Classroom Voting Questions: Statistics

Looking at Data - Distributions

Displaying Distributions with Graphs

1. A data set consists of fifty three-digit numbers ranging from 180 to 510. The best choice for stems in a stem-and-leaf display would be to use ______.
   (a) 1 digit stems (1, 2, . . ., 5)
   (b) 2 digit stems (18, 19, . . ., 51)
   (c) 3 digit stems (180, 181, . . ., 510)

2. Which of the following statements is most completely true in comparing an appropriately drawn histogram to a stem-and-leaf display of the same data?
   (a) Both convey the same information about the shape of the distribution.
   (b) Both convey the same information about gaps in the distribution.
   (c) Both convey the same information about outliers.
   (d) Both convey the same amount of information generally.
   (e) Two from (A)-(D) are correct.
   (f) Three from (A)-(D) are correct.
   (g) All from (A)-(D) are correct.

3. Suppose you take a random sample of 10 juniors who bought the same model laptop when they were freshmen. You test these 10 laptops to determine how long their batteries last before needing to be recharged, and you obtain the following data (in hours): 1.2, 1.3, 3.8, 3.9, 3.9, 4.0, 4.1, 4.1, 4.2, 4.3. What should be done with the values 1.2 and 1.3? Which of the following is the best course of action?
   (a) Delete them from the data set since they are outliers.
   (b) Keep them in the data set even though they are outliers.
   (c) Determine why these values were so much lower than the rest, then delete them.
   (d) Determine why these values were so much lower than the rest, then keep them in the data set, provided they weren't due to data entry errors.

4. You consider all of the adult patients in a large hospital. Which of the following variables would you expect to have a distribution that is left-skewed as revealed by a dot plot of the data?
   (a) height
   (b) annual income
   (c) eye color
   (d) age

5. You consider all of the adult patients in a large hospital. Which of the following variables is continuous?
   (a) height
   (b) weight
   (c) number of past surgeries
   (d) more than one of the above

6. You consider all of the adult patients in a large hospital. Which of the following variables is discrete?
   (a) height
   (b) eye color
   (c) number of siblings
   (d) more than one of the above

Describing Distributions with Numbers

7. Think carefully about the heights of adult men and women, then decide which of the following statements is true:
   (a) Women are taller than men.
   (b) Men are taller than women.
   (c) Men and women are the same height.
   (d) None of the above

8. Mary has two legs, and is a resident of this city. What would Mary learn if she compared her number of legs with the average number of legs of all the people in this city?

(a) She has the average number of legs.
(b) She has an above average number of legs.
(c) She has a below average number of legs.
(d) None of the above

9. In a certain university there are three types of professors. Their salaries are approximately normally distributed within each of the following types:

- Assistant Professors make a median salary of \$50K, with a minimum of \$40K and a maximum of \$60K.
- Associate Professors make a median salary of \$65K per year, with minimum of \$57K and a maximum of \$80K.
- Full Professors make a median salary of \$90K per year, with a minimum of \$70K and a maximum of \$110K.

There are 1600 total Professors at this University, with the following distribution: 50% of all Professors are Assistants, 30% are Associates, and 20% are Fulls.

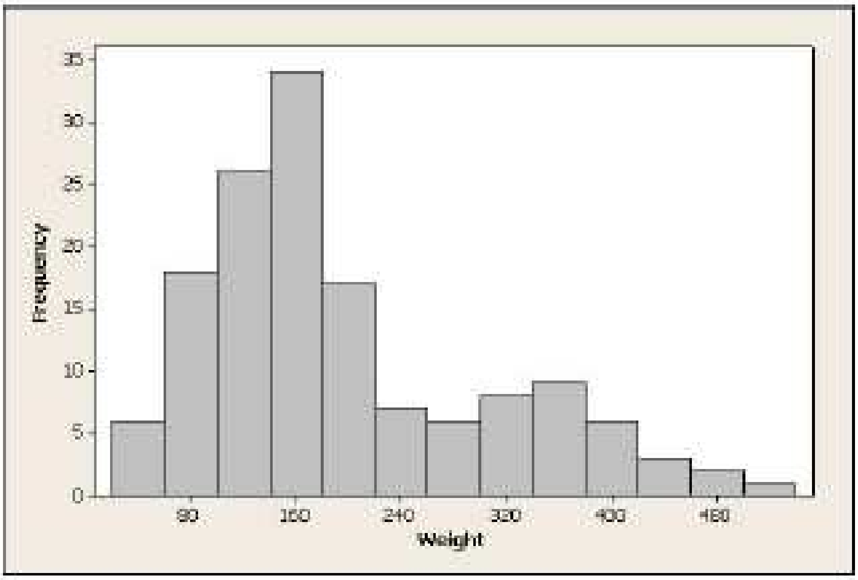
What can we say about the average salary at this university?

(a) mean $<$ median
(b) mean $=$ median
(c) mean $>$ median
(d) insufficient information

10. Many individuals, after the loss of a job, receive temporary pay unemployment compensation until they are re-employed. Consider the distribution of time to re-employment as obtained in an employment survey. One broadcast reporting on the survey said that the average time until re-employment was 4.5 weeks. A second broadcast reported that the average was 9.9 weeks. One of your colleagues wanted a better understanding of the situation and learned (through a Google search) that one report was referring to the mean and the other to the median and also that the standard deviation was about 14 weeks. Knowing that you are a statistically-savvy person, your colleague asked you which is most likely the mean and which is the median?

(a) 4.5 is the mean and 9.9 is the median.
(b) 4.5 is the median and 9.9 is the mean.
(c) Neither (A) nor (B) is possible given the SD of the data.
(d) I am not a statistically-savvy person, so how should I know?

11. For the data set displayed in the following histogram, which would be larger?



(a) mean

(b) median

(c) Can't tell from the given histogram.

12. Why is the term $(n-1)$ used in the denominator of the formula for sample variance?

    (a) There are $(n-1)$ observations.

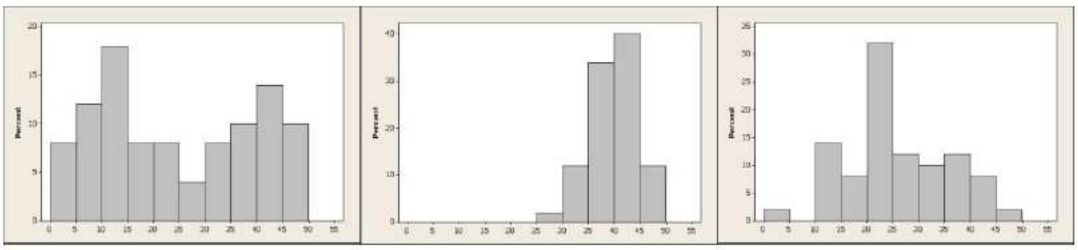
    (b) There are $(n-1)$ uncorrelated pieces of information.

    (c) The $(n-1)$ term gives the correct answer.

    (d) There are $(n-1)$ samples from the population.

    (e) There are $(n-1)$ degrees of freedom.

13. Which of the three histograms shown summarizes the data set with the smallest standard deviation?

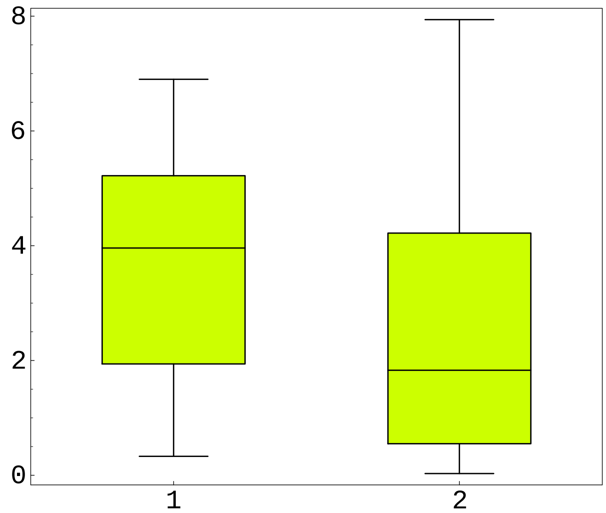


1 2 3

14. Suppose your statistics instructor tells you that you scored 70 on an exam and that the class mean was 74. You should hope that the standard deviation of exam scores was ______.

   (a) Small
   (b) Large

15. Below are boxplots for two data sets.



   TRUE or FALSE: There is a greater proportion of values outside the box for the set on the right than for the set on the left.

   (a) True, and I am very confident.
   (b) True, and I am not very confident.
   (c) False, and I am not very confident.
   (d) False, and I am very confident.

16. The five-number summary for all student scores on an exam is 29, 42, 70, 75, 79. Suppose 200 students took the test. How many students had scores between 42 and 70?

   (a) 25
   (b) 28
   (c) 50
   (d) 100

17. The five-number summary for all student scores on an exam is 40, 60, 70, 75, 79. Suppose 500 students took the test. What should you conclude about the distribution of scores?

    (a) Skewed to the left.
    (b) Skewed to the right.
    (c) Not skewed.
    (d) Not enough information given to determine skew.

18. Jack uses a calculator to find the sample standard deviation of a data set and ends up getting a negative number as the result. This implies that

    (a) on average, the deviations from the mean are negative.
    (b) the mean of the deviations is negative.
    (c) the mean is negative.
    (d) Jack made a mistake.

19. Which set of two observations would you expect to have the smaller standard deviation, the weights of two randomly-selcted professional ballerinas, or the weights of two randomly-selected sumo wrestlers?

    (a) The weights of the ballerinas, because ballerinas are lighter.
    (b) The weights of the sumo wrestlers, because sumo wrestlers are heavier.
    (c) The weights of the ballerinas; their weights are more likely to be closer together.
    (d) The weights of the ballerinas; their weights are more likely to be farther apart.
    (e) The weights of the sumo wrestlers; their weights are more likely to be closer together.
    (f) The weights of the sumo wrestlers; their weights are more likely to be farther apart.

20. When a professional sumo wrestler joined a group of people, the standard deviation of the weights of the new group members was substantially less than the standard deviation of the weights of the original group members. Which of the following is most likely?

    (a) The original group consisted of 3 professional sumo wrestlers.
    (b) The original group consisted of 100 professional sumo wrestlers.
    (c) The original group consisted of 3 professional ballerinas.
    (d) The original group consisted of 100 professional ballerinas.

21. A multi-billionaire decides to retire back in the small town in which she grew up. All of the houses in this town are modest and inexpensive. On the outskirts of town, she builds a huge, luxurious mansion. Consider house prices in the town before and after she builds her mansion. Which of the following measures of central tendency changes the most?

   (a) mean
   (b) median
   (c) mode

22. Which of the following measures is resistant to the influence of outliers?

   (a) mean
   (b) median
   (c) standard deviation
   (d) Q_3
   (e) interquartile range
   (f) two out of (a) through (e)
   (g) three out of (a) through (e)
   (h) four out of (a) through (e)

23. In a history class with over 500 students, a professor gave a very easy test, so that the distribution of scores was highly left-skewed. Which measure of central tendency and measure of variation should be used to summarize the scores?

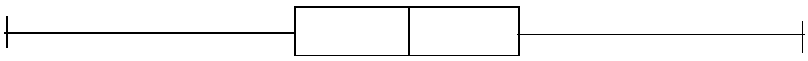
   (a) median and standard deviation
   (b) median and interquartile range
   (c) mean and standard deviation
   (d) mean and interquartile range

24. Consider a data set that consists of the following four numbers: 2, 5, 6, and a certain number that is less than negative one million. For this data set, rank the following from least to greatest: mean, median, standard deviation.

   (a) mean, median, standard deviation
   (b) mean, standard deviation, median
   (c) median, mean, standard deviation
   (d) median, standard deviation, mean
   (e) standard deviation, mean, median
   (f) standard deviation, median, mean

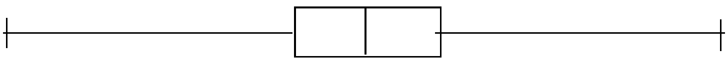
Density Curves and Normal Distributions

25. If a large sample were drawn from a normal distribution and accurately represented the population, which of the following is most likely to be a box plot of that sample?

(A) 

(B) 

(C) 

(D) 

(E) Two from (A)-(D) are correct.

(F) Three from (A)-(D) are correct.

(G) All from (A)-(D) are correct.

26. Consider the continuous random variable X = the weight in pounds of a randomly selected newborn baby born in the United States last year. Suppose that X can be modeled with a normal distribution with mean $\mu = 7.57$ and standard deviation $\sigma = 1.06$. If the standard deviation were $\sigma = 1.26$ instead, how would that change the graph of the pdf of X?

    (a) The graph would be narrower and have a greater maximum value.
    (b) The graph would be narrower and have a lesser maximum value.
    (c) The graph would be narrower and have the same maximum value.
    (d) The graph would be wider and have a greater maximum value.
    (e) The graph would be wider and have a lesser maximum value.
    (f) The graph would be wider and have the same maximum value.

27. Consider the continuous random variable X = the weight in pounds of a randomly newborn baby born in the United States during 2006. Suppose that X can be modeled with a normal distribution with mean $\mu = 7.57$ and standard deviation $\sigma = 1.06$. If the mean were $\mu = 7.27$ instead, how would that change the graph of the pdf of X?

    (a) The graph would be shifted to the left.

(b) The graph would be shifted to the right.

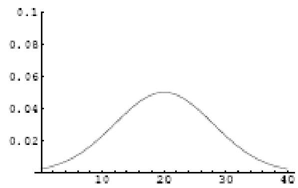
(c) The graph would become more negatively skewed.

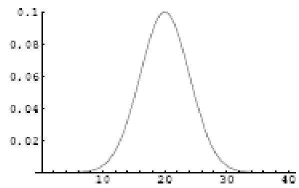
(d) The graph would become more positively skewed.

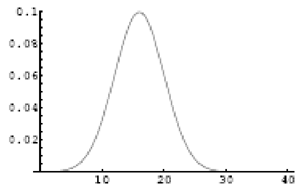
(e) The graph would have a greater maximum value.

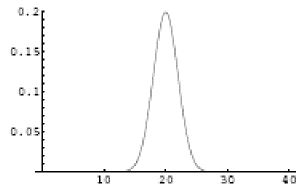
(f) The graph would have a lesser maximum value.

28. If X is a normal random variable with mean $\mu = 20$ and standard deviation $\sigma = 4$, which of the following could be the graph of the pdf of X?

(A) 

(B) 

(C) 

(D) 

29. Find $z_{0.15}$.

(a) 1.04

(b) -1.04

30. Yogurt is sold in cartons labeled as containing 6 oz, but the actual contents vary slightly from container to container. Suppose that the content distribution is approximately normal in shape with a mean of 6 oz and a standard deviation of 0.05 oz. What can be said about the percentage of cartons that have actual contents less than 5.95 oz?

(a) The percentage is approximately 68%

(b) The percentage is approximately 34%

(c) The percentage is approximately 32%

(d) The percentage is approximately 16%

31. The University of Oklahoma has changed its admission standards to require an ACT-score of 26. We know the ACT is normally distributed with a mean of 21 and an SD of 5. If we sample 100 students who took the ACT at random, how many would be expected to qualify for admission to OU?

(a) 5
(b) 16
(c) 34
(d) 84
(e) none of the above

32. A colleague has collected 1000 old VW vans for resale. The colleague, an old stats professor, will only sell a van to those who can answer the following question: The -2 SD sales price for one of these vans is set at \$550; and $+2$ SD sales price is set at \$1100. He tells you the distribution of sales prices is approximately normal. What is the expected number of vans for sale between \$550 and \$1100?

    (a) 500
    (b) 680
    (c) 750
    (d) 888
    (e) 950

33. The heights of women are normally distributed with a mean of 65 inches and an SD of 2.5 inches. The heights of men are also normal with a mean of 70 inches. What percent of women are taller than a man of average height?

    (a) 0.15%
    (b) 2.5%
    (c) 5%
    (d) 16%
    (e) insufficient information

34. Many psychological disorders (e.g. Depression, ADHD) are based on the application of the 2 SD rule assuming a normal distribution of reported symptoms. This means that anyone who reports a symptom count that is greater than the 2 SD point in a normal population can be considered to be "abnormal" or "disordered".

    Given this definition of "disorder", what is expected prevalence rate of these disorders based on the 2 SD rule?

    (a) 0.15%
    (b) 2.5%
    (c) 5%

(d) 16%

(e) 95%

35. The ACT has a mean of 21 and an SD of 5. The SAT has a mean of 1000 and a SD of 200. Joe Bob Keith took the ACT and he needs a score of 1300 on the SAT to get into UNC-Chapel Hill and a score of 1400 on the SAT to get into Duke. UNC and Duke both told Joe Bob Keith that they will convert the ACT to the SAT using a z-score (or standard-score) transformation. Joe Bob Keith has decided to go to the school with the highest standards that will accept him. If he doesn't qualify for either Duke or UNC, then it's Faber College for Joe Bob Keith. As it turns out, Joe Bob Keith got a 30 on the ACT, but he cannot figure out what that means for his choice of college. Help Joe Bob Keith out. Where is he going to school?

(a) UNC

(b) Duke

(c) Faber

36. Let Z be a standard normal random variable. Which of the following probabilities is the smallest?

(a) $P(-2 < Z < -1)$

(b) $P(0 < Z < 2)$

(c) $P(Z < 1)$

(d) $P(Z > 2)$

37. Let Z be a standard normal random variable. Which of the following probabilities is the smallest?

(a) $P(0 \leq Z \leq 2.07)$

(b) $P(-0.64 \leq Z \leq -0.11)$

(c) $P(Z > -1.06)$

(d) $P(Z < -0.88)$

38. 77% of the area under a normal curve lies to the left of what z-score?

(a) 0.74

(b) 0.77

(c) $z_{0.77}$

(d) 0.78

39. Intelligence quotients (IQs) are normally distributed with a mean of 100 and a standard deviation of 16. What percentage of the population has an IQ between 112 and 116?

    (a) 4%
    (b) 7%
    (c) 9%
    (d) 25%

40. Intelligence quotients (IQs) are normally distributed with a mean of 100 and a standard deviation of 16. Find Q_3 for IQ.

    (a) 111
    (b) 112
    (c) A continuous distribution does not have quartiles.

41. Jeannie works at the drive-through window at a local fast-food restaurant. In the middle of the afternoon, the mean time between customers arriving at the window is 5 minutes with a standard deviation of 5 minutes. As a customer drives up to the window, Jeannie is wondering what the probability is that the next customer will arrive more than 10 minutes from now. TRUE or FALSE: Jeannie should convert 10 into a z-score and then find the area to the right of that z-score under the standard normal curve.

    (a) True, and I am very confident.
    (b) True, and I am not very confident.
    (c) False, and I am not very confident.
    (d) False, and I am very confident.

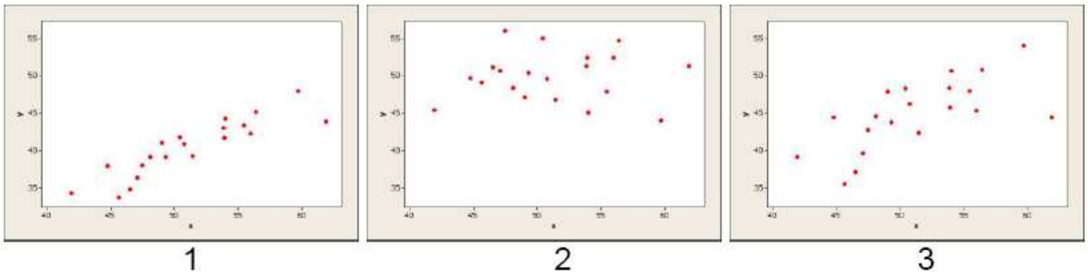
42. Which of the following is not a reason for constructing a normal probability plot or a normal quantile plot?

    (a) You are about to perform an inferential statistics procedure.
    (b) You have computed a five-number summary and would like to display the results.
    (c) You are concerned about outliers.
    (d) You are concerned about skewness.

Chapter 2: Looking at Data - Relationships

Correlation

43. The scatterplots below display three bivariate data sets. The correlation coefficients for these data sets are 0.03, 0.68, and 0.89. Which scatter plot corresponds to the data set with $r = 0.03$?

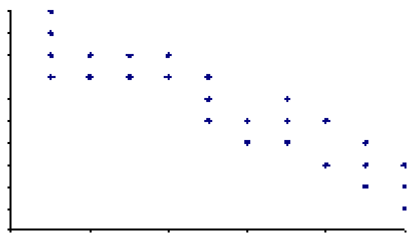


(a) Plot 1
(b) Plot 2
(c) Plot 3

44. Joe Bob found a strong correlation in an empirical study showing that individuals' physical ability decreased significantly with age. Which numerical result below best describes this situation?

(a) $r = -1.2$
(b) $r = -1.0$
(c) $r = -0.8$
(d) $r = +0.8$
(e) $r = +1.0$
(f) $r = +1.2$

45. Which correlation best describes the scatterplot?



(a) -0.7

(b) -0.3

(c) 0

(d) $+0.3$

(e) $+0.7$

46. If you believed strongly in the idea that the more hours per week full-time students work in a job, the lower their GPA would be, then which correlation would you realistically expect to find?

(a) -0.97

(b) -0.72

(c) -0.20

(d) $+0.20$

(e) $+0.72$

(f) $+0.97$

47. A researcher found that $r = +.92$ between the high temperature of the day and the number of ice cream cones sold at Cone Island. This result tells us that

(a) high temperatures cause people to buy ice cream.

(b) buying ice cream causes the temperature to go up.

(c) some extraneous variable causes both high temperatures and high ice cream sales.

(d) temperature and ice cream sales have a strong positive linear relationship.

48. You are conducting a correlation analysis between a response variable and an explanatory variable. Your analysis produces a significant positive correlation between the two variables. Which of the following conclusions is the *most* reasonable?

(a) Change in the explanatory variable causes change in the response variable.
(b) Change in the explanatory variable is associated with in change in the response variable.
(c) Change in the response variable causes change in the explanatory variable.
(d) All from (a)-(c) are equally reasonable conclusions.

49. The salary and the numbers of years of teaching experience were recorded for 20 social studies teachers in rural west Texas. When the data points were plotted, there was a roughly linear relationship and a positive correlation between salary and number of years of teaching experience, with $r = 0.8$. What percentage of the variation in the salaries is explained by the linear relationship between salary and years of service?

(a) 80%
(b) 64%
(c) 36%
(d) 20%

Least-Squares Regression

50. A store manager conducted an experiment in which he systematically varied the width of a display for toothpaste from 3 ft. to 6 ft. and recorded the corresponding number of tubes of toothpaste sold per day. The data was used to fit a regression line, which was

$$\text{tubes sold per day} = 20 + 10(\text{display width})$$

What is the predicted number of tubes sold per day for a display width of 12 feet?

(a) 120
(b) 140
(c) It would be unwise to use the regression line to make a prediction for a display width of 12 ft.

Cautions about Correlation and Regression

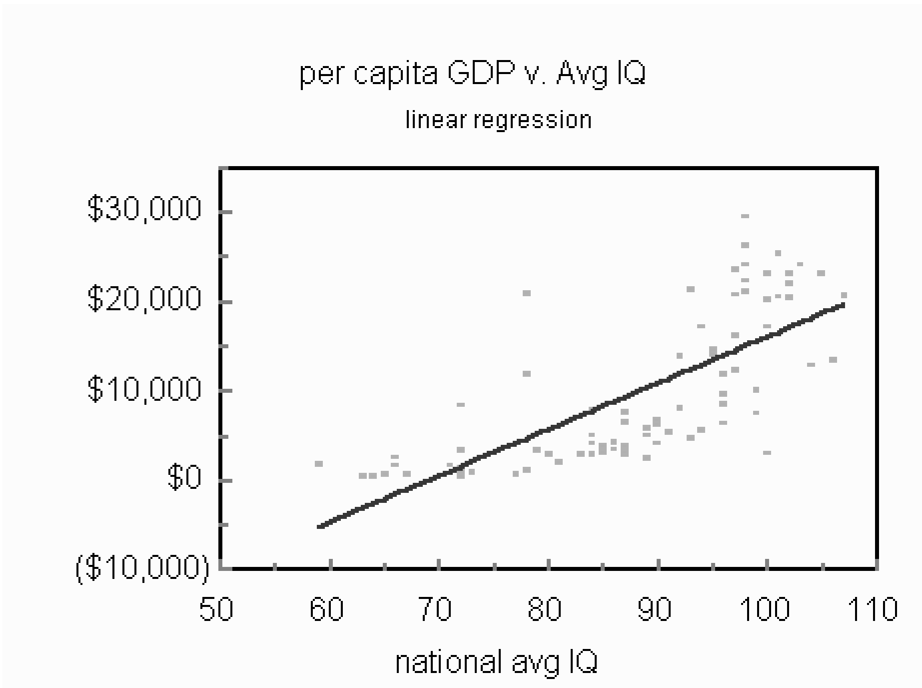
51. Gas mileage and weight were recorded for each automobile in a sample of 20 compact cars. There was a strong negative correlation, with r = -.87. Based on the value of r, it is reasonable to conclude that increasing the weight of a compact car causes a decrease in gas mileage.

(a) True, and I am very confident.
(b) True, and I am not very confident.
(c) False, and I am not very confident.
(d) False, and I am very confident.

52. Which of the following characteristics in a residual plot are indicative of potential problems?

(a) A strong pattern in the residual plot
(b) Isolated points in the residual plot
(c) A lack of any strong pattern in the residual plot
(d) Both (a) and (b) above are indicative of potential problems
(e) (a), (b), and (c) above are all indicative of potential problems

53. Which phrase best describes the scatterplot?



(a) strong $+r$
(b) strong $-r$
(c) weak $+r$
(d) weak $-r$
(e) influential outliers

(f) non-linearity
(g) Two from (A)-(F) are true.
(h) Three from (A)-(F) are true.

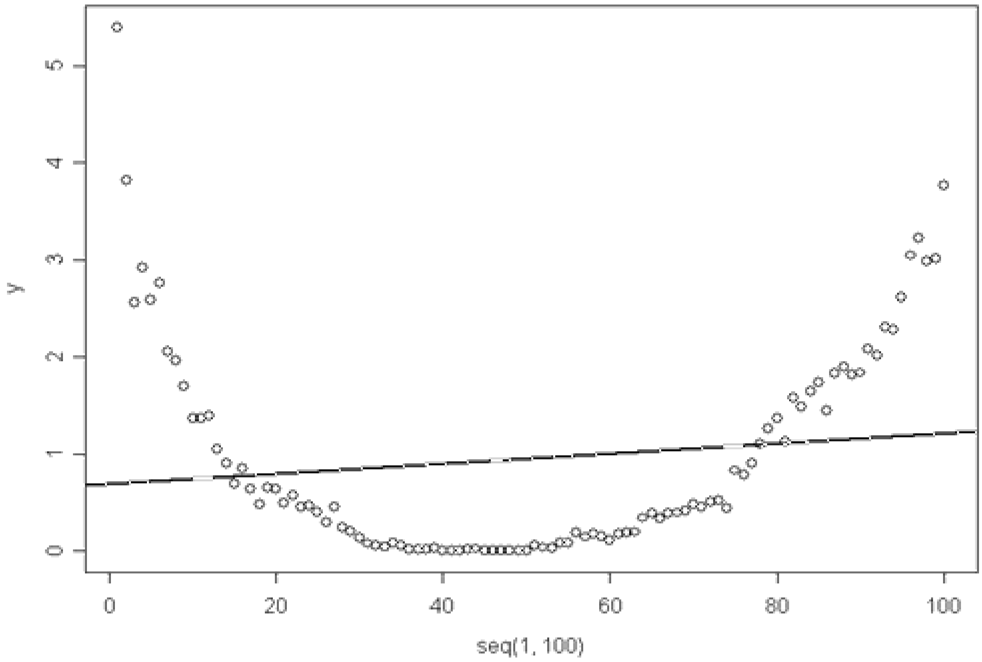
54. Why is it important to look for outliers in data prior to applying regression?

(a) Outliers always affect the magnitude of the regression slope.
(b) Outliers are always bad data.
(c) Outliers should always be eliminated from the data set.
(d) Outliers should always be considered because of their potential influence.
(e) We shouldn't look for outliers, because all the data must be analyzed.

55. Which of the following factors is *NOT important* to consider when interpreting a correlation coefficient?

(a) restriction of range
(b) problems associated with aggregated data
(c) outliers
(d) lurking variables
(e) unit of measurement

56. What is the greatest concern about the regression below?



(a) It has a small slope.

(b) It has a high R^2.

(c) The investigator should not be using a linear regression on these data.

(d) The residuals are too large.

(e) The regression line does not pass through the origin.

Data Analysis for Two-Way Tables

57. The following contingency table/two-way table classifies the members of a certain government into political party (Liberal or Conservative) and whether they support or oppose the spending bill that is currently up for adoption.

	Support	**Oppose**	**Total**
Liberal	47	11	58
Conservative	14	35	49
Total	61	46	107

What fraction of the government members are conservatives who support the bill?

(a) $\dfrac{14}{61}$

(b) $\dfrac{14}{49}$

(c) $\dfrac{14}{107}$

(d) None of the above

58. The following contingency table/two-way table classifies the members of a certain government into political party (Liberal or Conservative) and whether they support or oppose the spending bill that is currently up for adoption.

	Support	**Oppose**	**Total**
Liberal	47	11	58
Conservative	14	35	49
Total	61	46	107

What fraction of the liberals support the bill?

(a) $\dfrac{47}{61}$

(b) $\dfrac{47}{58}$

(c) $\dfrac{47}{107}$

(d) None of the above

59. The following contingency table/two-way table classifies the members of a certain government into political party (Liberal or Conservative) and whether they support or oppose the spending bill that is currently up for adoption.

	Support	**Oppose**	**Total**
Liberal	47	11	58
Conservative	14	35	49
Total	61	46	107

The following fractions are formed by dividing numbers in the table: $\frac{11}{58}, \frac{58}{107}, \frac{11}{107}$. In order, these numbers are part of which distributions?

(a) joint, marginal, conditional

(b) joint, conditional, marginal

(c) marginal, joint, conditional

(d) marginal, conditional, joint

(e) conditional, joint, marginal

(f) conditional, marginal, joint

60. Phoenix and Cassandra are professional basketball players who have dealt with injuries over the past two seasons. Two seasons ago, Phoenix made 5 out of 10 free throws, while Casssandra made 60 out of 100 free throws. Last season, Phoenix made 139 out of 200 free throws, while Cassandra made 7 out of 10 free throws. This leads to the two joint distributions below. (The first table is for two seasons ago; the second table is for last season.)

	Made	**Missed**
Phoenix	0.5	0.5
Cassandra	0.6	0.4

	Made	**Missed**
Phoenix	0.695	0.305
Cassandra	0.7	0.3

Which player was the best free-throw shooter?

(a) Phoenix

(b) Cassandra

Chapter 3: Producting Data

Design of Experiments

61. "Graduating is good for your health," according to a headline in the Boston Globe (3 April 1998). The article noted "According to the Center for Disease Control, college graduates feel better emotionally and physically than do high school dropouts." Do you think the headline is justified based on this statement?

(a) Yes, as long as the data was from random samples of college graduates and high school dropouts.

(b) Yes, because this must have been an observational study. As long as it was a well-designed study, the headline is justified.

(c) No, because the headline implies a cause and effect relationship, which is not justified based on an observational study.

(d) No, because this study must have been an experiment and we cant draw cause and effect conclusions from an experiment.

62. In a study of perceived importance of money, 100 attorneys were selected at random from those in private practice and 100 attorneys were selected at random from those employed by government agencies as district attorneys. The attorneys in each group were asked to respond to a set of questions designed to assess level of stress in the workplace. This study is ______.

    (a) an observational study
    (b) an experiment

63. When is it unreasonable to reach a cause-and-effect conclusion based on data from a statistical study?

    (a) Any time the study is based on a random sample from a population of interest.
    (b) When the study is observational.
    (c) When the study is a well-designed experiment that uses random assignment to experimental conditions (treatments).
    (d) It is always reasonable to reach a cause-and-effect conclusion based on data from a statistical study.

Sampling Design

64. Researchers believe that one possible cause of Very Low Birth Weight (VLBW) infants is the presence of undiagnosed infections in the mother. To assess this possibility, they collected data on all pregnant women presenting themselves for prenatal care at large urban hospitals. What is the *appropriate population* for this study?

    (a) All infants.
    (b) All infants born as VLBW infant.
    (c) All infants born in large urban centers.
    (d) All pregnant women.
    (e) All pregnant women living in large urban centers.

65. A Gallup survey was taken recently regarding peoples current preference for Democratic nominee for President for which there are 11 candidates. The survey also collected gender information, in order to capture male female differences in preference. For this poll, what is the *primary variable* of interest and *how many values* does it take?

(a) gender; 2
(b) gender; more than 2
(c) candidate preference; 2
(d) candidate preference; more than 2
(e) political party: 2
(f) political party; more than 2

66. Increasing sample size

(a) has no effect on bias.
(b) increases bias.
(c) decreases bias.

67. If you were trying to obtain a random sample of a population of interest for a political poll for a local mayoral race, which of the following approaches would be best to obtain the random sample?

(a) Randomly assign a number to local companies and, using random-number generation, go to those companies selected and conduct interviews.
(b) Randomly select a busy street corner in your city and conduct on-site interviews.
(c) Assign a number to people in the local phone book and, using random-number generation, call those randomly selected.
(d) Randomly select a couple of television stations from your local cable company using random number generation and ask people through advertising to call a polling line.
(e) Randomly dial phone numbers within the selected area and interview those who answer the phone.

68. In order to estimate the proportion of students at a small liberal arts college who watch reality TV for more than 4 hours per week, a random sample of students at the school is selected and each is interviewed about his or her reality TV viewing habits. The students conducting the survey are worried that people that watch reality TV might be embarrassed to admit it and that they may not respond to the survey with honest answers. What type of bias are the students conducting the survey worried about?

(a) They shouldn't worry - there is no obvious source of bias.
(b) Voluntary bias
(c) Nonresponse bias
(d) Response bias

69. A sample of 162 students at a large university is taken and it is found that 110 of them, that is, 68% of them, have considered changing their major at some point. Is the 68% a statistic or a parameter?

    (a) A parameter, because it is a measurement describing a characteristic of a sample.
    (b) A parameter, because it is measurement describing a characteristic of a population.
    (c) A statistic, because it is a measurement describing a characteristic of a sample.
    (d) A statistic, because it is a measurement describing a characteristic of a population.

70. Consider a standard 52-card deck, with four suits (hearts(red), diamonds(red), spades(black), clubs(black)), 13 cards per suit (2-10, J, Q, K, A). Define an event space on the standard deck such that it consists of 52 simple outcomes, one for each card in the deck.

    Which of the following is a true statement?

    (a) Black is not an event.
    (b) Black is an event with 1 simple outcome.
    (c) Black is an event with 26 simple outcomes.
    (d) Black is an event with 52 simple outcomes.
    (e) None of the above is true.

Random Variables

71. Draw the following dart board: A dart board is constructed from three concentric circles with radii 1 inch, 2 inches, and 3 inches, respectively. If a dart lands in the innermost circle, the player receives 4 points. If the dart lands between the innermost circle and the middle circle, the player receives 2 points. If the dart lands between the middle circle and the outermost circle, the player receives 1 point. Assume that the probability of a dart landing in any particular region is proportional to the area of that region.

    Define the random variable X to be the sum of the player's score on two successive throws. Then X is what type of random variable?

    (a) discrete
    (b) continuous

72. Draw the following dart board: A dart board is constructed from three concentric circles with radii 1 inch, 2 inches, and 3 inches, respectively. If a dart lands in the innermost circle, the player receives 4 points. If the dart lands between the innermost circle and the middle circle, the player receives 2 points. If the dart lands between the middle circle and the outermost circle, the player receives 1 point. Assume that the probability of a dart landing in any particular region is proportional to the area of that region.

Suppose that a player's score on a single dart throw is defined to be the distance between the dart and the center of the board. Define the random variable X to be the sum of the player's score on two successive throws. Then X is what type of random variable?

(a) discrete

(b) continuous

73. A radioactive mass emits particles at an average rate of 15 particles per minute. Define the random variable X to be the number of particles emitted in a 10-minute time frame. Then X is what type of random variable?

(a) discrete

(b) continuous

74. A radioactive mass emits particles at an average rate of 15 particles per minute. A particle is emitted at noon today. Define the random variable X to be the time elapsed between noon and the next emission. Then X is what type of random variable?

(a) discrete

(b) continuous

75. A randomly-selected kindergarten class in a large city will get to have a party on Friday of next week. At one point in the party, each child in the class will receive half of a candy bar. Define the random variable X to be the number of candy bars given out in the class next Friday. Then X is what type of random variable?

(a) discrete

(b) continuous

76. Consider the continuous random variable X = the weight in pounds of a randomly selected newborn baby born in the United States during 2006. Let f be the probability density function for X. It is probably safe to say that $P(X < 0) = 0$ and $P(X < 20) = 1$. Which of the following is *not* a justifiable conclusion about f given this information?

(a) No portion of the graph of f can lie below the x-axis.

(b) The area under the entire graph of f equals 1.

(c) The area under the graph of f between $x = 0$ and $x = 20$ is 1.

(d) The nonzero portion of the graph of f lies entirely between $x = 0$ and $x = 19$.

77. A randomly selected family has two kids. What is the probability that the family has one boy and one girl?

(a) $\frac{1}{2}$

(b) $\frac{1}{3}$

(c) $\frac{1}{4}$

(d) None of the above

78. Two standard, six-sided dice are rolled. What is the probability that the sum of the dice is 6?

(a) $\frac{1}{6}$

(b) $\frac{5}{6}$

(c) $\frac{1}{12}$

(d) $\frac{5}{12}$

(e) $\frac{1}{36}$

(f) $\frac{5}{36}$

79. Two standard, six-sided dice are rolled. What is the most probable sum?

(a) 2

(b) 6

(c) 7

(d) 12

80. Consider rolling a standard, six-sided die. Let A be the event that the number rolled is even. Let B be the event that the number rolled is a multiple of 3. The event $(not B)$ consists of

(a) $1, 3, 5$

(b) $1, 2, 4, 5$

(c) $2, 4, 6$

(d) $1, 3, 5$

81. Consider rolling a standard, six-sided die. Let A be the event that the number rolled is even. Let B be the event that the number rolled is a multiple of 3. The event (A and B) consists of

    (a) $2, 3, 4, 6$
    (b) $2, 3, 4, 6, 6$
    (c) 6

82. Consider rolling a standard, six-sided die. Let A be the event that the number rolled is even. Let B be the event that the number rolled is a multiple of 3. The event $(A or B)$ consists of

    (a) $2, 3, 4, 6$
    (b) $2, 3, 4, 6, 6$
    (c) 6

83. A standard, six-sided die is rolled. What is the probability of rolling an even number or a number divisible by 3?

    (a) $\frac{2}{3}$
    (b) $\frac{5}{6}$
    (c) 4
    (d) 5

84. A card is drawn at random from a standard deck of 52 playing cards. What is the probability that the card is a red card or a jack?

    (a) 28
    (b) 30
    (c) $\frac{7}{13}$
    (d) $\frac{15}{26}$

Means and Variances of Random Variables

85. Suppose that a random variable X has only two values, 0 and 1. If $P(X = 0) = 0.5$ then what can we say about $E(X)$?

   (a) $E(X) = 0$
   (b) $E(X) = 0.5$
   (c) $E(X) = 1$
   (d) Either (A) or (C) is possible.
   (e) Both (A) and (C).
   (f) insufficient information

86. Suppose that a random variable X has only two values, 0 and 1. If $P(X = 0) = 0.5$ then what can we say about $\text{Var}(X)$?

   (a) $\text{Var}\ (X) = -0.25$
   (b) $\text{Var}\ (X) = 0$
   (c) $\text{Var}\ (X) = 0.25$
   (d) $\text{Var}\ (X) = 0.5$
   (e) $\text{Var}\ (X) = 1$
   (f) insufficient information

87. Suppose that a random variable X has only two values, 3 and 4. If $P(X = 3) = 0.5$ then what can we say about $E(X)$?

   (a) $E(X) = 0.5$
   (b) $E(X) = 1$
   (c) $E(X) = 3$
   (d) $E(X) = 3.5$
   (e) $E(X) = 4$

88. Suppose that a random variable X has only two values, 3 and 4. If $P(X = 3) = 0.5$ then what can we say about $\text{Var}(X)$?

   (a) $\text{Var}\ (X) = 0.25$
   (b) $\text{Var}\ (X) = 0.5$
   (c) $\text{Var}\ (X) = 0.75$

(d) Var $(X) = 1.0$

(e) Var $(X) = 3.25$

(f) Var $(X) = 3.5$

89. Suppose your instructor asks you a multiple-choice question with three answer choices in class. You are to submit your answer and also rate the confidence (low, medium, or high) with which you believe in that answer. You will be scored based on the following chart.

Confidence	Correct Answer	Incorrect Answer
Low	3	2
Medium	4	1
High	5	0

If have no idea what the answer to the question is and you have to guess randomly among the three available answer choices, what confidence level should you choose in order to maximize your points?

(a) Low

(b) Medium

(c) High

(d) It doesn't matter.

90. Suppose your instructor asks you a multiple-choice question with *two* answer choices in class. You are to submit your answer and also rate the confidence (low, medium, or high) with which you believe in that answer. You will be scored based on the following chart.

Confidence	Correct Answer	Incorrect Answer
Low	3	2
Medium	4	1
High	5	0

If have no idea what the answer to the question is and you have to guess randomly among the two available answer choices, what confidence level should you choose in order to maximize your points?

(a) Low

(b) Medium

(c) High

(d) It doesn't matter.

91. A small store located not too far from a campground sells cartons containing six eggs. Each customer is limited to a maximum purchase of three cartons. Among customers who buy eggs, 50% buy one carton, 30% buy two cartons, and 20% buy three cartons. What is the mean number of eggs purchased by customers who purchase eggs?

   (a) 1.7
   (b) 3.4
   (c) 10.2
   (d) 12

92. Manuel, a biology major, works in the Admissions Office to help cover his tuition and other expenses. One day a mysterious stranger drops by the Admissions Office and offers Manuel the opportunity to play the following game one time: A fair coin will be flipped. If the coin comes up heads, Manuel will be given \$3. If the coin comes up tails, Manuel has to pay \$1. Should Manuel play the game?

   (a) Yes, because the expected value for Manuel is positive.
   (b) This is a fair game, so one can't really answer "Yes" or "No."
   (c) No, because the expected value for Manuel is negative.

93. Manuel, a biology major, works in the Admissions Office to help cover his tuition and other expenses. One day a mysterious stranger drops by the Admissions Office and offers Manuel the opportunity to play the following game one time: A fair coin will be flipped. If the coin comes up heads, Manuel will be given three million dollars. If the coin comes up tails, Manuel has to pay one million dollars. Should Manuel play the game?

   (a) Yes, because the expected value for Manuel is positive.
   (b) No, because the expected value for Manuel is negative.
   (c) Yes, because the standard deviation is large.
   (d) No, because the standard deviation is large.

General Probability Rules & Conditional Probability

94. In a certain semester, 500 students enrolled in both Calculus I and Physics I. Of these students, 82 got an A in calculus, 73 got an A in physics, and 42 got an A in both courses. Which of the following probabilities is the smallest? The probability that a randomly chosen student

(a) got an A in at least one of the two courses.
(b) got less than an A in at least one of the two courses.
(c) got an A in both of the two courses.
(d) got an A in calculus but not in physics.
(e) got an A in physics but not calculus.

95. Three cards are placed in a hat–one card is blue on both sides, one card is red on both sides, and one card has one side blue and one side red. A card is drawn at random from the hat and you see that one side is blue. What is the probability that the other side is also blue?

    (a) 1/3
    (b) 1/2
    (c) 2/3

96. Consider tossing a fair coin, that is, one that comes up heads half of the time and tails half of the time. Let A be the event "the first toss is a head," B be the event "the second toss is a tails," C be the event "the two outcomes are the same," D be the event "two heads turn up." Which of the following pairs of events is not independent?

    (a) A and B
    (b) A and C
    (c) A and D

97. Suppose A is the event that it rains today and B is the event that I brought my umbrella into work today. What is wrong with the following argument? "These events are independent because bringing an umbrella to work doesn't affect whether or not it rains today."

    (a) These events are not independent, because one's decision of bringing an umbrella is dependent on the likelihood of rain. (However, rain is definitely not dependent on one carrying an umbrella although Murphy's Law might prove the opposite.)
    (b) Although bringing an umbrella to work doesn't cause it to rain, given that you've brought your umbrella to work, the probability that it's a rainy day is higher than the chance of rain on any random day.
    (c) These events are independent because the probability of bringing an umbrella to work doesn't affect the probability of the event its rains today and vice versa.
    (d) It is false because the fact that it is raining today means that it was probably predicted to rain. If you checked that prediction then you would be more likely to bring in an umbrella making the events linked.

98. Assume that two events A and B are independent events. Which of the following statements is *false*?

   (a) $P(A \text{ and } B) = P(A) * P(B)$
   (b) $P(B|A) = [P(A|B) * P(B)]/P(A|B)$
   (c) A and B are mutually exclusive events.
   (d) $P(A|B) * P(B|A) = P(A \text{ and } B)$

99. Through accounting procedures, it is known that about 10% of the employees in a store are stealing. The managers would like to fire the thieves, but their only tool in distinguishing them from the honest employees is a lie detector test that is only 90% accurate. That is, if an employee is a thief, he or she will fail the test with probability 0.9, and if an employee is not a thief, he or she will pass the test with probability 0.9. If an employee fails the test, what is the probability that he or she is a thief?

   (a) 90%
   (b) 75%
   (c) 66 2/3%
   (d) 50%

100. A recent article in the Oklahoma Daily suggested that marijuana is a gateway drug for harder drug use. Suppose we have the following "facts". When asked, 90% of current "hard drug" users admit previously using marijuana; 40% of the general population admit using marijuana at some point during their lives; and 20% of the general population admit to using "hard drugs" at some point in their life. Given these three facts, what is the conditional probability of "hard drug" use given prior marijuana usage?

   (a) 0.16
   (b) 0.20
   (c) 0.25
   (d) 0.45
   (e) 0.90

101. A recent article in the Oklahoma Daily suggested that marijuana is a gateway drug for harder drug use. The following fact which we will take as accurate - was used to support their argument: 9 out of 10 of "hard drug" users have previously used marijuana. Additionally, the newspaper also reported that 4 out of every 10 persons in the general population have admitted using marijuana and that 2 out of 10 persons in the general population have admitted partaking of "harder" drugs.

   You now find out that one of your children has used marijuana. What is the probability of your child subsequently using some "hard drug" based on the information presented above?

(a) 0.16
(b) 0.20
(c) 0.25
(d) 0.45
(e) 0.90

102. A cab was involved in a hit and run accident at night. Only two cab companies, the Transporter and the Rock, operate in the city. You are given the following data:

(a) 85% of the cabs in the city are Transporters and 15% are Rocks.
(b) A witness identified the cab as a Rock. The court tested the reliability of the witness under the same circumstances that existed on the night of the accident and concluded that the witness correctly identified each one of the two cabs 80% of the time and failed 20% of the time.

What is the probability that the cab involved in the accident was indeed a Rock?

(a) 0.75
(b) 0.41
(c) 0.27
(d) 0.63
(e) 0.80

103. The following contingency table/two-way table classifies the members of a certain government into political party (Liberal or Conservative) and whether they support or oppose the spending bill that is currently up for adoption.

	Support	**Oppose**	**Total**
Liberal	47	11	58
Conservative	14	35	49
Total	61	46	107

Imagine randomly selecting one member of the government. Let L, C, S, and O denote the events of selecting a liberal, a conservative, a bill supporter, and a bill opposer, respectively. Find $P(C \,\&\, S)$.

(a) $\dfrac{14}{61}$
(b) $\dfrac{14}{49}$
(c) $\dfrac{14}{107}$

(d) None of the above

104. The following contingency table/two-way table classifies the members of a certain government into political party (Liberal or Conservative) and whether they support or oppose the spending bill that is currently up for adoption.

	Support	**Oppose**	**Total**
Liberal	47	11	58
Conservative	14	35	49
Total	61	46	107

Imagine randomly selecting one member of the government. Let L, C, S, and O denote the events of selecting a liberal, a conservative, a bill supporter, and a bill opposer, respectively. Find $P(S|L)$.

(a) $\frac{47}{61}$

(b) $\frac{47}{58}$

(c) $\frac{47}{107}$

(d) None of the above

105. A sample of sports fans from Canada and the Unites States were asked whether they would prefer to attend a professional basketball game or a professional (ice) hockey game. The following table gives a joint probability distribution for the sample.

	Basketball	**Hockey**
Canada	0.026	0.312
United States	0.512	0.150

Imagine randomly selecting one member of this sample. Let C, S, B, and H denote the events of selecting someone from Canada, someone from the United States, someone who chose basketball, and someone who chose hockey, respectively. Find $P(C|H)$.

(a) 0.312

(b) 0.675

(c) 0.923

(d) None of the above

106. A woman is the victim of a homicide, and her husband is on trial for her murder. It is known that prior to her murder, her husband had verbally threatened to kill her. During the trial, the defense attorney tells the court, "Only 1% of all men who threaten to kill their wives actually go on to kill them." There are, of course, many other pieces of evidence presented in the trial, but let's focus on this statement made by the defense attorney. TRUE or FALSE: The attorney's statement is a significant piece of evidence in favor of the man's innocence.

    (a) True, and I am very confident.
    (b) True, and I am not very confident.
    (c) False, and I am not very confident.
    (d) False, and I am very confident.

Chapter 5: Sampling Distributions

Sampling Distributions for Counts and Proportions

107. Consider the following experiment. On a Friday night, a highway patrol officer sets up a roadblock and stops 100 drivers. A given driver is considered a success if he or she is wearing a seat belt; the driver is considered a failure otherwise. Can we consider this experiment a binomial experiment?

    (a) Yes
    (b) No

108. Consider the following experiment. A particular car club has 100 members, 70 of which regularly wear their seat belts and 30 of which do not. Ten of these members are selected at random without replacement as they leave a car show. A given driver is considered a success if he or she is wearing a seat belt. The driver is considered a failure otherwise. Can we consider this experiment a binomial experiment?

    (a) Yes
    (b) No

109. In 1938, Duke University researchers Pratt and Woodruff conducted an experiment looking for evidence of ESP (extrasensory perception). In the experiment, students were presented with five standard ESP symbols (square, wavy lines, circle, star, cross). The experimenter shuffled a desk of ESP cards, each of which had one of the five symbols on it. The experimenter drew a card from this deck, looked at it, and concentrated on the symbol on the card. The student would then guess the symbol, perhaps

by reading the experimenter's mind. This experiment was repeated with 32 students for a total of 60,000 trials. The students were correct 12,489 times.

If the students were selecting one of the five symbols as random, the probability of success would be $p = 0.2$ and we would expect the students to be correct 12,000 times out of 60,000. Should we write off the observed excess of 489 as nothing more than random variation?

(a) Yes
(b) No

110. For which of the following probabilities would the binomial distribution be appropriate?

(a) The probability of a randomly selected professional basketball player making half of his free throws throughout a regular 82-game NBA season.
(b) The probability that a randomly selected student from a randomly selected high-school within the greater New York City metropolitan area will be accepted to an elite University.
(c) The probability that a randomly selected engineering student from a specific University will take at least 3 attempts to pass the licensure exam.
(d) Two of the above are appropriate for the binomial distribution.
(e) All of the above are appropriate for the binomial distribution.
(f) None of the above is appropriate for the binomial distribution.

111. Suppose a family is randomly selected from among all families with 3 children. What is the probability that the family has exactly one boy? You may assume that Pr(boy) = Pr(girl) for each birth.

(a) 1/8
(b) 1/6
(c) 1/3
(d) 3/8
(e) 1/2
(f) 5/6
(g) 7/8

112. Suppose a family is randomly selected from among all families with 4 children. What is the probability that the family has exactly two boys? You may assume that Pr(boy) = Pr(girl) for each birth.

(a) 1/24

(b) 1/16
(c) 1/6
(d) 3/8
(e) 1/2

113. There are five true/false questions on a quiz. If a student guesses randomly, what is the probability of getting exactly four questions right?

(a)
(b) 1/32
(c) 1/10
(d) 5/32
(e) 1/5
(f) None of the above

114. We roll two dice. What is the probability of getting exactly one four?

(a)
(b) 1/36
(c) 1/18
(d) 5/6
(e) 5/36
(f) 5/18
(g) None of the above

115. To measure the success of the latest treatment for iPod-related deafness among young adults, researchers measured the sound sensitivity of 100 young adults by having them stand 20 feet away from a speaker playing "Slim Whitman Favorite Hits." It was found that 35% of the sample could not repeat any song lyrics from the CD. What is the mean of this distribution?

(a) (20)(.35)
(b) (20)(.65)
(c) (20)(.35)(.65)
(d) (.35)(.65)
(e) (100)(.35)
(f) (100)(.65)

(g) (100)(.35)(.65)

(h) insufficient information

116. To measure the success of the latest treatment for iPod-related deafness among young adults, researchers measured the sound sensitivity of 100 young adults by having them stand 20 feet away from a speaker playing "Slim Whitman Favorite Hits." It was found that 35% of the sample could not repeat any song lyrics from the CD. What is the variance of this distribution?

(a) (20)(.35)

(b) (20)(.65)

(c) (20)(.35)(.65)

(d) (.35)(.65)

(e) (100)(.35)

(f) (100)(.65)

(g) (100)(.35)(.65)

(h) insufficient information

The Sampling Distribution of a Sample Mean

117. Jamie randomly selects 25 houses that are for sale in the U.S. The shape of the distribution of their prices is probably

(a) significantly right-skewed

(b) approximately bell-shaped

(c) significantly left-skewed

118. Jamie randomly selects 100 houses that are for sale in the U.S. The shape of the distribution of their prices is probably

(a) significantly right-skewed

(b) approximately bell-shaped

(c) significantly left-skewed

119. Jamie asks 25 of her friends to each randomly select 4 houses that are for sale in the U.S. and average their 4 prices together. The distribution of these 25 averages is probably

(a) significantly right-skewed
(b) approximately bell-shaped
(c) significantly left-skewed

120. Jamie asks 100 of her friends to each randomly select 4 houses that are for sale in the U.S. and average their 4 prices together. The distribution of these 100 averages is probably

(a) significantly right-skewed
(b) approximately bell-shaped
(c) significantly left-skewed

121. Jamie asks 25 of her friends to each randomly select 50 houses that are for sale in the U.S. and average their 50 prices together. The distribution of these 25 averages is probably

(a) significantly right-skewed
(b) approximately bell-shaped
(c) significantly left-skewed

122. Jamie asks 100 of her friends to each randomly select 50 houses that are for sale in the U.S. and average their 50 prices together. The distribution of these 100 averages is probably

(a) significantly right-skewed
(b) approximately bell-shaped
(c) significantly left-skewed

123. Why is there a μ in the symbol $\mu_{\bar{x}}$, which is used to denote the mean of the sampling distribution of the sample mean?

(a) Strictly speaking, the correct symbol is $\bar{x}_{\bar{x}}$, but $\mu_{\bar{x}}$ is used for simplicity.
(b) The μ refers to a parameter of the original population.
(c) The distribution whose mean is being taken consists of all sample means.

124. A physical therapy class has 10 students. The lightest student weights 110 pounds, the heaviest student weighs 240 pounds, the median weight of the 10 students is 140 pounds, and the mean weight of the 10 students is 160 pounds. Every student in the class pairs up with another student. In each pair, the two students find the mean of their two weights and then enter the mean into a spreadsheet on a computer in the classroom. Then the students pair off with different partners and again find the mean weight of their pair and type the mean into the spreadsheet. The students keep doing this until each student has been partnered with every other student. What is the most precise thing that can be said about the mean of all the numbers that were typed into the spreadsheet?

(a) It is between 110 pounds and 240 pounds.

(b) It is between 110 pounds and 240 pounds, but is most likely between 140 pounds and 160 pounds.

(c) It is 160 pounds.

125. A physical therapy class has 10 students. The lightest student weights 110 pounds, the heaviest student weighs 240 pounds, the median weight of the 10 students is 140 pounds, and the mean weight of the 10 students is 160 pounds. Every student in the class pairs up with another student. In each pair, the two students find the mean of their two weights and then enter the mean into a spreadsheet on a computer in the classroom. Then the students pair off with different partners and again find the mean weight of their pair and type the mean into the spreadsheet. The students keep doing this until each student has been partnered with every other student. The standard deviation of all the numbers that were typed into the spreadsheet is calculated. Then the entire experiment is repeated, except this time the students get into groups of five, taking the mean of all five weights, and keep doing this until every possible group of five students has recorded its mean weight. How will the standard deviation of all of the 5-student means compare to the standard deviation of all the 2-student means?

(a) It will be smaller.

(b) It will be the same.

(c) It will be larger.

(d) We would need to know the weights of all 10 students to answer this.

126. Your statistics professor says to you, “If you can guess a certain quantity within 7 points of its true value, I will give you some extra credit.” Which quantity would you prefer to guess?

(a) the score on next week’s exam of a randomly selected student

(b) the mean of the scores of all the students in the class on next week’s exam

127. The finishing times in a certain race are normally distributed with a mean of 25 minutes and a standard deviation of 4 minutes. What percentage of the samples of 4 finishers have means less than 21 minutes?

(a) Less than 1%
(b) 2.5%
(c) 16%
(d) None of the above is even close.

Chapter 6: Introduction to Inference

Estimating with Confidence

128. The fundamental concept underlying statistical inference is that

(a) through the use of sample data we are able to draw conclusions about a sample from which the data were drawn.
(b) through the examination of sample data we can derive appropriate conclusions about a population from which the data were drawn.
(c) when generalizing results to a sample we must make sure that the correct statistical procedure has been applied.
(d) Two of the above are true.
(e) All of the above are true.

129. A 95% confidence interval is an interval calculated from

(a) sample data that will capture the true population parameter for at least 95% of all samples randomly drawn from the same population.
(b) population data that will capture the true population parameter for at least 95% of all samples randomly drawn from the same population.
(c) sample data that will capture the true sample statistic for at least 95% of all samples randomly drawn from the same population.
(d) population data that will capture the true sample statistic for at least 95% of all samples randomly drawn from the same population.

130. A 95% confidence interval has been constructed around a sample mean of 28. The interval is (21, 35). Which of the following statement(s) is true?

(a) The margin of error in the interval is 7.

(b) 95 out of 100 confidence intervals constructed around sample means will contain the true population mean.

(c) The interval (21,35) contains the true population mean.

(d) Both (a) and (b) are true.

(e) (a), (b), and (c) are true.

131. A 95% confidence interval for the mean of a population is given as (6.85, 7.61). Is it correct to say that there is a 95% chance that μ is between 6.85 and 7.61?

(a) Yes

(b) No

132. Is it correct to say the following? If the process of selecting a sample of size 30 and then computing the corresponding 95% confidence interval is repeated 100 times, 95 of the resulting intervals will include μ.

(a) Yes

(b) No

133. A 95% confidence intervals for birthweights is found to be (6.85, 7.61). Is it correct to say that 95% of all birth weights will be between 6.85 and 7.61 pounds?

(a) Yes

(b) No

(c) About 95% of all birth weights will be in this range.

134. Suppose that a random sample of size 60 resulted in a 90% confidence interval for the proportion of students who carry more than 2 credit cards of (0.52, 0.76). Which of the following is a correct interpretation of the 90% confidence level?

(a) 90% of the time the population proportion will be between 0.52 and 0.76

(b) The method used to construct the interval will produce an interval that includes the value of the population proportion about 90% of the time in repeated sampling.

(c) If 100 different random samples of size 60 from this population were each used to construct a 90% confidence interval, 90 of them will contain the value population proportion.

(d) The probability that the population proportion is between 0.52 and 0.76 is 0.90.

135. Suppose you construct a 95% confidence interval from a random sample of size $n = 20$ with sample mean 100 taken from a population with unknown mean μ and known standard deviation $\sigma = 10$, and the interval is fairly wide. Which of the following conditions would NOT lead to a narrower confidence interval?

   (a) If you decreased your confidence level
   (b) If you increased your sample size
   (c) If the sample mean were smaller
   (d) If the population standard deviation were smaller

136. Which is wider, an 80% confidence interval, or a 90% confidence interval with both of them made from the same set of numerical data?

   (a) An 80% confidence interval is wider than a 90% confidence interval.
   (b) A 90% confidence interval is wider than an 80% confidence interval.
   (c) This depends on the data.

137. Each individual in a random sample of 40 cell phone users was asked how many minutes of airtime he or she used in a typical month. The data was then used to construct a 99% confidence interval for the mean monthly number of minutes of air time used. The confidence interval was $(207, 293)$. Which of the following could be the 95% confidence interval constructed using this same sample?

   (a) $(200, 300)$
   (b) $(218, 282)$
   (c) $(227, 313)$

138. Suppose that for babies born in the United States, birth weight is normally distributed about some unknown mean μ with standard deviation $\sigma = 1.06$ pounds. What is the minimum sample size necessary to ensure that the resulting 99% confidence interval has a width of at most 0.5?

   (a) 70
   (b) 119
   (c) 120
   (d) 140

Tests of Significance

139. Child and Protective Services, a branch of the Department of Health and Human Services is investigating the monthly average number of children in foster care over the last several years. They are interested in seeing if the average is dropping from 235 children per month in 2001.

     The null hypothesis for this problem would be:

     (a) $H_0 : \mu < 235$
     (b) $H_0 = 235$
     (c) $H_0 : p = 235$
     (d) $H_0 : \mu = 235$
     (e) None of the above

140. Which of the following pairs gives a legitimate null and alternative hypothesis for carrying out a hypothesis test?

     (a) $H_0 : \pi = 0.4 \quad H_a : \pi > 0.6$
     (b) $H_0 : \mu < 80 \quad H_a : \mu > 80$
     (c) $H_0 : \mu = 80 \quad H_a : \mu < 80$
     (d) $H_0 : \bar{x} = 25 \quad H_a : \bar{x} > 25$

141. A grocery store manager is interested in determining if the proportion of customers who pay by credit card at his store is greater than the reported national figure of 0.10. What hypotheses should the store manager test?

     (a) $H_0 : \mu > 0.1 \quad H_a : \mu = 0.1$
     (b) $H_0 : \mu = 0.1 \quad H_a : \mu > 0.1$
     (c) $H_0 : \pi = 0.1 \quad H_a : \pi > 0.1$
     (d) $H_0 : \pi > 0.1 \quad H_a : \pi = 0.1$
     (e) $H_0 : \pi > 0.1 \quad H_a : \pi < 0.1$

142. A drug company believes their newest drug for controlling cardiac arrhythmias is more effective and has less side effects than the current drug being used in the market. They submit their new drug to the FDA for a clinical trial to assess the efficacy of their drug in comparison to the current drug. What is the most appropriate null hypothesis for this clinical trial?

(a) H_0: Efficacy of new drug = Efficacy of old drug
(b) H_0: Efficacy of new drug > Efficacy of old drug
(c) H_0: Efficacy of new drug < Efficacy of old drug
(d) H_0: Efficacy of new drug $\neq$ Efficacy of old drug
(e) None of the above

143. If the conclusion in a hypothesis test is to fail to reject H_0, we can conclude that there is strong evidence that the null hypothesis is true.

(a) True, and I am very confident.
(b) True, and I am not very confident.
(c) False, and I am not very confident.
(d) False, and I am very confident.

144. If we reject the null hypothesis, does this mean we should accept the alternative hypothesis?

(a) Yes, and I am very confident.
(b) Yes, and I am not very confident.
(c) No, and I am not very confident.
(d) No, and I am very confident.

145. The p-value tells us the probability that the null hypothesis is true.

(a) True, and I am very confident.
(b) True, and I am not very confident.
(c) False, and I am not very confident.
(d) False, and I am very confident.

146. The p-value tells us the probability that our result is due to random chance.

(a) True, and I am very confident.
(b) True, and I am not very confident.
(c) False, and I am not very confident.
(d) False, and I am very confident.

Use and Abuse of Tests

147. Robert is asked to conduct a clinical trial on the comparative efficacy of Aleve versus Tylenol for relieving the pain associated with muscle strains. He creates a carefully controlled study and collects the relevant data. To be most informative in his presentation of the results, Robert should report

    (a) whether a statistically significant difference was found between the two drug effects.

    (b) a P-value for the test of no drug effect.

    (c) the mean difference and the variability associated with each drug's effect.

    (d) a confidence interval constructed around the observed difference between the two drugs.

148. A P-value represents

    (a) the probability, given the null hypothesis is true, that results like these could have been obtained purely on the basis of chance alone.

    (b) the probability, given the alternative hypothesis is true, that the results could have been obtained purely on the basis of chance alone.

    (c) the probability that the results could have been obtained purely on the basis of chance alone.

    (d) Two of the above are proper representations of a P-value.

    (e) None of the above is a proper representation of a P-value.

149. Two studies investigating the effect of motivation upon job performance found different results. With the exception of the sample size the studies were identical. The first study used a sample size of 500 and found statistically significant results, whereas the second study used a sample size of 100 and could not reject the null hypothesis. Which of the following is true?

    (a) The first study showed a larger effect than the second.

    (b) The first study was less biased than the second study for estimating the effect size because of the larger sample size.

    (c) The first study results are less likely to be due to chance than the second study results.

    (d) Two of the above are true.

    (e) All of the above are true.

Power and Inference as a Decision

150. We want to investigate whether people in Pottersville have more expensive cars than people in Baileytown. We gather a sample of people from both towns and compute the average values of the cars in each. Suppose that in reality the residents of Pottersville have more expensive cars, on average, but our sample sizes are not large enough for us to detect this difference. What type of error did we make?

    (a) Type I error
    (b) Type II error
    (c) Not enough information is given.

151. We want to investigate whether the trout in lake Able are the same size, on average, as the trout in lake Baker. We gather a sample of trout from each lake, and we find that the trout in lake Able are larger by a statistically significant amount. However, in reality both populations of trout have the same average size, and we just happened to get a sample from Able that were unusually large, due to bad luck. What type of error did we make?

    (a) Type I error
    (b) Type II error
    (c) Not enough information is given.

152. The manager of a university computing help line is trying to decide whether to hire additional staff. She has decided to hire if there is evidence that the average time callers to the help line must wait on hold before receiving assistance is greater than 5 minutes. She decides to collect data in order to test $H_0 : \mu = 5$ versus $H_a : \mu > 5$ where μ is the mean time on hold. From the callers' perspective, which type of error would be more serious?

    (a) Type I error
    (b) Type II error
    (c) Both types of error would be considered equally serious

153. Suppose that the P-value in a hypothesis test is 0.08. If the significance level for the test is $\alpha = 0.05$, which of the following is the appropriate decision?

    (a) Fail to reject H_0
    (b) Reject H_0

(c) There is not enough information given to know whether or not H_0 should be rejected.

154. In order to investigate a claim that the average time required for the county fire department to respond to a reported fire is greater than 15 minutes, county staff determined the response times for 40 randomly selected fire reports. The data was used to test $H_0 : \mu = 15$ versus $H_a : \mu > 15$ and the computed P-value was 0.12. If a 0.05 level of significance is used, what conclusions can be drawn?

(a) There is convincing evidence that the mean response time is 15 minutes (or less).

(b) There is convincing evidence that the mean response time is greater than 15 minutes.

(c) There is not convincing evidence that the mean response time is greater than 15 minutes.

155. Carol reports a statistically significant result ($P < 0.02$) in one of her journal articles. The editor suggests that because of the small sample size of the study ($n = 20$), the result cannot be trusted and she needs to collect more data before the article can be published. He is concerned that the study has too little power. How would you respond to the editor?

(a) The study has enough power to detect the effect since the significant result was obtained.

(b) Because the sample size so small, increasing the sample size to 200 should ensure sufficient power to detect a small effect.

(c) Setting the $\alpha = 0.01$ would be an alternative to collecting more data.

(d) Because the P-value is so close to $\alpha = 0.05$, the effect size is likely to be small and hence more information is needed.

Chapter 7: Inference for Distributions

Inference for the Mean of a Population

156. If you are testing two groups of individuals to see if they differ in regards to their working memory capacity, your alternative hypothesis would be that the two groups

(a) differ significantly in terms of working memory capacity.

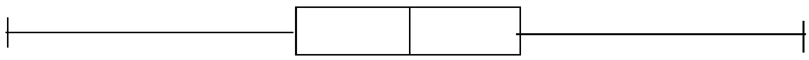
(b) differ in terms of working memory capacity.

(c) differ, but not significantly, in terms of working memory capacity.

(d) do not differ in terms of working memory capacity.

(e) do not differ significantly in terms of working memory capacity.

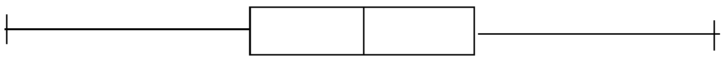
157. This box plot is for a sample that accurately represents a normal distribution:

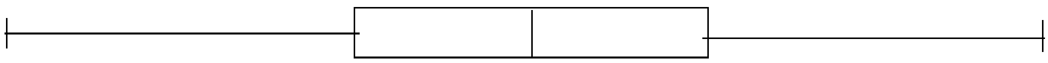


Which of the following box plots is for a sample that represents a Student's t-distribution with the same standard deviation and sample size as the normal distribution above?

(A) 

(B) 

(C) 

(D) 

(E) Two from (A)-(D) are correct.

(F) Three from (A)-(D) are correct.

(G) All from (A)-(D) are correct.

158. A random sample of 25 observations, with a mean of 44.4 and a sample standard deviation of 3.5, is drawn from a population that is approximately normally distributed. If one sets up a hypothesis test with population mean equal to 43 against an alternative that the population mean is not 43, using $\alpha = 0.01$, what is the value of the test statistic?

(a) 2.000

(b) 2.576

(c) 2.797

(d) 2.857

(e) 10.000

159. A random sample of 25 observations, with a mean of 44.4 and a sample standard deviation of 3.5, is drawn from a population that is approximately normally distributed. If one sets up a hypothesis test that the population mean is equal to 43 against an alternative that the population mean is not 43, using $\alpha = 0.01$, what is the 0.01 significance point (critical value) from the appropriate distribution?

(a) 2.576.
(b) 2.797.
(c) -2.576.
(d) -2.797.
(e) Both (a) and (c) are correct.
(f) Both (b) and (d) are correct.

160. A random sample of 25 observations, with a mean of 44.4 and a sample standard deviation of 3.5, is drawn from a population. If one sets up a hypothesis test that the population mean is equal to 43 against an alternative that the population mean is not 43, using $\alpha = 0.05$, what is the 0.05 significance point (critical value) from the appropriate distribution?

(a) 1.96.
(b) 2.064.
(c) -1.96.
(d) -2.064.
(e) None of the above

161. A random sample of 25 observations, with a mean of 44.4 and a sample standard deviation of 3.5, is drawn from a population that is approximately normally distributed. If one sets up a hypothesis test with population mean equal to 43 against an alternative that the population mean is not 43, using $\alpha = 0.01$, does one reject the null hypothesis and why?

(a) Yes, the test statistic is larger than the tabled critical value.
(b) No, the test statistic is larger than the tabled critical value.
(c) Yes, the test statistic is smaller than the tabled critical value.
(d) No, the test statistic is smaller than the tabled critical value.
(e) insufficient information

162. In a random sample of 2013 adults, 1283 indicated that they believe that rudeness is a more serious problem than in past years. Which of the test statistics shown below would be appropriate to determine if there is sufficient evidence to conclude that more than three-quarters of U.S. adults believe that rudeness is a worsening problem?

(a) $\dfrac{\hat{p} - .5}{\sqrt{(.5)(1 - .5)/2013}}$

(b) $\dfrac{\hat{p} - .75}{\sqrt{(.75)(1 - .75)/2013}}$

(c) $\dfrac{\bar{x} - .75}{\sqrt{s/2013}}$

163. A climate researcher sets up an experiment that the mean global temperature is $\mu = 60°$ F, looking for an indication of global warming in a climate model projection. For the year 2050, the series of 10 models predict an average temperature of 65° F. A standard one-tailed t-test is run on the data. Then the power of the test

(a) increases as μ decreases.

(b) remains constant as μ changes.

(c) increases as μ increases.

(d) decreases as μ increases.

164. Kim's husband is in residency to become a medical doctor. He doesn't seem to get much sleep. Neither do his resident friends at the hospital. Kim hypothesizes that medical residents get, on average, less than 5 hours of sleep each day. She conducts a hypothesis test and gets statistically significant results. Which of the following might be a reasonable summary of the results?

(a) Accept the null hypothesis. At the 5% significance level, the data provide sufficient evidence to conclude that the mean hours of sleep for medical residents is less than 5 hours.

(b) Accept the null hypothesis. At the 5% significance level, the data do not provide sufficient evidence to conclude that the mean hours of sleep for medical residents is less than 5 hours.

(c) Reject the null hypothesis. At the 5% significance level, the data provide sufficient evidence to conclude that the mean hours of sleep for medical residents is less than 5 hours.

(d) Reject the null hypothesis. At the 5% significance level, the data do not provide sufficient evidence to conclude that the mean hours of sleep for medical residents is less than 5 hours.

165. Kim's husband is in residency to become a medical doctor. He doesn't seem to get much sleep. Neither do his resident friends at the hospital. Kim hypothesizes that medical residents get, on average, less than 5 hours of sleep each day. She conducts a hypothesis test but does not get statistically significant results. Which of the following might be a reasonable summary of the results?

(a) Do not reject the null hypothesis. At the 5% significance level, the data provide sufficient evidence to conclude that the mean hours of sleep for medical residents is 5 hours.

(b) Do not reject the null hypothesis. At the 5% significance level, the data provide sufficient evidence to conclude that the mean hours of sleep for medical residents is less than 5 hours.

(c) Do not reject the null hypothesis. At the 5% significance level, the data do not provide sufficient evidence to conclude that the mean hours of sleep for medical residents is less than 5 hours.

(d) Reject the null hypothesis. At the 5% significance level, the data do not provide sufficient evidence to conclude that the mean hours of sleep for medical residents is less than 5 hours.

166. The state legislature ordered that a study be done to see whether the mean number of reported crimes at institutions of higher learning across the state differs from the national mean. A hypothesis test will be performed at the 1% significance level using the critical-value approach. What are the critical values?

(a) $\pm\frac{t_{0.01}}{2}$

(b) $\pm t_{\frac{0.01}{2}}$

(c) $\pm\frac{t_{0.99}}{2}$

(d) $\pm t_{\frac{0.99}{2}}$

167. Cody thinks that on average more than 30 students enter the gym between 12:00 p.m. and 12:30 p.m. on class days. If Cody performs a hypothesis test, a decrease in which of the following quantities—all other quantities remaining the same—would increase the probability of Cody's rejecting the null hypothesis?

(a) sample size

(b) sample variance

(c) sample mean

168. Cody thinks that on average more than 30 students enter the gym between 12:00 p.m. and 12:30 p.m. on class days. If Cody performs a hypothesis test, and gets a P-value of 0.02, what is the probability that more than 30 students will enter the gym between 12:00 pm and 12:30 pm on a randomly-selected class day?

(a) less than 0.02

(b) 0.02

(c) more than 0.02

(d) The answer cannot be determined from the information given.

169. A drug company claims that their new weight loss pill will cause obese people to lost an average of 15 pounds after six weeks of use. The null hypothesis is that the mean is 15 pounds, while the alternative hypothesis is that the mean is less than 15 pounds. We try out the pill on 75 obese people and find that, after six weeks, the mean weight loss is only 8.2 pounds. The P-value of our result is 0.00216. What do we conclude?

(a) This proves that the pill works as claimed by the drug company.

(b) This proves that the pill does not work as claimed by the drug company.

(c) The results are ambiguous, so we can draw no conclusions.

(d) None of the above

170. A tire company claims that their tires last an average of 30,000 miles. Our null hypothesis is that the claim is true, while our alternative hypothesis is that the tires last for an average of less than 30,000 miles. We test 50 of their tires finding that this sample lasts only for an average of 28,000 miles. We calculate a P-value of 0.389 for this result. What do we conclude?

(a) We conclude that the company's claim is correct.

(b) We conclude that the company's claim is incorrect.

(c) The results are ambiguous, so we can draw no conclusions.

(d) None of the above

171. Zoe wants to know the average height of trees in her city. She randomly selects thirty trees in her city and measures their heights, obtaining a mean of 37.1 feet and a standard deviation of 15.6 feet. Which statistical procedure should she perform? (Assume that all assumptions for the procedure are satisfied.)

(a) confidence interval for one mean with σ known (z-interval procedure)

(b) confidence interval for one mean with σ unknown (t-interval procedure)

(c) hypothesis test for one mean with σ known (z-test)

(d) hypothesis test for one mean with σ unknown (t-test)

172. Stan and Priscilla head up the family's traditional Christmas tamale making, in which everyone gets together and makes an unbelievable number of tamales. You can always count on Tio Carlos insisting that the tamales be big. Are the tamales this year averaging 3.5 ounces like they always do? (The standard deviation in weights each year is around 0.31 ounces.) Stan and Priscilla randomly select twenty of the tamales made so far and find that their mean weight is 3.53 ounces. Which statistical procedure should they perform? (Assume that all assumptions for the procedure are satisfied.)

    (a) confidence interval for one mean with σ known (z-interval procedure)
    (b) confidence interval for one mean with σ unknown (t-interval procedure)
    (c) hypothesis test for one mean with σ known (z-test)
    (d) hypothesis test for one mean with σ unknown (t-test)

173. David wants to know whether the mean brix level of Haden mangoes available in his city's supermarkets differs from 14. He randomly selects thirty Haden mangoes from his city's supermarkets and measures their brix levels, finding the mean to be 13.3, with a standard deviation of 1.8. Which statistical procedure should he perform? (Assume that all assumptions for the procedure are satisfied.)

    (a) confidence interval for one mean with σ known (z-interval procedure)
    (b) confidence interval for one mean with σ unknown (t-interval procedure)
    (c) hypothesis test for one mean with σ known (z-test)
    (d) hypothesis test for one mean with σ unknown (t-test)

Comparing Two Means

174. You are trying to decide whether to use a pooled t-test or a nonpooled t-test to compare two population means. Which of the following statements is not true?

    (a) If the population standard deviations are very different, the Type I error probability can be much larger than the specified level for the pooled t-test.
    (b) If the population standard deviations are somewhat different, the pooled t-test is preferable if the sample sizes are very different.
    (c) If the population standard deviations are nearly the same, the pooled t-test is more powerful than the nonpooled t-test.

175. Thirty people suffering from obesity sign up to be subjects in a clinical trial. The subjects are weighed and then given a promising new supplement to take daily for six weeks. At the end of the six weeks, the subjects are weighed again. Which test is more appropriate?

   (a) pooled t-test
   (b) paired t-test

176. Two methods are used to predict the shear strength for steel plate girders. Each method is applied to nine specific girders and the ratio of predicted load to observed load is calculated for each method and each girder. What kind of t-test should we use to compare these data?

   (a) Independent t-test
   (b) Paired t-test

177. Two catalysts are being analyzed to determine how they affect the mean yield of a chemical process. Catalyst 1 is used in the process eight times and the yield in percent is measured each time. Then catalyst 2 is used in the process eight times and the yield is measured each time. What kind of t-test should be used to compare these data?

   (a) Independent t-test
   (b) Paired t-test

178. Six river locations are selected and the zinc concentration is determined for both surface water and bottom water at each location. What kind of t-test should be used to compare these data?

   (a) Independent t-test
   (b) Paired t-test

179. We want to know if residents of Missoula have more life insurance than residents of Great Falls. We randomly select 100 Missoula residents and 100 Great Falls residents, finding that the Missoulians have an average of $125,000 of coverage, while the Great Falls residents have an average of $110,000 of coverage. We then use the t distribution to calculate a two-tailed p-value based on $H_0 : \mu_1 = \mu_2$. We find that $p = 0.16$. What do we conclude?

   (a) On average, Missoula residents have more life insurance than Great Falls residentes.
   (b) On average, Missoula residents have less life insurance than Great Falls residentes.

(c) On average, Missoula residents have the same amount of life insurance as Great Falls residentes.

(d) The average amount of life insurance that Missoula residents have is not different from Great Falls residents.

(e) This data doesn't tell us whether or not Missoula residents have more life insurance than Great Falls residents, on average.

(f) This is not the correct way of analyzing this data.

180. Was the weather last summer hotter than the previous summer? We look up the temperature at noon every day for a week in June during last summer, and for a week in June during the previous summer. We find that the average temperature from last summer was 81°, while the average temperature from the previous summer was 75°. We then use the t distribution to calculate a two-tailed p-value based on $H_0 : T_1 = T_2$. We find that $p = 0.0014$. What do we conclude?

(a) On average, last summer was hotter than the previous summer.

(b) On average, last summer was less hot than the previous summer.

(c) On average, last summer was equally hot as the previous summer.

(d) This data doesn't tell us whether or not last summer was hotter than the previous summer, on average.

(e) The average temperature last summer was different from the previous summer.

Chapter 8: Inference for Proportions

Inference for a Single Proportion

181. To estimate the proportion of students at a university who watch reality TV shows, a random sample of 50 students was selected and resulted in a sample proportion of .3. A 95% confidence interval for the proportion that watches reality TV would be ______ a 90% confidence interval.

(a) narrower than

(b) the same width as

(c) wider than

182. Suppose we wish to estimate the percentage of students who smoke marijuana at each of several liberal arts colleges. Two such colleges are StonyCreek (enrollment 5,000) and Whimsy (enrollment 13,000). The Dean of each college decides to take a random sample of 10% of the entire student population. The margin of error for a simple random sample of 10% of the population of students at each school will be

(a) smaller for Whimsy than for StonyCreek.
(b) smaller for StonyCreek than for Whimsy.
(c) the same for each school.
(d) insufficient information

183. Suppose we wish to estimate the percentage of people who speed while driving in a college town. We choose to sample the populations of Austin, TX (University of Texas) and Norman, OK (University of Oklahoma). We know that both cities have populations over 100,000 and that Austin is approximately 5 times bigger (in population) than Norman. We also expect the rates of speeding to be about the same in each city. Suppose we were to take a random sample of 1000 drivers from each city. The margin of error for a simple random sample of the population of drivers from each city will be

(a) smaller for the Austin sample than the Norman sample.
(b) smaller for the Norman sample than the Austin sample.
(c) the same for both samples.
(d) not possible to determine without more precise information about the population sizes.

184. A parachute manufacturer is concerned that the failure rate of 0.1% advertised by his company may in fact be higher. What is the null hypothesis for the test he would run to address his worries.

(a) $H_0 : \mu = 0.001$
(b) $H_0 : p > 0.001$
(c) $H_0 : \mu < 0.001$
(d) $H_0 : p = 0.001$

185. A parachute manufacturer is concerned that the failure rate of 0.1% advertised by his company may in fact be higher. A hypothesis test was run and the result was a P-value of 0.03333. The most likely conclusion the manufacturer might make is:

(a) My parachutes are safer than I claim.
(b) My parachutes are not as safe as I claim them to be.

(c) I can make no assumption of safety based on a statistical test.
(d) The probability of a parachute failure is 0.03333.
(e) Both (b) and (d) are true.

186. To explain the meaning of a P-value of 0.033, you could say:

(a) There is approximately a 96.7% chance of obtaining my sample results.
(b) Assuming the null hypothesis is accurate, results like those found in my sample should occur only 3.3% of the time.
(c) We can't say anything for sure without knowing the sample results.
(d) There is approximately a 3.3% chance of obtaining my sample results.

187. Suppose we have the results of a Gallup survey (simple random sampling) which asks participants for their opinions regarding their attitudes toward technology. Based on 1500 interviews, the Gallup report makes confidence statements about its conclusions. If 64% of those interviewed favored modern technology, we can be 95% confident that the percent of those who favored modern technology is

(a) 95% of 64%, or 60.8%
(b) 95% $\pm$ 3%
(c) 64%
(d) 64% $\pm$ 3%

188. A confidence interval for a proportion is constructed using a sample proportion of 0.5. If the sample proportion was 0.9 instead of 0.5, what would happen to the width of the resulting confidence interval?

(a) The new CI would be narrower.
(b) The new CI would have the same width.
(c) The new CI would be wider.

189. A sample needs to taken to answer the question Have you ever shoplifted? Assuming a random sample can be found, how many people would need to be polled to insure a margin of error of no more than 3% with 90% confidence?

(a) 1068
(b) 752
(c) 23
(d) None of the above

190. A sample needs to be taken to answer the question, "Have you jaywalked in the past month?" Assume that based on past studies we can be almost certain that the actual percentage of the population that jaywalks in any given month is between 70% and 90%. Assuming a random sample can be obtained, how many people would need to be polled to insure a margin of error of no more than 3% with 95% confidence?

(a) 1068
(b) 897
(c) 385
(d) None of the above

191. A sample needs to be taken to answer the question, "Have you voted in a local election in the past 10 years?" Assume that based on past studies we can be almost certain that the actual population percentage is between 45% and 75%. Assuming a random sample can be obtained, how many people would need to be polled to insure a margin of error of no more than 3% with 99% confidence?

(a) 1844
(b) 1825
(c) 1383
(d) None of the above

192. The margin of error is computed for a poll with a sample size of 50. Approximately what sample size would you need if you wanted to cut the margin of error in half?

(a) 25
(b) 100
(c) 200
(d) 400

193. Which of the following does *not* result in a larger margin of error?

(a) Increasing the confidence level
(b) Decreasing the sample size
(c) Having a larger population size

194. Terrance hypothesizes that less than 20% of the population has a favorable view of the latest flavor-of-the-month boy band. He conducts a hypothesis test and gets $P = 0.0001$. Which of the following might be a reasonable summary of the results?

(a) Accept the null hypothesis. The data provide sufficient evidence to conclude that less than 20% of the population has a favorable view of the band ($P = 0.0001$).

(b) Accept the null hypothesis. The data do not provide sufficient evidence to conclude that less than 20% of the population has a favorable view of the band ($P = 0.0001$).

(c) Reject the null hypothesis. The data provide sufficient evidence to conclude that less than 20% of the population has a favorable view of the band ($P = 0.0001$).

(d) Reject the null hypothesis. The data do not provide sufficient evidence to conclude that less than 20% of the population has a favorable view of the band ($P = 0.0001$).

Comparing Two Proportions

195. A two proportion z interval was constructed for the difference in the two population proportions, p_1 and p_2. The resulting 99% confidence interval was $(-0.004, 0.12)$. A conclusion that could be drawn is:

(a) There is no significant difference between p_1 and p_2.

(b) There is a significant difference between p_1 and p_2.

(c) The range of possible differences between the two proportions could be from a 0.4% difference with p_2 being larger up to a 12% difference with p_1 being larger.

(d) Both (a) and (c) are correct.

(e) Both (b) and (c) are correct.

Chapter 9: Analysis of Two-Way Tables

Inference for Two-Way Tables

196. In a 2×2 table of the frequency of sexual intercourse by age, we observe a chi-square (χ^2) statistic of 2.5. What should be the conclusion?

(a) There is observed evidence that sex and age are associated.

(b) There is little observed evidence of anything but a chance association.

(c) It is not possible to obtain an observed chi-square statistic this large.

(d) It would be unlikely to obtain an observed chi-square statistic this large.

(e) No conclusion is appropriate without sample size information.

197. Two quantitative variables can be either (linearly) correlated or not (linearly) correlated. Fill in the blank with a roughly analagous more general dichotomy, a dichotomy that applies to both quantitative and qualitative variables: "correlated is to not correlated" as __________________.

(a) "independent is to not independent"

(b) "associated is to not associated"

(c) "resistant is to not resistant"

198. TRUE or FALSE: Two quantitative variables that are not (linearly) correlated are independent.

(a) True, and I am very confident.

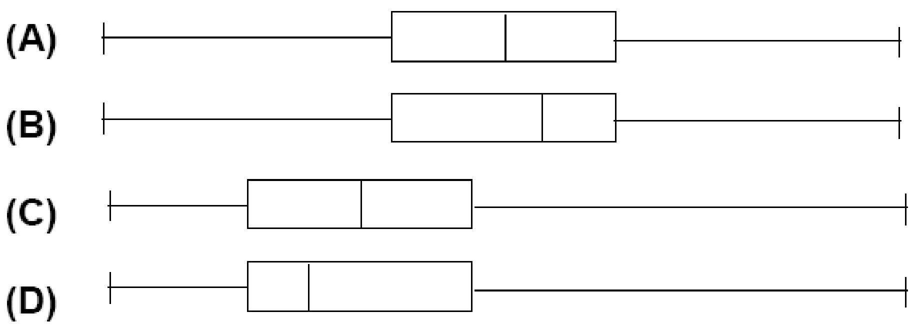
(b) True, and I am not very confident.

(c) False, and I am not very confident.

(d) False, and I am very confident.

Goodness of Fit

199. If a large sample were drawn from a chi-square (χ^2) distribution (with degrees of freedom ≤ 10) and accurately represented the population, which of the following is most likely to be a box plot of that sample?



(E) Two from (A)-(D) are correct.

(F) Three from (A)-(D) are correct.

(G) All from (A)-(D) are correct.

200. Which of the following statements concerning the Chi-Square Goodness-of-Fit test is true?

   (a) The test basically involves converting observed frequencies into relative frequencies so that, roughly speaking, they can be compared to established or hypothesized relative frequencies.
   (b) One explicit assumption of the test is that we have a large sample.
   (c) One explicit assumption of the test is that all of the expected frequencies be at least 1.
   (d) The test is usually two-tailed.

201. Note that a chi-square random variable with one degree of freedom is the square of a standard normal variable. Let X be a chi-square random variable with one degree of freedom. Find $P(X \geq 1)$.

   (a) 0.05
   (b) 0.16
   (c) 0.32
   (d) The answer cannot be determined from the information given.

Chapter 10: Inference for Regression

Simple Linear Regression

202. What is the most common rationale for significance testing of simple linear regression?

   (a) to test if the intercept is significantly large
   (b) to test is the slope of the regression line is positive
   (c) to test if the slope of the regression line is negative
   (d) to test if the slope is different from zero
   (e) to appease an editor or reviewer when publishing the results

203. Which of the following does *not* result in more accurate estimates for the slope of the regression line?

   (a) An increase in the sample size
   (b) An increase in the coefficient of determination
   (c) An increase in the variance of the observed x-values

(d) An increase in the variance of the observed y-values

204. You have performed a linear regression with age as the predictor variable and hours per week spent online as the response variable. You have also constructed a 95% confidence interval for the mean number of hours spent online for 21-year olds. The confidence interval is $(18, 28)$. Which of the following could be a 95% prediction interval for the number of hours spent online for a randomly selected 21-year old?

   (a) $(20, 26)$
   (b) $(20, 30)$
   (c) $(14, 32)$
   (d) all of the above
   (e) two of the above

205. You have performed a linear regression with age as the predictor variable and hours per week spent online as the response variable. You have also constructed a 95% confidence interval for the mean number of hours spent online for 21-year olds. The confidence interval is $(18, 28)$. Which of the following could be a 90% prediction interval for the number of hours spent online for a randomly selected 21-year old?

   (a) $(20, 26)$
   (b) $(18, 28)$
   (c) $(16, 30)$
   (d) all of the above
   (e) two of the above

Chapter 12: One-Way Analysis of Variance

One Way ANOVA

206. One-Way ANOVA is a generalization of which test?

   (a) pooled t-test
   (b) nonpooled t-test
   (c) paired t-test
   (d) multiple comparisons

207. The test statistic in a One-Way ANOVA basically measures the ratio of

(a) the variance within the samples to the differences among the standard deviations of the samples
(b) the differences among the standard deviations of the samples to the variance within the samples
(c) the variance within the samples to the differences among the means of the samples
(d) the differences among the means of the samples to the variance within the samples

208. For college students consider the variable "classification" which takes on the values "freshman," "sophomore," "junior," and "senior," and the variable "frequency of binge drinking," which takes on the values "never," "less than once per month," and "more than once per month." If you suspect an association between "classification" and "frequency of binge drinking," which statistical procedure should you perform? (Assume that all assumptions for the procedure are satisfied.)

(a) pooled t-test
(b) nonpooled t-test
(c) paired t-test
(d) confidence interval for two proportions
(e) hypothesis test for two proportions
(f) chi-square goodness-of-fit test
(g) chi-square independence test
(h) ANOVA

209. Your friend Myra is looking forward to seeing her two young cousins again this weekend. She says all they talk about is princesses. "They say the word 'princess' 10 times every five minutes," Myra chuckles. You tell Myra you think that's an exaggeration, and she agrees. Which statistical procedure could you both perform to see whether you are correct in your suspicion that Myra's statement was an exaggeration? (Assume that all assumptions for the procedure are satisfied.)

(a) hypothesis test for one mean
(b) pooled t-test
(c) nonpooled t-test
(d) hypothesis test for one proportion
(e) hypothesis test for two proportions
(f) chi-square goodness-of-fit test
(g) chi-square independence test

(h) ANOVA

210. You are in Washington D.C. working as an intern for a congresswoman. She asks you to get an estimate of the percentage of voters in her district who favor raising taxes on the rich. Which statistical procedure should you perform? (Assume that all assumptions for the procedure are satisfied.)

(a) confidence interval for one mean
(b) hypothesis test for one mean
(c) confidence interval for one proportion
(d) hypothesis test for one proportion
(e) confidence interval for two proportions
(f) hypothesis test for two proportions
(g) chi-square goodness-of-fit test
(h) chi-square independence test

211. You are wondering whether there is a difference in the average heights of men in Canada, the United States, and Mexico. Which statistical procedure should you perform? (Assume that all assumptions for the procedure are satisfied.)

(a) hypothesis test for one mean
(b) pooled t-test
(c) nonpooled t-test
(d) hypothesis test for one proportion
(e) hypothesis test for two proportions
(f) chi-square goodness-of-fit test
(g) chi-square independence test
(h) ANOVA

212. You just finished reading about the percentages of vehicles that got low gas mileage, medium gas mileage, and high gas mileage five years ago. You're wondering whether these percentages are different this year. Which statistical procedure should you use? (Assume that all assumptions for the procedure are satisfied.)

(a) hypothesis test for one mean
(b) pooled t-test
(c) nonpooled t-test
(d) hypothesis test for one proportion

(e) hypothesis test for two proportions
(f) chi-square goodness-of-fit test
(g) chi-square independence test
(h) ANOVA

213. Elliot is trying to deterime whether there is a difference in the average number of hours of sleep between adults who live in farming communities and adults who live in cities. He has just created boxplots of his sample data and notices that the cities data is much more spread out than the farming communities data. Which statistical procedure should Elliot use? (Assume that all assumptions for the procedure are satisfied.)

(a) hypothesis test for one mean
(b) pooled t-test
(c) nonpooled t-test
(d) paired t-test
(e) hypothesis test for two proportions
(f) chi-square goodness-of-fit test
(g) chi-square independence test
(h) ANOVA

Chapter 15: Nonparametric Tests

214. Which of the following best describes a situation in which you would use a nonparametric method (as opposed to another type of method)?

(a) The parameters of the population distribution are unknown.
(b) The parameters of the population distribution are not resistant.
(c) The population distribution is not normal.
(d) The shape of the population distribution is unknown.