

AP Statistics Midterm Review

Use the following to answer question 1.

There are 20 multiple-choice questions on an exam, each having responses a, b, c, or d. Suppose a student *guesses* the answer to each question, and the guesses from question to question are independent. Let X be the number of questions for which the student has the same answer as the person sitting next to him on his right.

- The distribution of X is
 - $B(20, .2)$.
 - $B(20, .25)$.
 - $B(4, .25)$.
 - impossible to determine unless the student sitting next to him is also guessing.
- Suppose X is a random variable with the $B(n = 9, p = 1/3)$ distribution. The probability X is either a 0 or a 1 is
 - 0.6667.
 - 0.3333.
 - 0.1431.
 - 0.1111.
- If X is $B(n = 9, p = 1/3)$, the standard deviation σ_X of X is
 - 6.
 - 3.
 - 2.
 - 1.414.
- In a test of ESP (extrasensory perception), the experimenter looks at cards that are hidden from the subject. Each card contains either a star, a circle, a wavy line, or a square. The subject tries to read the experimenter's mind and name the shape on the cards. What is the probability that the subject gets the first four correct before giving a wrong answer if he is just guessing?
 - 0.00293.
 - 0.00391.
 - 0.07910.
 - 0.31641.
- In a certain game of chance, your chances of winning are 0.2. If you play the game five times and outcomes are independent, the probability that you win all five times is
 - 0.6723.
 - 0.3277.
 - 0.04.
 - 0.00032.
- If X is $B(n = 9, p = 1/3)$, the mean μ_X of X is
 - 6.
 - 3.
 - 2.
 - 1.414.
- An airplane has a front and a rear door that are both opened to allow passengers to exit when the plane lands. The plane has 100 passengers seated. The number of passengers exiting through the front door should have
 - a binomial distribution with mean 50.
 - a binomial distribution with 100 trials but success probability not equal to 0.5.
 - a normal distribution with a standard deviation of 5.
 - none of the above.
- A multiple-choice exam has 100 questions, each with five possible answers. If a student is just guessing at all the answers, the probability that he or she gets more than 30 correct is
 - 0.3100.
 - 0.2000.
 - 0.1020.
 - 0.0062.

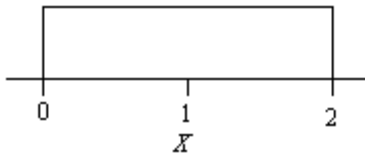
Use the following to answer question 9.

Let the random variable X represent the profit made on a randomly selected day by a certain store. Assume X is normal with a mean of \$360 and standard deviation \$50.

9. Referring to the information above, the value of $P(X > \$400)$ is
A) 0.2881. B) 0.8450. C) 0.7881. D) 0.2119.

Use the following to answer question 10.

The probability density of a random variable X is given in the figure below.



10. Referring to the information above, the probability that X is at least 1.5 is
A) 0. B) 1/4. C) 1/3. D) 1/2.
11. In a particular game, a ball is randomly chosen from a box that contains three red balls, one green ball, and six blue balls. If a red ball is selected you win \$2, if a green ball is selected you win \$4, and if a blue ball is selected you win nothing. Let X be the amount that you win. The expected value of X is
A) \$1. B) \$2. C) \$3. D) \$4.
12. I toss a fair coin a large number of times. Assuming tosses are independent, which of the following is true?
A) Once the number of flips is large enough (usually about 10,000) the number of heads will always be exactly half of the total number of tosses. For example, after 10,000 tosses I should have 5000 heads.
B) The proportion of heads will be about ?, and this proportion will tend to get closer and closer to ? as the number of tosses increases.
C) As the number of tosses increases, any long run of heads will be balanced by a corresponding run of tails so that the overall proportion of heads is ?.
D) All of the above.
13. Suppose we have a *loaded* die that gives the outcomes 1–6 according to the probability distribution

| | | | | | | |
|--------|-----|-----|-----|-----|-----|-----|
| X | 1 | 2 | 3 | 4 | 5 | 6 |
| $P(X)$ | 0.1 | 0.2 | 0.3 | 0.2 | 0.1 | 0.1 |

Note that for this die all outcomes are *not* equally likely, as they would be if this die were fair. If this die is rolled 6000 times, then \bar{x} , the sample mean of the number of spots on the 6000 rolls, should be about

- A) 3. B) 3.30. C) 3.50. D) 4.50.
14. Suppose we roll a red die and a green die. Let A be the event that the number of spots showing on the red die is 3 or less and B be the event that the number of spots showing on the green die is more than 3. The events A and B are
A) disjoint. B) complements. C) independent. D) reciprocals.

Use the following to answer question 15.

A stack of four cards contains two red cards and two black cards. I select two cards, one at a time, and do *not* replace the first card selected before selecting the second card.

15. Referring to the information above, consider the events

A = the first card selected is red

B = the second card selected is black

The events A and B are

- A) independent. B) disjoint. C) complements. D) none of the above.

Use the following to answer question 16.

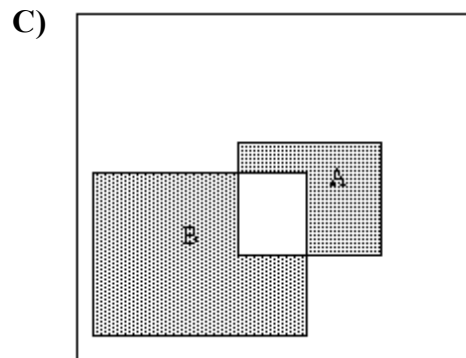
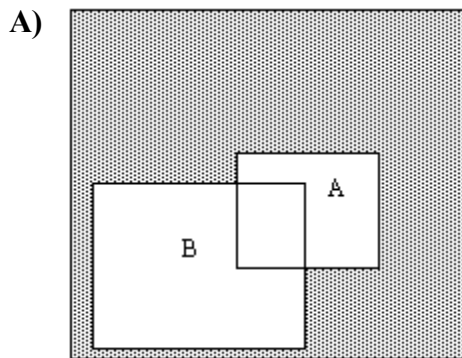
If you draw an M&M candy at random from a bag of the candies, the candy you draw will have one of six colors. The probability of drawing each color depends on the proportion of each color among all candies made. Assume the table below gives the probability that a randomly chosen M&M has each color.

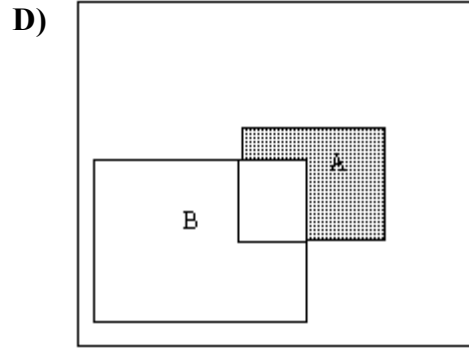
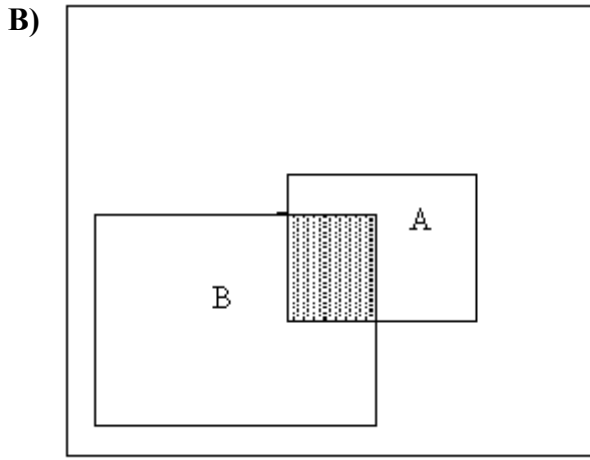
| Color | Brown | Red | Yellow | Green | Orange | Tan |
|-------------|-------|-----|--------|-------|--------|-----|
| Probability | .3 | .3 | ? | .1 | .1 | .1 |

16. The probability of drawing a yellow candy is

- A) .1.
B) .2.
C) .3.
D) impossible to determine from the information given.

17. A plumbing contractor puts in bids in on two large jobs. Let the event that the contractor wins the first contract be A and the event that the contractor wins the second contract be B. Which of the Venn diagrams on this and the next page has shaded the event that the contractor wins exactly one of the contracts?



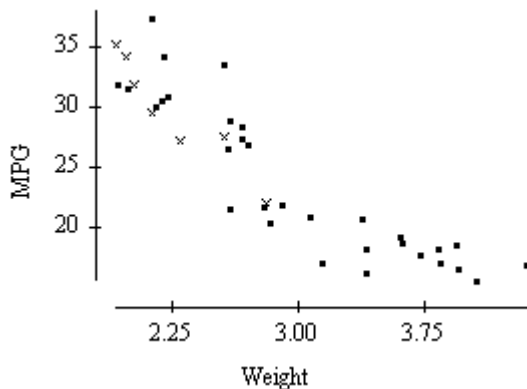


18. I toss a thumb tack 60 times and it lands point up on 35 of the tosses. The approximate probability of landing point up is
- A) 35. B) 0.35. C) 0.58. D) 0.65.

Use the following to answer question 19.

Ignoring twins and other multiple births, assume babies born at a hospital are independent events with the probability that a baby is a boy and the probability that a baby is a girl both equal to 0.5.

19. Referring to the information above, the probability that at least one of the next three babies is a boy is
- A) 0.125. B) 0.333. C) 0.75. D) 0.875.
20. Event A occurs with probability 0.8. The conditional probability that event B occurs given that A occurs is 0.5. The probability that both A and B occur
- A) is 0.3.
 B) is 0.4.
 C) is 0.8.
 D) cannot be determined from the information given.
21. The graph below plots the gas mileage (miles per gallon, or MPG) of various 1978 model cars versus the weight of these cars in thousands of pounds.



In the graph, the points denoted by the plotting symbol x correspond to cars made in Japan. From this plot, we may conclude that

- A) in 1978 there was little difference between Japanese cars and cars made in other countries.

- B) in 1978 Japanese cars tended to be lighter in weight than other cars.
- C) in 1978 Japanese cars tended to get poorer gas mileage than other cars.
- D) the plot is invalid. A scatterplot is used to represent quantitative variables, and the country that makes a car is a qualitative variable.

Use the following to answer question 22.

A television station is interested in predicting whether voters in its viewing area are in favor of federal funding for abortions. It asks its viewers to phone in and indicate whether they support/are in favor of or are opposed to this. Of the 2241 viewers who phoned in, 1574 (70.24%) were opposed to federal funding for abortions.

22. Referring to the information above, the sample obtained is
- A) a simple random sample.
 - B) a single-stage sample.
 - C) a census.
 - D) probably biased.

Use the following to answer question 23.

Researchers wish to determine if a new experimental medication will reduce the symptoms of allergy sufferers without the side effect of drowsiness. To investigate this question, the researchers give the new medication to 50 adult volunteers who suffer from allergies. Forty-four of these volunteers report a significant reduction in their allergy symptoms without any drowsiness.

23. Referring to the information above, the experimental units are
- A) the researchers.
 - B) the 50 adult volunteers.
 - C) the 44 volunteers who reported a significant reduction in their allergy symptoms without any drowsiness.
 - D) the six volunteers who did not report a significant reduction in their allergy symptoms without any drowsiness.
24. In order to investigate whether women are more likely than men to prefer Democratic candidates, a political scientist selects a large sample of registered voters, both men and women. She asks every voter whether they voted for the Republican or the Democratic candidate in the last election. This is
- A) an observational study.
 - B) a multistage sample.
 - C) a double-blind experiment.
 - D) a block design.
25. Will a fluoride mouthwash used after brushing reduce cavities? Twenty sets of twins were used to investigate this question. One member of each set of twins used the mouthwash after each brushing; the other did not. After six months, the difference in the number of cavities of those using the mouthwash was compared with the number of cavities of those who did not use the mouthwash. This experiment uses
- A) random placebos.
 - B) double-blinding.
 - C) double replication.
 - D) a matched-pairs design.
26. A call-in poll conducted by *USA Today* concluded that Americans love Donald Trump. *USA Today* later reported that 5640 of the 7800 calls for the poll came from the offices owned by one man, Cincinnati financier Carl Lindner, who is a friend of Donald Trump. The results of this poll are probably

- A) surprising, but reliable since it was conducted by a nationally recognized organization.
- B) biased, but only slightly since the sample size was quite large.
- C) biased understating the popularity of Donald Trump.
- D) biased overstating the popularity of Donald Trump.

27. Sickle-cell disease is a painful disorder of the red blood cells that affects mostly blacks in the United States. To investigate whether the drug hydroxyurea can reduce the pain associated with sickle-cell disease, a study by the National Institute of Health gave the drug to 150 sickle-cell sufferers and a placebo to another 150. The researchers then counted the number of episodes of pain reported by each subject. The response is
- A) the drug hydroxyurea.
 - B) the number of episodes of pain.
 - C) the presence of sickle-cell disease.
 - D) the number of red blood cells.
28. A study to determine whether or not a football filled with helium traveled farther when kicked than one filled with air found that, while the football filled with helium went, on average, farther than the one filled with air, the difference was not statistically significant. The response
- A) is the gas, air or helium, with which the football is filled.
 - B) does not exist without statistical significance.
 - C) is the number of kickers.
 - D) is the distance the football traveled.

Use the following to answer question 29.

One hundred volunteers who suffer from severe depression are available for a study. Fifty are selected at random and are given a new drug that is thought to be particularly effective in treating severe depression. The other 50 are given an existing drug for treating severe depression. A psychiatrist evaluates the symptoms of all volunteers after four weeks in order to determine if there has been substantial improvement in the severity of the depression.

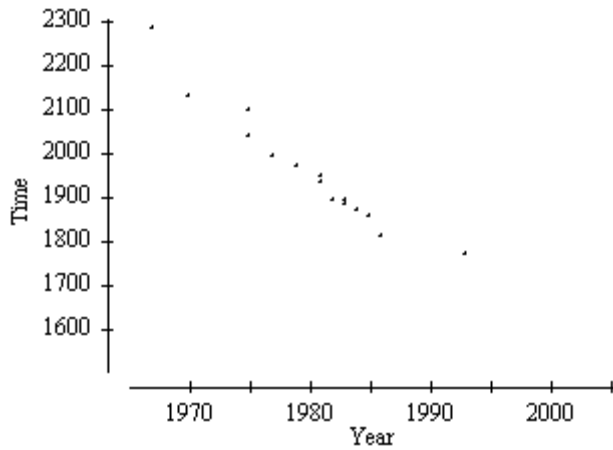
29. The factor in the study above is
- A) which treatment the volunteers receive.
 - B) the use of randomization and the fact that this was a comparative study.
 - C) the extent to which the depression was reduced.
 - D) the use of a psychiatrist to evaluate the severity of depression.
30. A news release for a diet products company reports: "There's good news for the 65 million Americans currently on a diet." Its study showed that people who lose weight can keep it off. The sample was 20 graduates of the company's program who endorse it in commercials. The results of the sample are probably
- A) biased, overstating the effectiveness of the diet.
 - B) biased, understating the effectiveness of the diet.
 - C) unbiased since these are nationally recognized individuals.
 - D) unbiased, but they could be more accurate. A larger sample size should be used.

Use the following to answer question 31.

A television station is interested in predicting whether voters in its viewing area are in favor of federal

funding for abortions. It asks its viewers to phone in and indicate whether they support/are in favor of or are opposed to this. Of the 2241 viewers who phoned in, 1574 (70.24%) were opposed to federal funding for abortions.

31. Referring to the information above, the viewers who phoned in are
- A) a voluntary response sample.
 - B) a convenience sample.
 - C) a probability sample.
 - D) a population.
32. Recent data show that states that spend an above-average amount of money X per pupil in high school tend to have below-average mean Verbal SAT scores Y of all students taking the SAT in the state. In other words, there is a negative association between X (spending per pupil) and Y (mean Verbal SAT score). High spending per pupil and low mean Verbal SAT scores are particularly common in states that have a large percentage of all high school students taking the exam. Such states also tend to have larger populations. The most plausible explanation for the observed association between X and Y is that
- A) X causes Y . Overspending generally leads to extra, unnecessary programs, diverting attention from basic subjects. Inadequate training in these basic subjects generally leads to lower SAT scores.
 - B) Y causes X . Low SAT scores create concerns about the quality of education. This inevitably leads to additional spending to help solve the problem.
 - C) changes in X and Y are due to a common response to other variables. If a higher percentage of students take the exam, the average score will be lower. Also, states with larger populations have large urban areas where the cost of living is higher and more money is needed for expenses.
 - D) the association between X and Y is purely coincidental. It is implausible to believe the observed association could be anything other than accidental.
33. Two variables, x and y , are measured on each of several individuals. The correlation between these variables is found to be 0.88. To help us interpret this correlation we should do which of the following?
- A) Compute the least-squares regression line of y on x and consider whether the slope is positive or negative.
 - B) Interchange the roles of x and y (i.e., treat x as the response and y as the predictor variable) and recompute the correlation.
 - C) Plot the data.
 - D) All of the above.
34. A scatterplot of the world record time for women in the 10,000-meter run versus the year in which the record was set appears below. Note that the time is in seconds and the data are for the period 1965–1995.



Based on this plot, we can expect

- A) that by 2005 the world record time for women will be well below 1500 seconds.
- B) that about every decade, we can expect the world record time to decrease by at least 100 seconds.
- C) that about every decade, we can expect the world record time to decrease by about 50 seconds.
- D) none of the above.

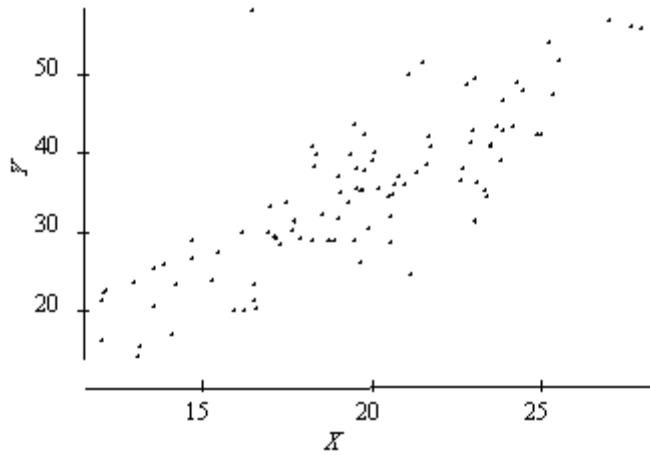
35. A researcher computed the average Math SAT score of all high school seniors who took the SAT exam for each of the 50 states. The researcher also computed the average salary of high school teachers in each of these states and plotted these average salaries against the average Math SAT scores for each state. The plot showed a distinct negative association between average Math SAT scores and teacher salaries. The researcher may legitimately conclude which of the following?

- A) Increasing the average salary of teachers will cause the average of Math SAT scores to decrease, but it is not correct to conclude that increasing the salaries of individual teachers causes the Math SAT scores of individual students to increase.
- B) States that pay teachers high salaries tend to do a poor job of teaching mathematics, on average.
- C) States whose students tend to perform poorly in mathematics probably have a higher proportion of problem students and thus need to pay teachers higher salaries in order to attract them to teach in those states.
- D) The data used by the researcher do not provide evidence that increasing the salary of teachers will cause the performance of students on the Math SAT to get worse.

36. When exploring very large sets of data involving many variables, which of the following is true?

- A) Extrapolation is safe because it is based on a greater quantity of evidence.
- B) Associations will be stronger than would be seen in a much smaller subset of the data.
- C) A strong association is good evidence for causation because it is based on a large quantity of information.
- D) None of the above.

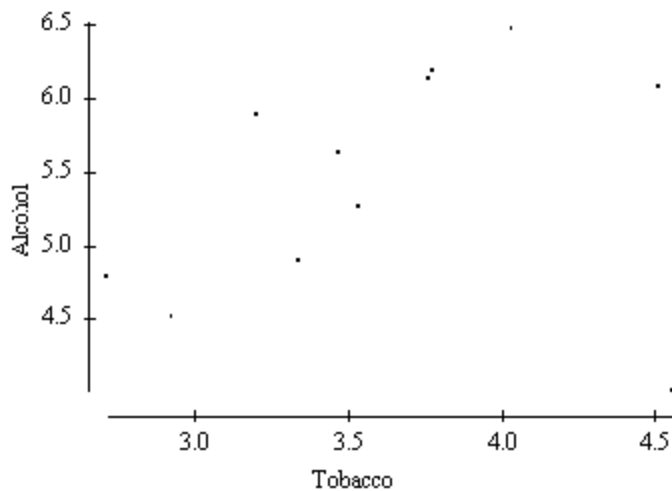
37. Consider the following scatterplot.



From this plot we can conclude

- A) that there is evidence of a modest cause-and-effect relation between X and Y with increases in X causing increases in Y .
- B) that there is an outlier in the plot.
- C) that there is a strongly influential point in the plot.
- D) all of the above.

38. The British government conducts regular surveys of household spending. The average weekly household spending on tobacco products and alcoholic beverages for each of 11 regions in Great Britain was recorded. A scatterplot of spending on tobacco versus spending on alcohol is given below.



Which of the following statements is true?

- A) The observation in the lower right corner of the plot is influential.
 - B) There is clear evidence of negative association between spending on alcohol and tobacco.
 - C) The equation of the least-squares line for this plot would be approximately $y = 10 - 2x$.
 - D) The correlation coefficient for this data is 0.99.
39. Which of the following is true of the least-squares regression line?
- A) The slope is the change in the response variable that would be predicted by a unit change in the explanatory variable.
 - B) It always passes through the point (\bar{x}, \bar{y}) , the means of the explanatory and response variables, respectively.
 - C) It will only pass through all the data points if $r = \pm 1$.

D) All of the above.

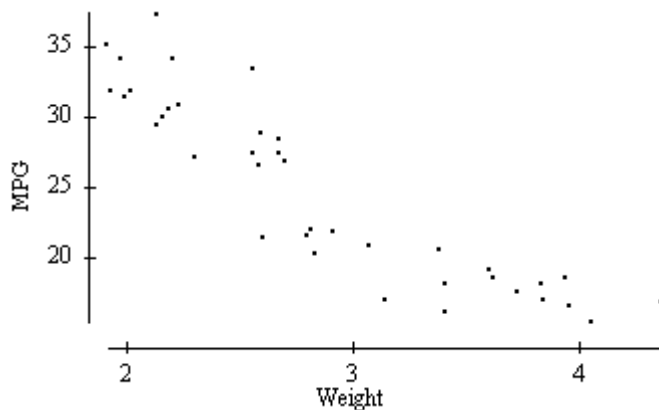
40. Volunteers for a research study were divided into three groups. Group 1 listened to Western religious music, group 2 listened to Western rock music, and group 3 listened to Chinese religious music. The blood pressure of each volunteer was measured before and after listening to the music, and the change in blood pressure (blood pressure before listening minus blood pressure after listening) was recorded. To explore the relationship between type of music and change in blood we could

- A) see if blood pressure decreases as type of music increases by examining a scatterplot.
- B) make a histogram of the change in blood pressure for all of the volunteers.
- C) make side-by-side boxplots of the change in blood pressure, with a separate boxplot for each group.
- D) do all of the above.

41. Which of the following statements is true?

- A) The correlation coefficient equals the proportion of times two variables lie on a straight line.
- B) The correlation coefficient will be +1.0 only if all the data lie on a perfectly horizontal straight line.
- C) The correlation coefficient measures the fraction of outliers that appear in a scatterplot.
- D) The correlation coefficient has no unit of measurement and must always lie between -1.0 and $+1.0$, inclusive.

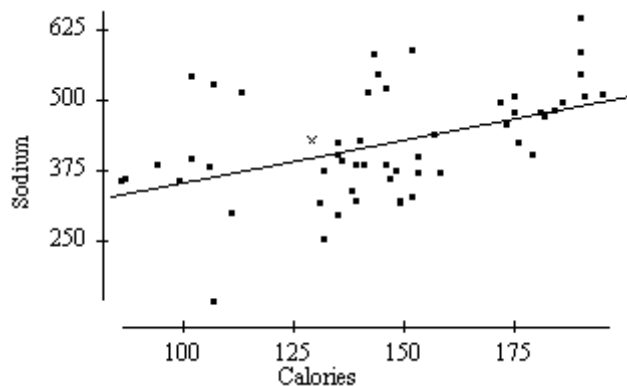
42. Consider the scatterplot below.



According to the scatterplot, which of the following is a plausible value for the correlation coefficient between weight and MPG?

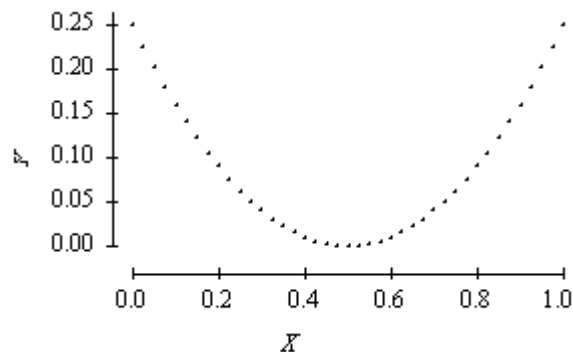
- A) +0.2.
- B) -0.9.
- C) +0.7.
- D) -1.0.

43. A scatterplot of the calories and sodium content of several brands of meat hot dogs is shown below. The least-squares regression line has been drawn in on the plot. Referring to this scatterplot, the value of the residual for the point labeled x



- A) is about 40.
- B) is about 1300.
- C) is about 425.
- D) cannot be determined from the information given.

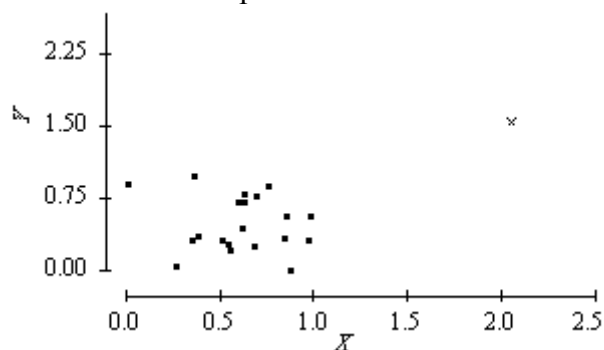
44. Consider the scatterplot below of two variables X and Y .



We may conclude that

- A) the correlation between X and Y must be close to 1 since there is a nearly perfect relation between them.
- B) the correlation between X and Y must be close to -1 since there is a nearly perfect relation between them, but it is not a straight-line relation.
- C) the correlation between X and Y is close to 0.
- D) the correlation between X and Y could be any number between -1 and $+1$. Without knowing the actual values of X and Y we can say nothing more.

45. Consider the scatterplot below.



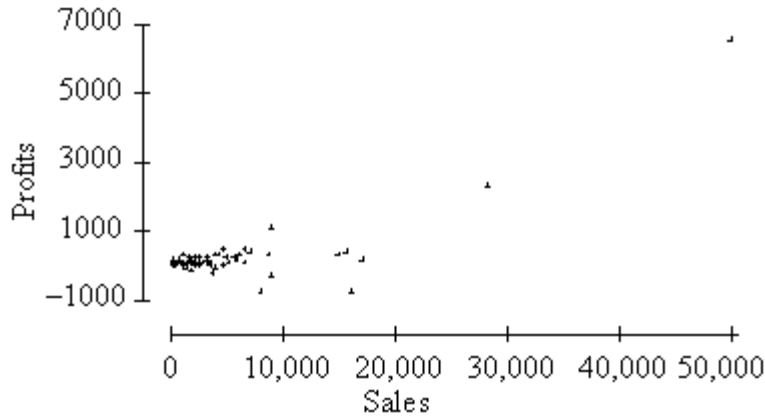
The point indicated by the plotting symbol x would be

- A) a residual.
- C) a z -score.

B) influential.

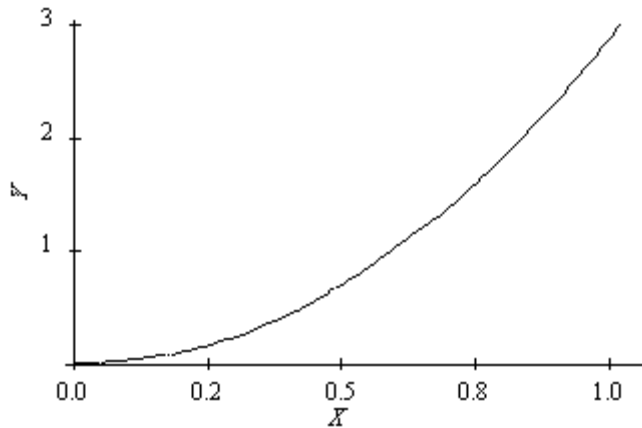
D) a least-squares point.

46. The profits (in multiples of \$100,000) versus the sales (in multiples of \$100,000) for a number of companies are plotted below.



Notice that in the plot profits is treated as the response variable and sales the explanatory variable. The correlation between profits and sales is 0.814. Suppose we had taken sales to be the response variable and profits to be the explanatory variable. In this case, the correlation between sales and profits would be

- A)** 0.814.
 - B)** -0.814 .
 - C)** 0.000.
 - D)** any number between 0.814 and -0.814 , but we can't state the exact value.
47. For the density curve below, which of the following is true?



- A)** The mean and median are equal.
- B)** The mean is greater than the median.
- C)** The mean is less than the median.
- D)** The mean could be either greater than or less than the median.

Use the following to answer questions 48-49.

The distribution of actual weights of 8.0-ounce chocolate bars produced by a certain machine is normal with a mean of 8.1 ounces and a standard deviation of 0.1 ounces.

48. The proportion of chocolate bars weighing less than 8.0 ounces is

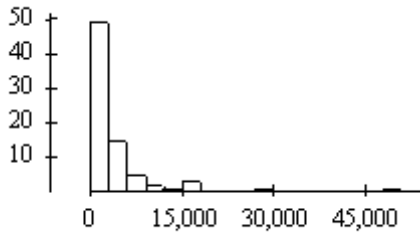
- A) 0.500. B) 0.159. C) 0.341. D) 0.841.

49. The proportion of chocolate bars weighing between 8.2 and 8.3 ounces is
 A) 0.819. B) 0.636. C) 0.477. D) 0.136.

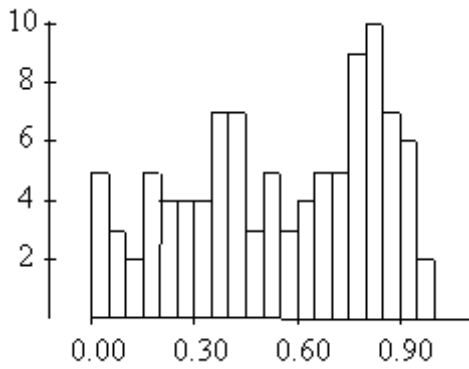
50. Using the standard normal distribution tables, what is the area under the standard normal curve corresponding to $-0.5 < Z < 1.2$?
 A) 0.3085. B) 0.8849. C) 0.5764. D) 0.2815.

51. Which of the following four histograms would best be approximated by a normal distribution?

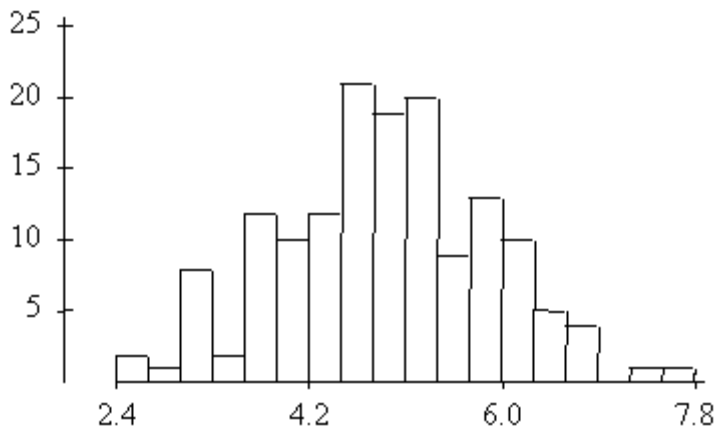
A)

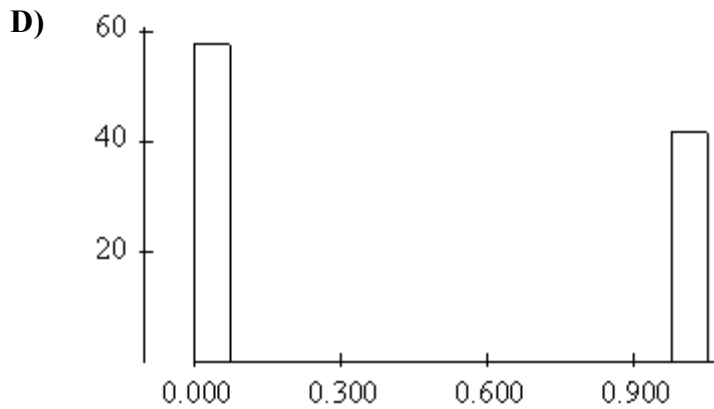


B)



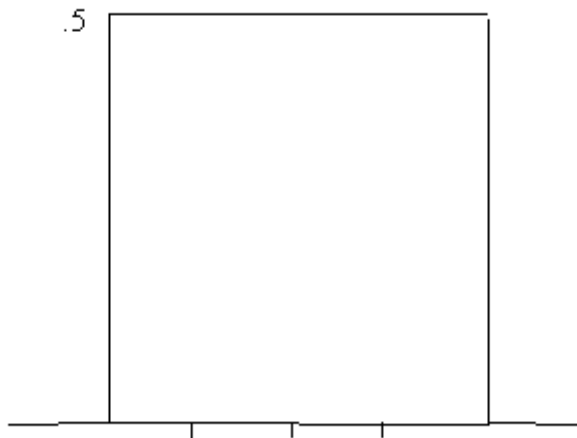
C)





52. Birthweights at a local hospital have a normal distribution with a mean of 110 ounces and a standard deviation of 15 ounces. The proportion of infants with birthweights between 125 ounces and 140 ounces is
- A) 0.819. B) 0.636. C) 0.477. D) 0.136.
53. A normal density curve has which of the following properties?
- A) It is symmetric.
 B) It has a peak centered above its mean.
 C) The spread of the curve is proportional to the standard deviation.
 D) All of the above.

Use the following to answer questions 54-55.



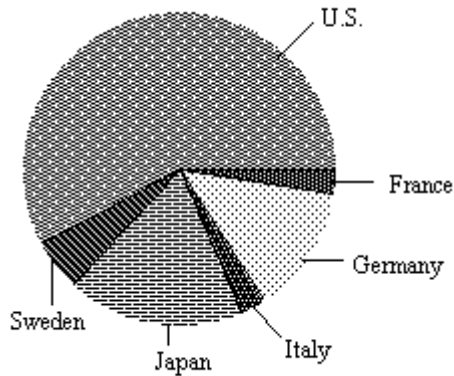
54. For the density curve shown on the previous page, what percentage of the observations lies between 0.5 and 1.2?
- A) 25%. B) 35%. C) 50%. D) 70%.
55. For this density curve, which of the following is true?
- A) It is symmetric. C) The median is 1.
 B) The total area under the curve is 1. D) All of the above.
56. The time to complete a standardized exam is approximately normal with a mean of 70 minutes and a standard deviation of 10 minutes. How much time should be given to complete the exam so that 80% of

the students will complete the exam in the time given?

- A) 84 minutes. B) 78.4 minutes. C) 92.8 minutes. D) 79.8 minutes.

Use the following to answer question 57.

Consumers Union measured the gas mileage in miles per gallon of 38 1978–79 model automobiles on a special test track. The pie chart below provides information about the country of manufacture of the model cars used by Consumers Union.



57. Based on this pie chart, we may conclude that
- A) Japanese cars get significantly lower gas mileage than cars from other countries. We can tell this because their slice of the pie is at the bottom of the chart.
 - B) more than half of the cars in the study were from the United States.
 - C) Swedish cars get gas mileages that are between those of Japanese and U.S. cars.
 - D) Mercedes Benz, Audi, Porsche, and BMW represent approximately a quarter of the cars tested.
58. There are three children in a room, ages three, four, and five. If a four-year-old child enters the room, the
- A) mean age will stay the same but the variance will increase.
 - B) mean age will stay the same but the variance will decrease.
 - C) mean age and variance will stay the same.
 - D) mean age and variance will increase.
59. The Insurance Institute for Highway Safety publishes data on the total damage suffered by compact automobiles in a series of controlled, low-speed collisions. A sample of the data in dollars, with brand names removed, is
- 1000 600 800 1000
- The variance of the above data is
- A) 191.5. B) 36,667. C) 165.8. D) 27,500.
60. The ages of people in a class (to the nearest year) are as follows:

| Age | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 32 |
|--------------------|----|-----|-----|-----|----|----|----|----|----|
| Number of Students | 14 | 120 | 200 | 200 | 90 | 30 | 10 | 2 | 1 |

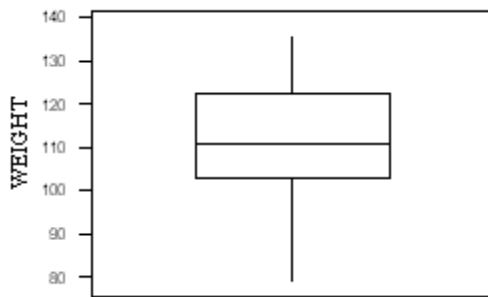
What is true about the median age?

- A) It must be 20.
- B) It could be any number between 19 and 21.
- C) It must be 21.
- D) It must be over 21.

61. When drawing a histogram it is important to
- A) have a separate class interval for each observation to get the most informative plot.
 - B) make sure the heights of the bars exceed the widths of the class intervals, so the bars are true rectangles.
 - C) label the vertical axis, so the reader can determine the counts or percent in each class interval.
 - D) make certain the mean and median are contained in the same class interval, so the correct type of skewness can be identified.

Use the following to answer question 62.

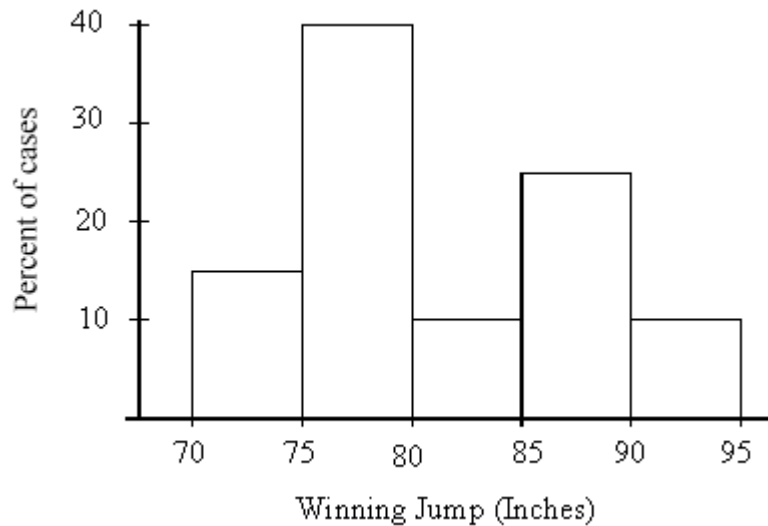
The boxplot below is of the birthweights (in ounces) of a sample of 160 infants born in a local hospital.



62. The number of children with birthweights between 100 and 120 ounces is approximately
- A) 40.
 - B) 50.
 - C) 80.
 - D) 100.
63. A consumer group surveyed the prices for a certain item in five different stores and reported the average price as \$15. We visited four of the five stores and found the prices to be \$10, \$15, \$15, and \$25. Assuming that the consumer group is correct, what is the price of the item at the store that we did not visit?
- A) \$10.
 - B) \$15.
 - C) \$20.
 - D) \$25.

Use the following to answer question 64.

Below is a histogram of the gold medal–winning high jumps for the Olympic Games.



64. The mean of this histogram is approximately
 A) 75 inches. B) 77.5 inches. C) 82 inches. D) 90 inches.
65. The standard deviation of 16 measurements of people's weights (in pounds) is computed to be 5.4. The units for the variance of these measurements is
 A) pounds. C) pounds squared.
 B) square root pounds. D) no units; variance never has units.
66. The standard deviation of 16 measurements of people's weights (in pounds) is computed to be 5.4. The variance of these measurements is
 A) 2.24. B) 29.16. C) 52.34. D) 256.

Answer Key - MidtermReview2

1. B
2. C
3. D
4. A
5. D
6. B
7. D
8. D
9. D
10. B
11. A
12. B
13. B
14. C
15. D
16. A
17. C
18. C
19. D
20. B
21. B
22. D
23. B
24. A
25. D
26. D
27. B
28. D
29. A
30. A
31. A
32. C
33. C
34. D
35. D
36. D
37. B
38. A
39. D
40. C
41. D
42. B
43. A
44. C
45. B

- 46. A
- 47. C
- 48. B
- 49. D
- 50. C
- 51. C
- 52. D
- 53. D
- 54. B
- 55. D
- 56. B
- 57. B
- 58. B
- 59. B
- 60. A
- 61. C
- 62. C
- 63. A
- 64. C
- 65. C
- 66. B