# Practice Midterm 2

(Chapters 1-9)

1. 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.
   E) choose class intervals that vary in width with the narrowest intervals on the ends.

2. In drawing a histogram, which of the following suggestions should be followed?
   A) Leave large gaps between bars. This allows room for comments.
   B) The heights of bars should equal the class frequency.
   C) Generally, bars should be square so that both the height and width equal the class count.
   D) There should be enough intervals so that the correct type of skewness can be identified.
   E) The scale of the vertical axis should be that of the variable whose distribution you are displaying.

Use the following to answer questions 3 and 4:
Each of the following two histograms represents the distribution of acceptance rates (percent accepted) among 25 business schools in 1995. The histograms use different class intervals but are based on the same data. In each class interval, the left endpoint is included but not the right.
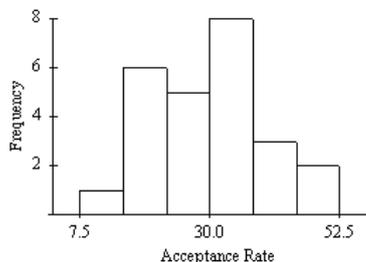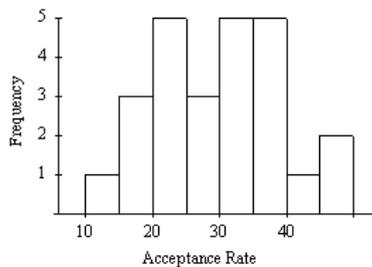

3. What percent of the schools have an acceptance rate of less than 20%?
   A) 3%.   B) 4%.   C) 12%.   D) 16%.   E) 20%.

4. Which interval contains fewer than half of all the observations?
   A) 20% ≤ acceptance rate < 35%.
   B) 22.5% ≤ acceptance rate < 37.5%.
   C) 25% ≤ acceptance rate < 40%.
   D) 30% ≤ acceptance rate < 45%.
   E) 30% ≤ acceptance rate < 52.5%.

Use the following to answer question 5:
In a statistics class with 136 students, the professor records how much money each student has in his or her possession during the first class of the semester. The histogram below is of the data collected.
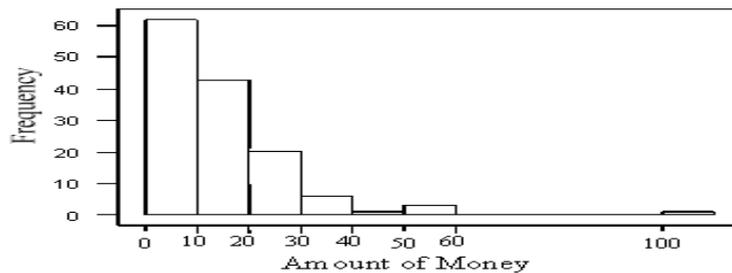

5. The percentage of students with under $10 in their possession is closest to
   A) 30%.   B) 35%.   C) 50%.   D) 60%.   E) 70%.

6. Items produced by a manufacturing process are supposed to weigh 90 grams. However, the manufacturing process is such that there is variability in the items produced and they do not all weigh exactly 90 grams. The distribution of weights can be approximated by a normal distribution with a mean of 90 grams and a standard deviation of 1 gram. Using the 68–95–99.7 rule, what percentage of the items will either weigh less than 87 grams or more than 93 grams?
   A) 0.3%.   B) 3%.   C) 6%.   D) 94%.   E) 99.7%.

7. For the density curve below, which of the following is true?
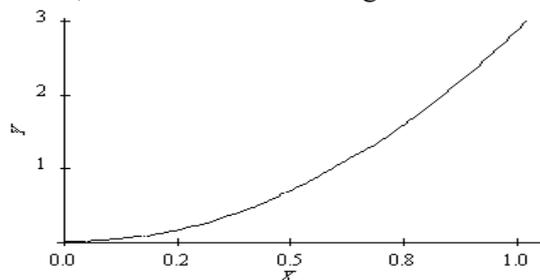
   A) The mean and median are equal.
   B) The mean is greater than the median.
   C) The mean is less than the median.
   D) The mean could be either greater than or less than the median.
   E) The mean is 0.5.

8. Increasing the frequency of observations in the tails of a distribution will
   A) not affect the standard deviation as long as the increases are balanced on either side of the mean.
   B) not affect the standard deviation under any circumstances.
   C) increase the standard deviation.
   D) decrease the standard deviation.
   E) skew the standard deviation.

9. The time it takes for students to complete a standardized exam is approximately normal with a mean of 70 minutes and a standard deviation of 10 minutes. Using the 68–95–99.7 rule, what percentage of students will complete the exam in under an hour?
   A) 68%.   B) 47.5%.   C) 32%.   D) 16%.   E) 5%.

10. Using the standard normal distribution tables, what is the area under the standard normal curve corresponding to $Z < 1.1$?
    A) 0.1357.   B) 0.2704.   C) 0.3643.   D) 0.8413.   E) 0.8643.

11. At a large university, the office responsible for scheduling classes notices that demand is low for classes that meet before 10:00 AM or after 3:00 PM and is high for classes that meet between 10:00 AM and 3:00 PM. Which of the following may we conclude?
    A) There is an association between demand for classes and the time the classes meet.
    B) The association between demand for classes and time for classes is linear.
    C) There is a *negative* association between demand for classes and the time the classes meet.
    D) There is no association between demand for classes and the time the classes meet.
    E) There is a *positive* association between demand for classes and the time the classes meet.

12. The graph below plots the gas mileage (miles per gallon) of various 1978 model cars versus the weight of these cars in thousands of pounds.
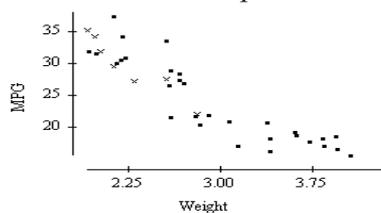
In the graph, the points denoted by the plotting symbol x correspond to cars made in Japan. From this plot, we may conclude that
    A) in 1978 there was little difference between Japanese cars and cars made in other countries.
    B) in 1978 Japanese cars tended to be lighter in weight than other cars.
    C) in 1978 Japanese cars tended to get poorer gas mileage than other cars.
    D) there is a positive association between weight and gas mileage for Japanese cars.
    E) the plot is invalid. A scatterplot is used to represent quantitative variables, and the country that makes a car is a qualitative variable.

3

13. 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 listened to and change in blood pressure, 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) make a pie chart displaying the distribution of type of music listened to for all of the volunteers.
    E) do all of the above.

14. A school guidance counselor examines the number of extracurricular activities of students and their grade point average. The guidance counselor says, "The evidence indicates that the correlation between the number of extracurricular activities a student participates in and his or her grade point average is close to zero." A correct interpretation of this statement would be that
    A) active students tend to be students with poor grades, and vice versa.
    B) students with good grades tend to be students that are not involved in many activities, and vice versa.
    C) students involved in many extracurricular activities are just as likely to get good grades as bad grades. The same is true for students involved in few extracurricular activities.
    D) as a student becomes more involved in extracurricular activities, there will be a change in his/her grades.
    E) involvement in many extracurricular activities and good grades go hand in hand.

15. A student wonders if people of similar heights tend to date each other. She measures herself, her dormitory roommate, and the women in the adjoining rooms; then she measures the next man each woman dates. Here are the data (heights in inches):

Women	66	64	66	65	70	65
Men	72	68	70	68	74	69

    Which of the following statements is true?
    A) The variables measured are all categorical.
    B) There is a strong negative association between the heights of men and women, since the women are always smaller than the men they date.
    C) Tall women tend to date short men.
    D) Any height above 70 inches must be considered an outlier.
    E) There is a positive association between the heights of men and women who date each other.

16. Using least-squares regression, I determine that the (base 10) logarithm of the population of a country is approximately described by the equation
$$\widehat{\log(\text{population})} = -13.5 + 0.01 \times (\text{year})$$
Based on this equation, the population of the country in the year 2010 should be about
A) 6.6. B) 735. C) 2,000,000. D) 3,981,072. E) 33,000,000.

17. Which of the following would provide evidence that a power law model describes the relationship between a response variable y and an explanatory variable x?
A) A scatterplot of y versus x looks approximately linear.
B) A scatterplot of $\log y$ versus x looks approximately linear.
C) A scatterplot of y versus $\log x$ looks approximately linear.
D) A scatterplot of $\log y$ versus $\log x$ looks approximately linear.
E) A scatterplot of the square root of y versus x looks approximately linear.

18. Which of the following scatterplots would indicate that Y is growing exponentially over time?

A)

C)

B)

D)

E)

5
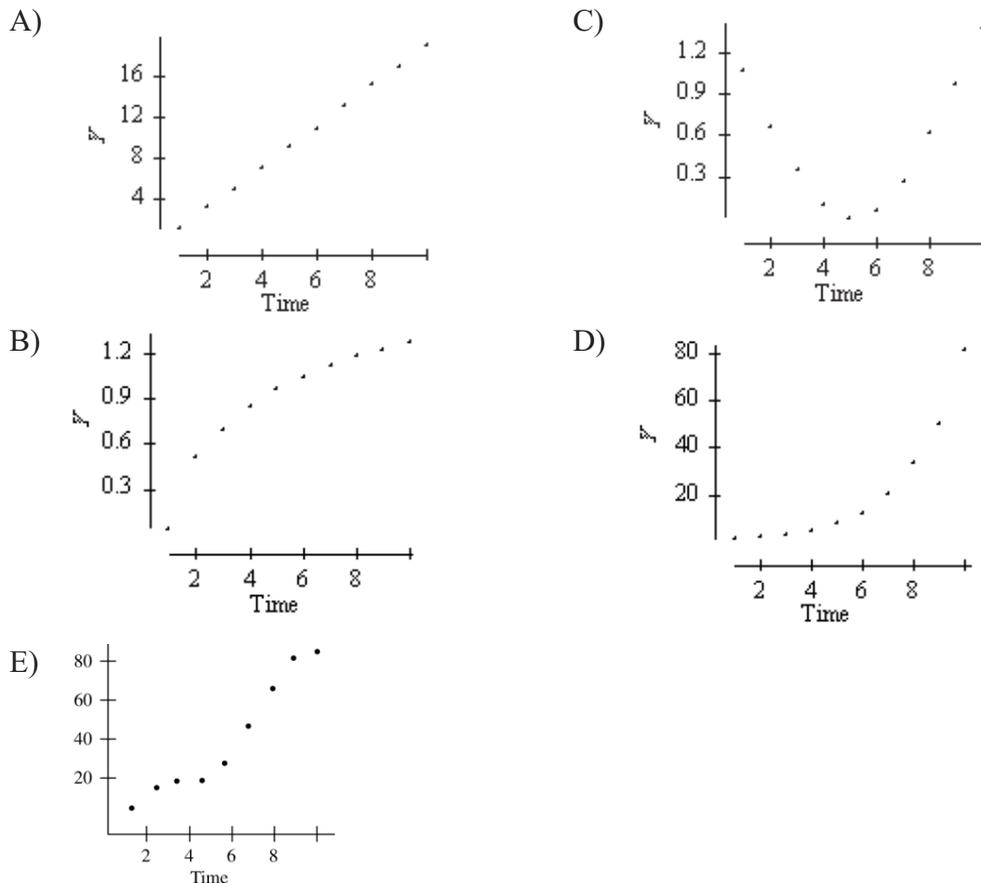
19. A scatterplot of a response variable Y versus an explanatory variable X is given below.
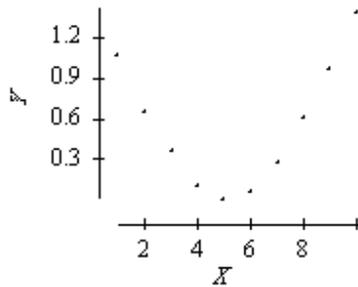
Which of the following is true?
A) There is a nonlinear relationship between Y and X.
B) There is a very strong positive correlation between Y and X because there is an obvious relationship between these variables.
C) There is a monotonic relationship between Y and X.
D) There is a strong quadratic relationship between Y and X.
E) All of the above.

20. Suppose the relationship between a response variable y and a predictor variable x is approximately
$$y = 2.7 \times 10^{-0.5x}$$
Which of the following plots would approximately follow a straight line?
A) A plot of y against x.
B) A plot of y against $\log x$.
C) A plot of $\log y$ against x.
D) A plot of 10^y against x.
E) A plot of $\log y$ against $\log x$.

Use the following to answer questions 21 and 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 policy. Of the 2241 viewers who phoned in, 1574 (70.24%) were opposed to federal funding for abortions.

21. 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.
E) a simple random sample.

22. Referring to the information above, the sample obtained is
A) a simple random sample.
B) a single-stage sample.
C) a census.
D) close to the actual proportion of the population who support/oppose federal funding for abortions.
E) probably biased.

23. In order to assess the opinion of students at the University of Minnesota on campus snow removal, a reporter for the student newspaper interviews the first 12 students he meets who are willing to express their opinion. The method of sampling used is
   A) numerical sampling.
   B) convenience sampling.
   C) voluntary response.
   D) a census.
   E) simple random sampling.

Use the following to answer questions 24 and 25:
Choose a simple random sample of size three from the following employees of a small company.

1. Bechhofer   4. Kesten   7. Taylor
2. Brown   5. Kiefer   8. Wald
3. Ito   6. Spitzer   9. Weiss

Use the numerical labels attached to the names above and the list of random digits below. Read the list of random digits from left to right, starting at the beginning of the list.

11793 20495 05907 11384 44982 20751 27498 12009 45287 71753 98236 66419 84533

24. Referring to the information above, the simple random sample is
   A) 117.
   B) Bechhofer, Ito, Weiss.
   C) Bechhofer, Taylor, Weiss.
   D) Kesten, Kiefer, Taylor.
   E) Bechhofer, then Bechhofer again, then Taylor.

25. If we used another list of random digits to select the sample, which of the following statements would be true?
   A) We would get the same result that we obtained with the list used here.
   B) We would get the same names as those obtained from the sample listed here.
   C) We would get at most one name in common with the sample obtained here.
   D) It would be just as likely that the sample that we obtained here would be selected as any other set of three names.
   E) We would get a completely different sample than that obtained with the list used here.

26. Suppose we have a *loaded* die that gives the outcomes 1–6 according to the following 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 the die were fair. If this die is rolled 6000 times, the number of times we get a 2 or a 3 should be about
A) 1000.   B) 2000.   C) 3000.   D) 4000.
E) The answer cannot be determined since the probabilities are only approximate.

27. 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.
E) dependent.

28. In a particular game, a fair die is tossed. If the number of spots showing is either 4 or 5, you win $1; if the number of spots showing is 6, you win $4; and if the number of spots showing is 1, 2, or 3, you win nothing. If it costs you $1 to play the game, the probability that you win more than the cost of playing is
A) 0.   B) 1/6.   C) 1/3.   D) 2/3.   E) 5/6.

29. I select two cards from a deck of 52 cards and observe the color of each (26 cards in the deck are red and 26 are black). Which of the following is an appropriate sample space S for the possible outcomes?
A) $S = \{red, black\}$.
B) $S = \{(red, red), (red, black), (black, red), (black, black)\}$, where, for example, (red, red) stands for the event "the first card is red and the second card is red."
C) $S = \{0, 1, 2\}$.
D) All of the above.
E) The results will vary since the cards are drawn without replacement.

30. A game consists of drawing three cards at random from a deck of playing cards. You win $3 for each red card that is drawn. It costs $2 to play. For one play of this game, the sample space S for the net amount you win (after deducting the cost of play) is
A) $S = \{\$0, \$1, \$2, \$3\}$ D) $S = \{-\$2, \$3, \$6, \$9\}$
B) $S = \{-\$6, -\$3, \$0, \$6\}$ E) $S = \{\$0, \$3, \$6, \$9\}$
C) $S = \{-\$2, \$1, \$4, \$7\}$

31. Suppose there are three balls in a box. On one of the balls is the number 1, on another is the number 2, and on the third is the number 3. You select two balls at random and without replacement from the box and note the two numbers observed. The sample space S consists of the three equally likely outcomes {(1, 2), (1, 3), (2, 3)}. X, the sum of the numbers on the two balls selected, has the following probability distribution:

X	3	4	5
Probability	1/3	1/3	1/3

The probability that X is at least 4 is
A) 0.   B) 1/3.   C) 9/20.   D) 2/3.   E) 1.

Use the following to answer questions 32 through 34:
The probability density curve of a random variable X is given in the figure below.
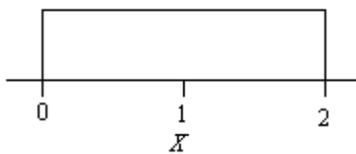
32. Referring to the information above, the probability that X is between 0.5 and 1.5 is
A) 1/4.   B) 1/3.   C) 1/2.   D) 3/4.   E) 1.

33. 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.   E) 2/3.

34. Referring to the information above, the probability that $X = 1.5$ is
A) 0.   B) 1/4.   C) 1/3.   D) 1/2.   E) 3/4.

35. A random variable is
A) a hypothetical list of the possible outcomes of a random phenomenon.
B) any phenomenon in which outcomes are equally likely.
C) any number that changes in a predictable way in the long run.
D) a variable used to represent the outcome of a random phenomenon.
E) a variable whose value is a numerical outcome associated with a random phenomenon.

36. If X is binomial with parameters $n = 9$ and $p = 1/3$, the mean μ_X of X is
A) 6.   B) 3.   C) 2.   D) 1.414.   E) 1.732.

37. If X is binomial with parameters $n = 9$ and $p = 1/3$, the standard deviation σ_X of X is
A) 6.   B) 3.   C) 2.   D) 1.414.   E) 1.732.

38. 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 at most once is
    A) 0.0819.   B) 0.2.   C) 0.3277.   D) 0.4096.   E) 0.7373.

39. 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
    A) 1.   B) 0.6723.   C) 0.3277.   D) 0.04.   E) 0.00032.

40. In a certain game of chance, your chances of winning are 0.2. You play the game five times and outcomes are independent. Suppose it costs $1 to play the game each time. Each time you win, you receive $4 (for a net gain of $3). Each time you lose, you receive nothing (for a net loss of $1). Your expected winnings for five plays are
    A) $3.   B) $1.   C) $0.   D) –$1.   E) –$2.

41. The distribution of the values taken on by a statistic in all possible samples from the same population is called
    A) the sample distribution.   D) a table of random digits.
    B) the bias of the statistic.   E) the parameter.
    C) the sampling distribution.

42. I flip a coin 10 times and record the proportion of heads I obtain. I then repeat this process of flipping the coin 10 times and recording the proportion of heads obtained many, many times. When done, I make a histogram of my results. This histogram represents
    A) the bias, if any, that is present.
    B) a binomial distribution.
    C) simple random sampling.
    D) the sampling distribution of the proportion of heads in 10 flips of the coin.
    E) the true population parameter.

43. If a statistic used to estimate a parameter is such that the mean of its sampling distribution is equal to the true value of the parameter being estimated, the statistic is said to be
    A) random.   B) biased.   C) variable.   D) unbiased.   E) a proportion.

44. A statistic is said to be unbiased if
    A) the data was collected from a double-blind study.
    B) the mean of its sampling distribution is equal to the true value of the parameter being estimated.
    C) both the person who calculated the statistic and the subjects whose responses made up the statistic were truthful.
    D) it is only used for honest purposes.
    E) the survey used to obtain the statistic was designed so as to avoid even the hint of racial or sexual prejudice.

45. The variability of a statistic is described by
    A) the spread of its sampling distribution.
    B) the amount of bias present.
    C) the different values the statistic will take from each of the samples.
    D) the stability of the population it describes.
    E) the vagueness in the wording of the question used to collect the sample data.

MC Practice Midterm #2 Key

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