

## Exam 2: Nov. 6, 1997, Quant II Form A

The 36 exam questions and answer sheet are both to be turned in to your Discussion Section instructor at the end of the exam. Code your name and ID number on the answer sheet. **Code your Section number under OPTIONAL CODES in positions L M N.** Please notice the Defective Question report, the Normal Distribution table, and the Formula Sheet at the end of the exam.

1. A student collects some data on college students. For 100 students, he obtains two variables: College GPA and Average Weekly Hours Watching Television. He finds the correlation coefficient between the two variables to be  $+0.95$ . We can therefore conclude that watching television must cause students to get better grades.

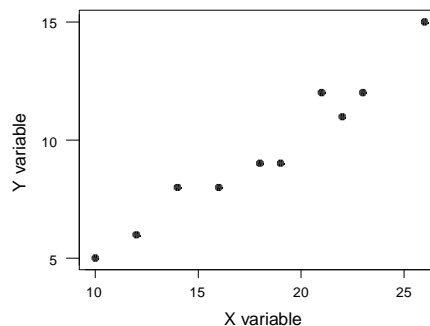
A) True B) False

2. A basketball coach has an idea that training in ballet might help her players jump higher. To test this idea she divides her team randomly into two equal groups. One group is required to take six weeks of ballet training before the basketball season begins. The other group receives no ballet training. This study would be classified as

A) a probability survey  
B) an observational study  
C) an experiment  
D) an observational survey  
E) None of the above.

3. Consider the scatterplot of two variables shown at the right. From this plot, which of the following is the most reasonable estimate of the correlation coefficient between the two variables?

A)  $-1.00$   
B)  $-0.80$   
C)  $+0.32$   
D)  $+0.97$   
E)  $+1.00$



4. If a sequence of longitudinal data is random we expect the lag one autocorrelation coefficient to be near zero.

A) True B) False

5. Twenty pairs of data produced a least squares regression line with equation

$$\hat{y} = 0.0263 + 0.547x \quad \text{Line 1}$$

The boss prefers simplicity so she wants to use the line with equation

$$\hat{y} = 0.03 + 0.5x \quad \text{Line 2}$$

Which one of the following statements is true?

- A) Line 1 will have a smaller value for the residual standard deviation.
  - B) Line 2 will have a smaller value for the residual standard deviation.
  - C) Line 1 will have a larger value for the total sum of squares.
  - D) Line 2 will have a larger value the total sum of squares.
  - E) Without having the data we cannot say which line will have smaller residual standard deviation nor larger total sum of squares.
6. What can we say about the correlation coefficient for the data at the right?

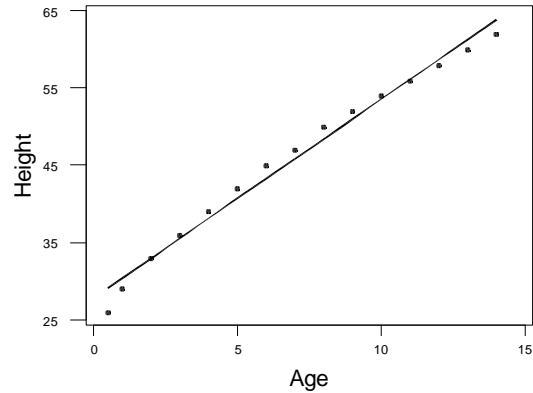
**(Do not calculate!)**

x	2.02	2.95	4.89	3.21	3.92
y	4.11	2.10	1.21	2.98	2.07

- A) It is  $-1$ .
  - B) It is strongly negative.
  - C) It is near zero.
  - D) It is strongly positive.
  - E) It is  $+1$ .
7. A placebo is
- A) a dummy treatment
  - B) the control group
  - C) the experimental group
  - D) an aspirin treatment
  - E) the treatment the researcher hopes will prove effective.
8. Researchers have observed that drinking red wine,  $x$ , seems to lead to fewer men having heart attacks,  $y$ . More recently, others have noted that drinking red wine leads to headaches and people with headaches tend to take aspirin. Furthermore, aspirin is known to reduce the chances of having heart attacks. Given these facts, the relationship between  $x$  and  $y$  would be best described as being due to
- A) cause-and-effect
  - B) mechanism
  - C) consistency
  - D) strong correlation
  - E) a lurking variable

9. One of the following is the equation of the least squares line for the data displayed at the right. Which one is it?

- A) Height = 10.3 + 2.75 Age
- B) Height = 28.8 + 4.01 Age
- C) Height = 37.8 + 2.37 Age
- D) Height = 20.5 + 1.22 Age
- E) Height = 27.8 + 2.57 Age



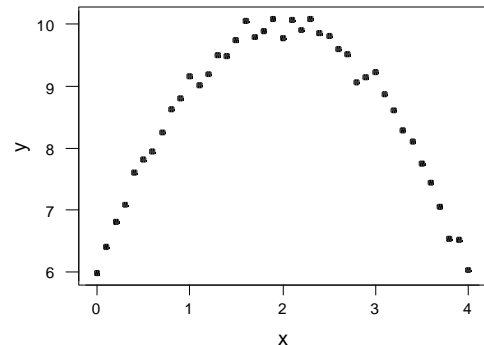
10. A portion of an ANOVA (analysis-of-variance) table from a multiple regression calculation is shown below. What is the value of  $s$  (the residual standard deviation) for this regression?

Source	SS	df
Regression	1501.3	
Error	591.1	56
Total		59

- A) 3.249
- B) 10.555
- C) 22.370
- D) 500.433
- E) None of the above.

11. For the scatterplot shown at the right, which of the following values is closest to the correlation coefficient?

- A) -0.9
- B) 0
- C) 0.4
- D) 0.9
- E) Cannot calculate a correlation coefficient for curved relationships.



12. Your friend tosses a pair of dice where one die is red and the other is green. She reports to you that she got a 6 total. Given this information, what is the chance that she rolled a 1 on the green die?
- A)  $1/6$   
 B)  $2/6$   
 C)  $3/6$   
 D)  $1/5$   
 E)  $2/5$
13. A set of data pairs has the following summary statistics:  $\bar{x} = 40$ ,  $\bar{y} = 100$ ,  $s_x = 3$ ,  $s_y = 12$ , and  $r = 0.5$ . What is the equation for the least squares regression line in *original terms* of  $x$  and  $y$ ?
- A)  $\hat{y} = 0.5x$   
 B)  $\hat{y} = 2x$   
 C)  $\hat{y} = 20 + 2x$   
 D)  $\hat{y} = 100 + 0.5x$   
 E) None of the above.
14. An interested company worker wants to model the relationship between years of job experience and salary. She uses least-squares to obtain the following fitted equation, where  $y$  is the annual salary (in dollars) and  $x$  is the number of years of experience:  $y = 20,000 + 2,500x$ . Sarah, an economist for the company, has 14 years experience and is making \$54,000 a year. What is the residual associated with Sarah?
- A)  $-2,000$   
 B)  $-1,000$   
 C)  $+1,000$   
 D)  $+2,000$   
 E)  $+2,500$
15. The table at the right gives the joint distribution of two variables  $x$  and  $y$ . What is the expected value of  $y$ ?

$y$	1	1/4	1/4	0
	0	0	1/4	1/4
		0	1	2
		$x$		

16. Suppose you are given the following least-squares regression equation, where  $y$  is the final exam score in a Spanish class,  $x_1$  is the midterm exam score and  $x_2$  is the gender of the student (0 = male, 1 = female):  $y = 0.83 + 1.01x_1 - 3.63x_2$ . What is the predicted final exam score (rounded to the nearest whole point) for a male student who got a 75 on the midterm exam?
- A) 73
  - B) 74
  - C) 75
  - D) 77
  - E) Cannot determine from the information given.
17. you are conducting an experiment on popcorn. The factors of the experiment are Brand and Popper type. You have selected three brands of popcorn and two types of popper. You will measure how many cups of popped popcorn are produced. If you perform two replications of the experiment, how many total batches of popcorn will you have to make?
- A) 2
  - B) 3
  - C) 6
  - D) 7
  - E) 12
18. In the Physician's Health Study, a coin was flipped to decide which doctors took the aspirin and which took the placebo. Using the coin flip to decided which subjects get which treatment is an example of:
- A) randomization
  - B) replication
  - C) blinding
  - D) stratification
  - E) placebo
19. Two variables  $x$  and  $y$  have an approximate straight-line relationship between them with a correlation coefficient of 0.7. If  $x$  were to increase by amount  $s_x$  how much would we predict that  $y$  would increase?
- A)  $s_x$
  - B)  $s_y$
  - C)  $0.7s_x$
  - D)  $0.7s_y$
  - E)  $0.7s_y/s_x$

20. Find the lag one autocorrelation coefficient for the sequence: 8, 6, 4, 3, 4.
- A) 0.125
  - B) 0.250
  - C) 0.375
  - D) 0.500
  - E) 0.750
21. Suppose  $y$  is a response variable of annual salary,  $x$  is a continuous predictor variable of years of experience, and  $z$  is a binary indicator variable indicating gender ( $z=1$  means female,  $z=0$  means male). Consider the regression model with equation:  $\hat{y} = b_0 + b_1x + b_2z$ . What coefficient (or combination of coefficients) represents the predicted *starting* annual salary (experience=0) for a male?
- A)  $b_0$
  - B)  $b_1$
  - C)  $b_2$
  - D)  $b_0+b_1$
  - E)  $b_0+b_2$
22. In the multiple regression example given in the British Steel video, they used two continuous predictor variables (amount of carbon in the steel plate and amount of carbon in the welding wires) to predict the amount of carbon in the final weld. Their regression model is best described as
- A) one line
  - B) two parallel lines
  - C) one quadratic curve
  - D) a plane
  - E) a quadratic surface
23. You flip a coin twice and each time record whether heads or tails came up. What is the conditional probability that the first flip was a head, given that at least one of the flips was a head?
- A) 1/4
  - B) 1/3
  - C) 1/2
  - D) 2/3
  - E) 3/4

24. The table below gives the breakdown of shoe sale transactions over a long period for a large shoe chain. An auditor selects one transaction at random from all of the transactions.

		Brand Purchased		
		Reebok	Nike	Other
Type of Purchase	Cash	5%	10%	5%
	Credit Card	40%	15%	25%

What is the chance that a Cash transaction is selected?

- A) 20%
- B) 35%
- C) 55%
- D) 80%
- E) None of the above.

25. Refer to the previous problem. Given that a Credit Card transaction is selected, what is the chance that it is a Reebok transaction?

- A) 20%
- B) 30%
- C) 40%
- D) 50%
- E) None of the above.

26. The table at the right gives the joint distribution of two variables  $x$  and  $y$ . What is the covariance between  $x$  and  $y$ ?

$y$	1	$1/2$	0
	0	0	$1/2$
		0	1
		$x$	

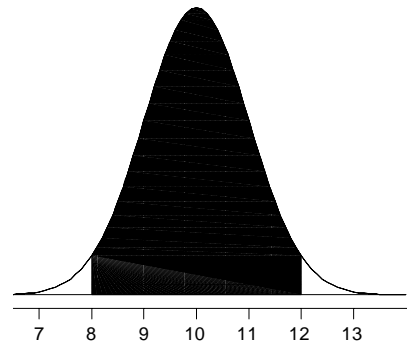
- A)  $-1/2$
- B)  $-1/4$
- C) 0
- D)  $1/4$
- E)  $1/2$

27. In the general U.S. adult population, Stanford-Binet IQ scores are approximately normally distributed with a mean of 100. Just 2.5% of U.S. adults score above 132 on this IQ test. What is the standard deviation of the IQ scores?

- A) 1
- B) 2.5
- C) 16
- D) 32
- E) 116

28. The graph at the right shows a normal curve with mean 10 and standard deviation 1. What is the value of the shaded area?

- A) 0.05
- B) 0.32
- C) 0.68
- D) 0.95
- E) 0.997



29. The 381 students in *Quant II* took a 35 question exam. The average score was 28.4 and the standard deviation was 2.6. Class policy is that you must get 88% or more correct to get an A. Here that means you must get 31 questions or more correct. The distribution of scores is approximately a normal curve. About how many students got A's on this test?

- A) 20
- B) 29
- C) 37
- D) 48
- E) 61

30. A balanced coin is flipped once. Let  $y$  be the number of tails obtained. What is the variance of  $y$ ?

- A)  $1/8$
- B)  $1/4$
- C)  $1/2$
- D) 1
- E) None of the above.

31. A manufacturer produces bags of candy with a printed label weight of 20 ounces. The actual weights vary according to a normal distribution with mean 20.49 ounces and standard deviation 0.21 ounces. What percentage of the bags weigh less than the label weight?

- A) 0%
- B) 1%
- C) 2.5%
- D) 5%
- E) 16%

32. A manufacturer can adjust the mean,  $\mu$ , of a process variable,  $y$ , by changing a dial on the machinery. The  $y$  has a normal distribution with standard deviation 3.9 inches. What should  $\mu$  be set to if we want 90% of the values of  $y$  to be above 25 inches? (Round to the nearest tenth of an inch.)
- A) 30.0
  - B) 31.4
  - C) 32.9
  - D) 34.1
  - E) 35.0

33. Which one of the following statements is **true**?

- A) Consider an experiment involving human subjects. If the experimenter is aware of what treatment a subject is getting, yet the subject is unaware, then this is called double blinding.
- B) Stratification is seldom used to make comparisons.
- C) A probability survey is an example of a longitudinal study.
- D) In an experiment with 2 factors, each with 3 levels, there are a total of 9 treatments.
- E) Data production quality refers to the question “Are the data relevant to the problem we wish to solve?”

34. A company fit a regression model to their salary data with the resulting regression equation

$$\widehat{\text{salary}} = 10.2 + 0.24\text{years} - 0.04\text{gender}$$

where salary is annual salary in thousands of dollars, years is years of experience, and gender is 1 for females and 0 for males. If Judy has 2 years more experience than Joe, by how much would the model predict her salary to differ from Joe’s?

- A) Judy would get \$440 more than Joe.
- B) Judy would get \$480 more than Joe.
- C) Joe would get \$440 more than Judy.
- D) Joe would get \$480 more than Judy.
- E) Cannot be determined from the information given.

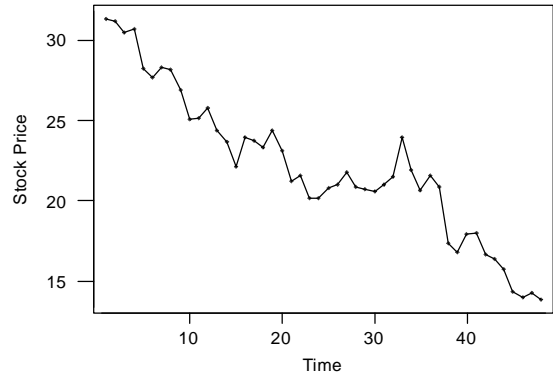
35. Which one of the following statements is **false**?

- A) A residual plot that shows a random scatter indicates adequacy of the model that was fit.
- B) In least-squares regression, the residuals always sum to zero.
- C) Outliers affect the least-squares regression line.

- D) The least-squares regression line is found by minimizing the sum of squared residuals.
- E) After fitting a linear regression model, we can see how the predictor variable causes the response variable to change, i.e. there is a cause-and-effect relationship between the predictor and the response.

36. The graph at the right shows the movements of a certain stock price over 48 days. Which of the following gives the value of the lag one autocorrelation coefficient?

- A)  $-0.841$
- B)  $-0.494$
- C)  $+0.271$
- D)  $+0.893$
- E)  $+1.000$



# Defective Question Report

Name: \_\_\_\_\_

Section: \_\_\_\_\_

ID: \_\_\_\_\_

If you believe that a test question is defective in some way, please list your complaint here. All complaints will be considered in our interpretation of the test results.

To correctly identify the test question we must know which **form** of the test you have taken. *We also must know how you answered the question.*

**Remove this last page from the exam questions and turn it in to one of the instructors in the course.**

**Form (circle one): A B C D E**

Question number: \_\_\_\_\_ Your choice of answer on answer sheet: \_\_\_\_\_

Your complaint:

Question number: \_\_\_\_\_ Your choice of answer on answer sheet: \_\_\_\_\_

Your complaint:

Question number: \_\_\_\_\_ Your choice of answer on answer sheet: \_\_\_\_\_

Your complaint: