

Chapter 11

PRACTICE EXAMS

11.1 Midterm Exam 1: Fall 2000

Multiple choice: 4 points each

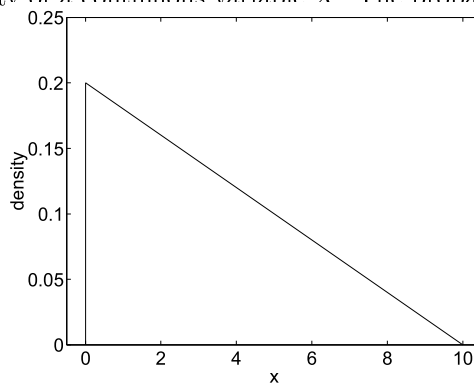
1. The presence of _____ bias is possible in an observational study if some individuals in the sample can not be contacted by the investigator and therefore can not give a response.
 - (a) response
 - (b) nonresponse
 - (c) selection
 - (d) under-coverage
2. Which experimental design strategy is useful (and necessary) for dealing with the effects of unknown extraneous variables (i.e., extraneous variables that the researcher has not measured and may be unaware of)?
 - (a) direct control
 - (b) blocking
 - (c) random assignment
 - (d) replication
3. Which two experimental design strategies are useful for dealing with the effects of a known extraneous variable (i.e., an extraneous variable that the researcher has measured)? Note: choose two answers.
 - (a) direct control
 - (b) blocking
 - (c) random assignment
 - (d) replication
4. Which of the following graphical techniques should not be used because it requires that users make difficult area judgments.
 - (a) pie chart
 - (b) stem and leaf display
 - (c) box plot
 - (d) bar chart
5. Which of the following graphical techniques is not useful for judging symmetry versus skewness.
 - (a) histogram
 - (b) stem and leaf display
 - (c) box plot
 - (d) bar chart

6. A histogram that is positively skewed
 - (a) is multi-modal
 - (b) is symmetric
 - (c) has a longer upper tail than lower tail
 - (d) has a longer lower tail than upper tail
7. The mean is larger than the median when
 - (a) the distribution is symmetric
 - (b) the distribution is positively skewed
 - (c) the distribution is negatively skewed
 - (d) the distribution is approximately normal
8. Which of the following is not a measure of variability in a data set?
 - (a) IQR
 - (b) variance
 - (c) upper quartile
 - (d) standard deviation
9. Which of the following is most sensitive to outliers in a data set?
 - (a) mean
 - (b) IQR
 - (c) median
 - (d) trimmed mean
10. The percentage of a distribution falling above the lower quartile
 - (a) depends on the shape of the distribution
 - (b) depends on the median of the distribution
 - (c) is 25%
 - (d) is 75%
11. Chebyshev's rule states that the proportion of observations that are within 3 standard deviations of the mean is at least
 - (a) $\frac{1}{3}$
 - (b) $\frac{2}{3}$
 - (c) $\frac{1}{9}$
 - (d) $\frac{8}{9}$
12. The empirical rule states that if a distribution is approximately normal, then approximately 95% of the observations are within ____ standard deviations of the mean.
 - (a) 1
 - (b) 2
 - (c) 3
 - (d) 4

13. The probability that a new tire will have a blow-out in the first year is 0.10. If the four tires in a new car function independently of each other, then what is the probability that at least one tire blows out during the first year?
- (a) 0.10^4
 - (b) $1 - 4 \times 0.10$
 - (c) $1 - 0.10^4$
 - (d) $1 - 0.90^4$
14. Outcomes that cannot occur simultaneously are
- (a) mutually exclusive or disjoint
 - (b) independent
 - (c) impossible
 - (d) equally probable
15. Which of the following is a continuous variable?
- (a) L = the lifetime of a 9 volt battery
 - (b) C = the number of cracked eggs in a carton of 12 eggs
 - (c) F = the number of people in a family
 - (d) B = the brand of laundry detergent used

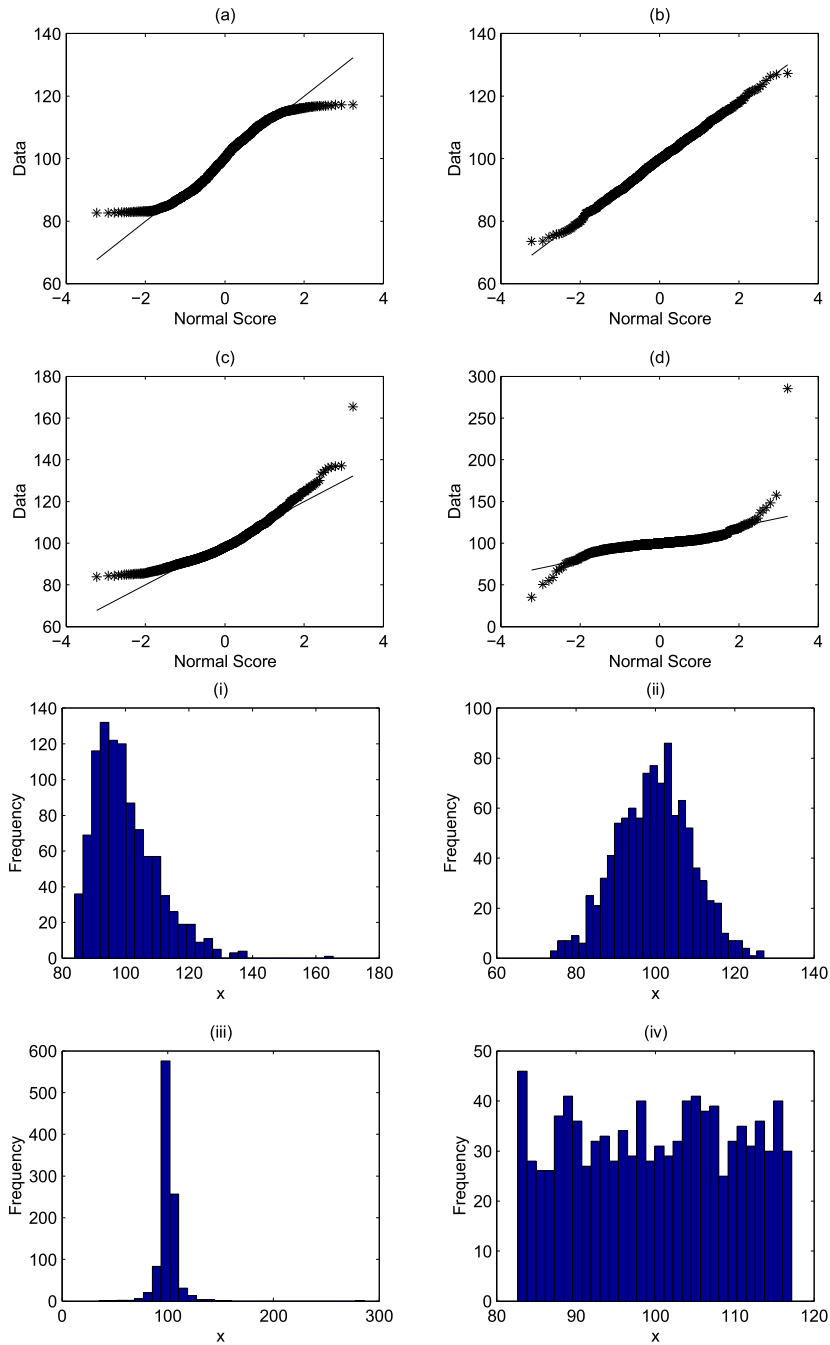
16. The plot below displays the density of a continuous variable X . The probability $P(X > 5)$ is

- (a) 0.50
- (b) 0.15
- (c) 0.75
- (d) 0.25



Matching, short answer and computational questions

17. Below are normal plots and histograms for four data sets of 1,000 observations each. Match the normal plot to the histogram. (2 pts each)



18. Refer to the four data sets in question 17.

- (a) For which data set would a power transformation be most useful in modifying the distribution from non-normal to approximately normal? Report your answer either as one of (a), (b), (c), (d); or as one of (i), (ii), (iii), (iv). (4 pts)
- (b) Would the best value of λ (the power) be larger than 1 or smaller than 1? (2 pts)

19. Let X be the weight of a one year old hatchery rainbow trout. Suppose that the probability distribution of X is normal with mean $\mu = 6$ ounces and standard deviation $\sigma = 1.5$ ounces.
- (a) What is the probability that a randomly selected one year old hatchery rainbow trout weighs between 3 and 6 ounces? (4 pts)
 - (b) What value of X is exceeded by only 10% of the population values? (4 pts)
20. The minimum temperatures on July 4 in Bozeman for the years 1994 to 2000 were the following: 34.0, 50.0, 57.2, 39.2, 48.0, 48.2, and 50.0.
- (a) Compute Q_1 , the median, and Q_3 . Also, sketch a boxplot. (6 pts)
 - (b) The smallest value in the sample is 34.0. How small would this value need to be before it would be called an extreme outlier? (2 pts)
21. List four characteristics that are especially important to look for when examining displays of distributions. (4 pts)

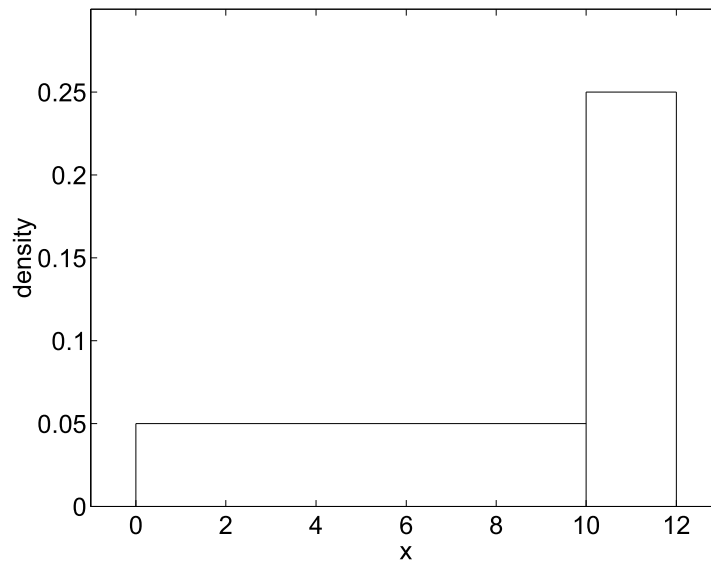
11.2 Midterm Exam 1: Spring 2001

Multiple choice: 4 points each.

1. If a distribution has a strong positive skew, then which of the following is false?
 - (a) The distribution has longer tail on the right than on the left.
 - (b) The median is larger than the mean.
 - (c) The empirical rule does not apply.
 - (d) The distribution is not symmetric.
2. An investigator was interested in the relationship between alcohol use and GPA among MSU undergraduates. The investigator obtained responses to questions concerning alcohol use and GPA from each student in two classes that the investigator was currently teaching. Which of the following sources of bias could not be present in this study?
 - (a) response/measurement
 - (b) selection
 - (c) non-response
 - (d) Each of the above answers is wrong because all sources of bias are present in this study.
3. Outcomes that are independent
 - (a) also must be disjoint (mutually exclusive).
 - (b) cannot be disjoint.
 - (c) could be either disjoint or non-disjoint.
 - (d) are both disjoint and non-disjoint.
4. Which of the following is most important if the investigator would like to conclude that changes in the explanatory variable cause changes in the response?
 - (a) Randomly select cases from the population.
 - (b) Randomly assign cases to the different levels of the response.
 - (c) Randomly select cases from the different levels of the explanatory variable.
 - (d) Randomly assign cases to the different levels of the explanatory variable.

5. Which of the following measures of center is most resistant to the influence of outliers?
- sample mean
 - trimmed mean
 - median
 - average of largest and smallest value in the sample
6. Which of the following experimental design principles is most useful for dealing with the effects of a known extraneous variable without holding it constant.
- direct control
 - blocking
 - random assignment
 - replication
7. The plot below displays the density of a continuous variable X . Find $P(X \leq 11)$

- 0.75
- 0.25
- 0.175
- 0.125



8. In the powerball lottery, the probability of winning a dollar or more is approximately 0.03. George plays the powerball lottery each week for four consecutive weeks. What is the probability that George wins a dollar or more at least once?
- $0.03 \times 0.97^3 = 0.02738$
 - $0.03 \times 4 = 0.12$
 - $0.03^4 = 8.1 \times 10^{-7}$
 - $1 - 0.97^4 = 0.1147$
9. Which of the following is a continuous variable?
- X = the score on a 20 point true/false exam.
 - P = the proportion of males in a statistics class having 20 students.
 - S = the number of grains of sand in a cubic mile.
 - W = the actual weight of a 12 ounce box of breakfast cereal.

10. An investigator is interested in determining whether physical exercise is useful in alleviating psychological depression in elderly people. The investigator divided her sample into three groups depending on their pre-existing level of depression. Subjects within each group were randomly assigned to exercise or control treatments. After three months of treatment the depression level in each subject was measured.
- This study is an example of stratified random sampling. The strata are the three depression groups.
 - This study is flawed because the investigator did not randomly assign the subjects to the three depression groups.
 - This study is an example of a randomized block design. The blocks are the three depression groups.
 - This study is an example of a completely randomized design.
11. According to Chebyshev's rule, at least 50% of a distribution is contained in the interval from $\mu - k\sigma$ to $\mu + k\sigma$, where the value of k
- is 50
 - is 1
 - is $\sqrt{2}$
 - cannot be determined; the value of k must be given

Matching, short answer and computational questions

12. Suppose that the weight of a human brain is normally distributed with mean 1320 grams and standard deviation 100 grams.
- Find the probability that a randomly selected human brain weighs between 1300 and 1400 grams. (5 pts)
 - How much does a human brain weigh if it is at the 95th percentile of the distribution? (5 pts)
13. Twenty senior male athletes competed in the shot-put in English area championships in June, 1991. The following numbers are distances in meters of throws for the 20 athletes. I have ordered the distances for your convenience.
- 11.23 13.22 13.66 13.70 14.05 14.23 14.30 14.46 14.52 14.63
- 14.72 14.92 14.98 15.53 15.81 16.27 16.31 16.47 17.21 17.79
- Compute the five number summary and the IQR for these data. Do not round (6 pts)
 - Sketch a box and whisker plot. Plot mild or extreme outliers using circles and stars. Do not draw whiskers to outliers. (6 pts)
14. A small random sample of males and a small random sample of females was obtained from the population of all MSU undergraduates. Below are heights (in inches) and weights (in pounds) for each case in the samples.

Males		Females	
Height	Weight	Height	Weight
70	160	64	140
68	165	62	110
74	230	66	135

- List the variables in this study and classify them as either categorical or numerical. (5 pts)
- List the data as though it was a standard data file suitable for analysis in SAS. (5 pts)
- What sampling plan was used? Justify your answer. (4 pts)

(d) Compute the sample mean and the sample variance of the male heights. (6 pts)

15. Below are density plots from four distributions. For each plot, sketch the normal probability plot. (2 pts per plot)

