A) 1/1.3
    B) 1.3
    C) either positive or negative. It is impossible to say anything about the correlation from the information given.
    D) positive, but we cannot say what the exact value is.

13. Consider the following scatterplot of two variables *x* and *y*.



    We may conclude

    A) the correlation between *x* and *y* must be close to 1 since there is nearly a perfect relation between them.
    B) the correlation between *x* and *y* must be close to –1 since there is nearly a perfect relation between them but it is not a straight line relation.
    C) the correlation between *x* and *y* could be any number between –1 and +1. Without knowing the actual values we can say nothing more.
    D) the correlation between *x* and *y* is close to 0 because although there is a strong relationship between these variables, it isn't a linear relationship.

14. What can be said of the correlation between the brand of an automobile and its quality?

    A) The correlation is negative because smaller cars tend to have higher quality and larger cars tend to have lower quality.
    B) Correlation makes no sense here because “brand” is a categorical variable.
    C) If correlation is negative, an arithmetic mistake was made. Correlation must be positive.
    D) The correlation is positive because better brands have higher quality.