Review Materials for Exam 1 (10/19/01)

Exam coverage: Chapter 1, Chapter 2, Chapter 3.1-3.5, Chapter 4 (but not 4.4), Chapter 5, and Chapter 13

PLEASE BRING SOME FORM OF PHOTO IDENTIFICATION (e.g. Bruin Card, Drivers License, etc.) TO THE EXAM. ATTENDANCE WILL BE TAKEN. PLEASE REMEMBER TO BRING WRITING INSTRUMENTS AND A CALCULATOR. WE WILL PROVIDED AN EXAM PACKET AND TABLE A-105 FROM THE TEXT.

A ONE SIDED 8.5" x 11" piece of paper with formulas is allowed into the exam. Typed, laser print, cut and paste, handwritten is OK.

Suggested Extra Problems From Your Textbook:

Chapter 3:
Exercise Set A #3, #4, #7
Exercise Set D #2
Review Exercises #1 and #4

Chapter 4:
Exercise Set A: #9
Exercise Set B: #1 - #5
Exercise Set E: #1, #4, #10

Read the technical note on page 74 and Using a Statistical Calculator (if it applies to you)

Chapter 5:
Exercise Set A #1
Exercise Set B #3, #4
Exercise Set C #1
Exercise Set E # 1 and #2
Exercise Set F #1

Page 428. #1, #3 (a & b), #5

Page 568. #4, #6

Chapter 13:
Exercise Set A: #2, #3
Exercise Set B: #1 - #3
Exercise Set C: #2
Review Exercises: 10, 11

WHAT FOLLOWS ARE TWO EXAMS FROM PAST CLASSES

1. In a hypothetical experiment, a new drug was compared with "standard therapy" treatment for patients suffering from inoperable cancer. The result measured was difference in survival time (in months). Which of the following best describes the primary reason to randomize patients into treatment or control groups? (Choose one best answer) (3pts)

(a) to create two groups that are similar at baseline on both known and unknown factors associated with survival time.

(b) prevent bias introduced when the patients know what type of treatment they are receiving

(c.) prevent bias introduced when the investigators know what type of treatment the patients are receiving

(d.) Both b and c

2. A team of researchers reported on the effectiveness of surgically implanted electrodes for reducing facial pain. 34 patients with "chronic medically intractable facial pain" underwent the procedure; 20 of the received implanted electrodes, 14 of them did not. 19 of the cases (56%) were deemed successful, in the sense that "there was a reduction of pain by at least 50% whenever the stimulator was on." Those 19 received permanent implants; 7 of them subsequently developed infections. (3pts)

i) The placebo effect is not likely to be a confounding factor in this design.
ii) The fact that the surgeons knew who had received treatment is likely to be a confounding factor

a) i is true, ii is false.
b) i is false, ii is true.
c) Both i and ii are true.
d) Both i and ii are false.

To determine the prevalence of sexually transmitted diseases (STD) and high risk sexual behavior for STD among adolescent males admitted to a juvenile detention facility, a survey was obtained from interviews. The results are tabled below.

Table 1. Behavioral variables in 966 subjects

Variable	Mean (SD)	Range	Median
Age at first coitus	12.3 (2.0)	6-18	13
No. lifetime partners	13.7 (16.8)	1-100	8
No. partners past 4 months	2.9 (3.4)	0-30	2
No. weeks since last sex	5.8 (15.1)	1-260	2

Note: SD is the abbreviation for standard deviation

3. Which of the descriptive statistics in Table 1 (mean, SD, range, median) is most susceptible to being influenced by a single extreme value? (Choose **one** best answer.) (3 pts)

a. mean
b. SD
c. range
d. median

4. Of the four variables in the table, which has the most symmetrical (normal-like) distribution based on the statistics presented? (Choose **one** best answer.) (3 pts)

a. age at first coitus
b. number of lifetime partners
c. number of partners in the past 4 months
d. number of weeks since last sex
e. none of the above, they are all non-normal based on the statistics given

The next three questions refer to the list {-8, -5, -3, 0,1, 3, 0, 4}.

5. (3pts)The mean is:
   (a) –1.0
   (b) 0
   (c) 0.5
   (d) 1.0
   (e) none of the above.

6. (3pts)The median is:
   (a) –1.0
   (b) 0
   (c) 0.5
   (d) 1.0
   (e) none of the above.

7 (3pts)The inter-quartile range is:
   (a) -4
   (b) 4
   (c) -6
   (d) 6
   (e) none of the above.

8. (3pts) Consider two events, one called A, the other B. Suppose we know that the probability that A occurs is 90%. Suppose we know the probability that B occurs is 60%. Here are two statements:

   i. A and B cannot be mutually exclusive.
   ii. A and B cannot be independent.

a) i is true, ii is false.
b) i is false, ii is true.
c) both i and ii are true.
d) both i and ii are false.
e) There is not enough information to answer this question.

9. (3pts) From long experience, it is known that people default on their credit card (that is, fail to pay their credit card bill) with a probability 0.2 or 20%. Suppose you pick 5 credit card holders at random with replacement. What is the probability that at least one will default on their credit card?
   a) Less than 20.0%
   b) About 20.0%
   c) About 33%
   d) About 67%
   e) Over 68%

10. (3pts) The SD can never be larger than the mean:
   (a) True
   (b) False

The last three questions refer to this statement, but each question is separate (i.e. you can get the first one wrong and that won't affect the others): Corporate securities (or publicly traded stocks) are an investment opportunity for individuals as well as institutions. The 10,000 stocks available for investment to U.S. residents are normally distributed with a mean one-year return of -1% and a standard deviation of 12%. SHOW YOUR WORK FOR FULL CREDIT.

11. What percentage of stocks had one-year returns between -16% and +2%? (5 points)

12. A stock is at 5^{th} percentile (i.e. 5% of the stocks have returns lower than this stock), what is its one-year return? (5 points)

13. In order to meet your retirement goals, you need to buy stocks that have a return of 8% or more. Approximately how many stocks out of the 10,000 qualify? (5 points)

Ten temperature readings from a freezer unit (in degrees Fahrenheit):

-9 -10 +1 +2 -5 0 0 +3 -11 +1

14. What are the average and median of this list? (5 points)

A. average = -2.8 and median = 0
B. average = -2.8 and median = -2.5
C. average = -3.5 and median = 0
D. average = -3.5 and median = -2.5
E. average = 0 and median = -3.5

15. The formula for converting temperature readings into degrees Celsius (a temperature reading on the metric scale) is:

Celsius = 5/9 (Fahrenheit - 32)

Which statement below is correct about a dataset of temperature readings after the change of scale from Fahrenheit to Celsius? (5 points)

A. The mean will change, but the median will remain the same
B. The mean and median will change, but the standard deviation will remain the same.
C. The range will remain the same but the standard deviation will change.
D. The mean and standard deviation will change, but the median will remain the same.
E. All of the above are false.

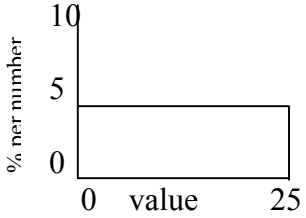
16. A medical observational study investigated the safety of anesthetics used in surgery. Records of over 1,000,000 operations performed in major hospitals showed the following death rates:

Anesthetic	1	2	3
Death Rate	1.7%	1.8%	5.6%

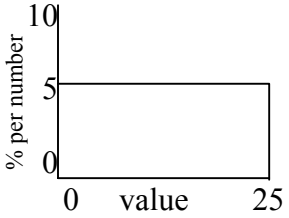
Clearly, anesthetic 3 appears dangerous, but follow-up studies showed that we could not conclusively determine this. What kinds of problems might be present in the original study? (5 points)

A. Confounding Factors
B. Bias
C. Simpson's Paradox
D. A, B, and C are present
E. Only 2 of the 3 (A, B, C) are present

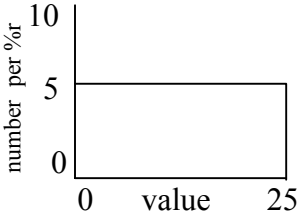
17. Three students sketched the histogram for lottery numbers which are uniformly (or evenly distributed) ranging from 0 to 25, only one histogram is correct, which one? (5 points)



(I)



(II)



(III)

18. Explain the choice you made in problem #4. Why is it right? (be brief your answer should fit in the space below) (5 points)

19. Jack and Jill, brother and sister, both take the SAT. On the mathematics section, Jack scored a 580 and Jill scored a 480. Assume that the math SAT is normally distributed with a mean of 500 and a standard deviation of 100 for all people taking the exam. What percentage of test takers did better than Jill but not as well as Jack? (5 points)

A. About 16%
B. About 37%
C. About 42%
D. About 58%
E. About 73%

20. What is the inter-quartile range for this distribution? Calculate it using the information given in problem #19.