SECTION 2.1 Exercises

1.

Data Sets

pg 85   CrunchIt!

Shoes How many pairs of shoes do students have? Do girls have more shoes than boys? Here are data from a random sample of 20 female and 20 male students at a large high school:

Female:	50	26	26	31	57	19	24	22	23	38
	13	50	13	34	23	30	49	13	15	51
Male:	14	7	6	5	12	38	8	7	10	10
	10	11	4	5	22	7	5	10	35	7

- (a) Find and interpret the percentile in the female distribution for the girl with 22 pairs of shoes.

- (b) Find and interpret the percentile in the male distribution for the boy with 22 pairs of shoes.

- (c) Who is more unusual: the girl with 22 pairs of shoes or the boy with 22 pairs of shoes? Explain.

Hide Answer
Correct Answer

(a) The girl with 22 pairs of shoes is the 6th smallest. Her percentile is 0.25. 25% of girls have fewer pairs of shoes. **(b)** The boy with 22 pairs has more shoes than 17 people. His percentile is 0.85. 85% of boys have fewer pairs of shoes. **(c)** The boy is more unusual because only 15% of the boys have as many or more than he has, while the girl has a value that is more centered in the distribution. 25% have fewer and 75% have as many or more.

2.

Old folks Here is a stemplot of the percents of residents aged 65 and older in the 50 states:

```
 7 | 0
 8 | 8
 9 | 8
10 | 019
11 | 16777
12 | 01122456778999
13 | 0001223344455689
14 | 023568
15 | 24
16 | 9
```

Key: 15|2 means 15.2% of this state's residents are 65 or older

- (a) Find and interpret the percentile for Colorado, which has 10.1% of its residents aged 65 or older.

- (b) Find and interpret the percentile for Rhode Island, with 13.9% of residents aged 65 or older.

- (c) Which of these two states is more unusual? Explain.

3.

Speed limits According to the *Los Angeles Times*, speed limits on California highways are set at the 85th percentile of vehicle speeds on those stretches of road. Explain what that means to someone who knows little statistics.

Hide Answer

Correct Answer

According to the *Los Angeles Times,* the speed limits on California highways are such that 85% of the vehicle speeds on those stretches of road are less than the speed limit.

4.

Blood pressure Larry came home very excited after a visit to his doctor.

He announced proudly to his wife, "My doctor says my blood pressure is at the 90th percentile among men like me. That means I'm better off than about 90% of similar men." How should his wife, who is a statistician, respond to Larry's statement?

5.

Growth charts We used an online growth chart to find percentiles for the height and weight of a 16-year-old girl who is 66 inches tall and weighs 118 pounds. According to the chart, this girl is at the 48th percentile for weight and the 78th percentile for height. Explain what these values mean in plain English.

Correct Answer

The girl in question weighs more than 48% of girls her age, but is taller than 78% of the girls her age. Since she is taller than 78% of girls, but only weighs more than 48% of girls, she is probably fairly skinny.

6.

Run fast Peter is a star runner on the track team. In the league championship meet, Peter records a time that would fall at the 80th percentile of all his race times that season. But his performance places him at the 50th percentile in the league championship meet. Explain how this is possible. (Remember that lower times are better in this case!)

*Exercises 7 and 8 involve a new type of graph called a **percentile plot.** Each point gives the value of the variable being measured and the corresponding percentile for one individual in the data set.*

7.

Text me The *percentile plot* below shows the distribution of text messages sent and received in a two-day period by a random sample of 16 females from a large high school.

- (a) Describe the student represented by the highlighted point.

- (b) Use the graph to estimate the median number of texts. Explain your method.

Text messaging — Percentile Plot
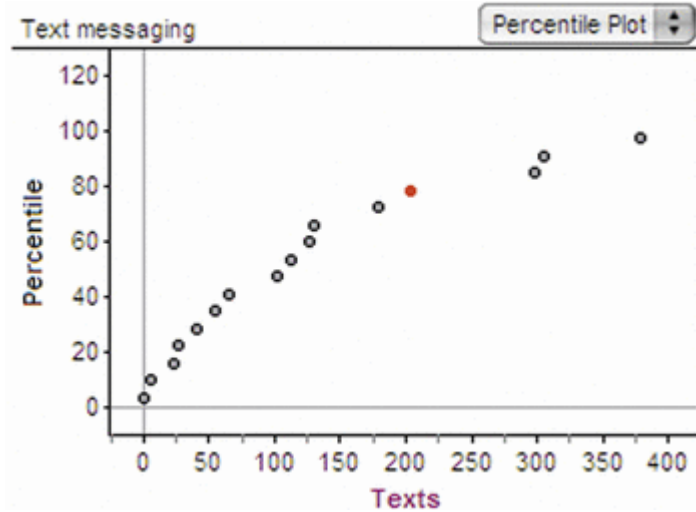
Correct Answer

(a) The highlighted student sent about 212 text messages in the two-day period which placed her at about the 80th percentile. **(b)** The median number of texts is the same as the 50th percentile. Locate 50% on the *y* axis, read over to the points and then find the relevant place on the *x* axis. The median is approximately 115 text messages.

8.

Foreign-born residents The *percentile plot* below shows the distribution of the percent of foreign-born residents in the 50 states.

- (a) The highlighted point is for Maryland. Describe what the graph tells you about this state.

- (b) Use the graph to estimate the 30th percentile of the distribution. Explain your method.

Foreign-born residents — Percentile Plot

9.

pg 88
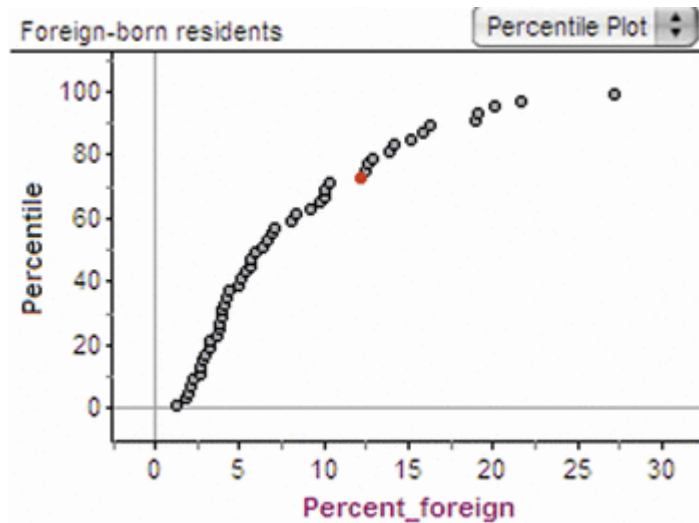
Shopping spree The figure below is a cumulative relative frequency graph of the amount spent by 50 consecutive grocery shoppers at a store.
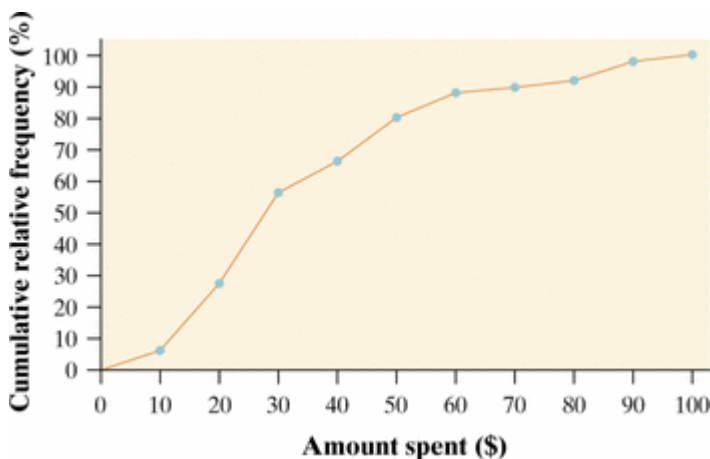

- (a) Estimate the interquartile range of this distribution. Show your method.

- (b) What is the percentile for the shopper who spent $19.50?

- (c) *Challenge:* Draw the histogram that corresponds to this graph.

**Hide Answer**
Correct Answer

(a) First find the quartiles. The first quartile is the 25th percentile. Find 25 on the *y* axis, read over to the line and then down to the *x* axis to get about $19. The 3rd quartile is the 75th percentile. Find 75 on the *y* axis, read over to the line and then down to the *x* axis to get about $50. So the interquartile range is $50 − $19 = $31.
(b) Approximately the 26th percentile. **(c)** Here is a histogram.
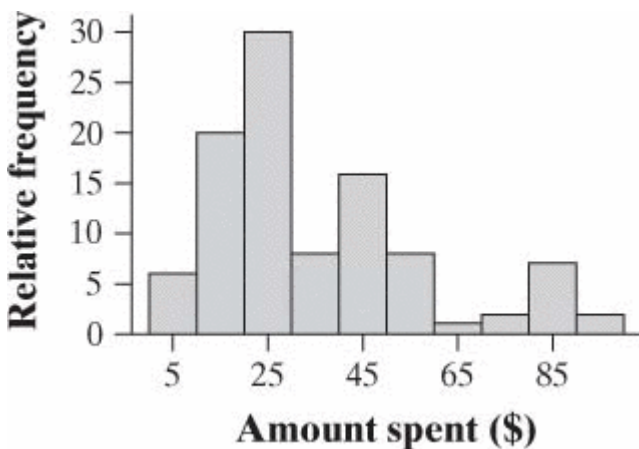
10.

Light it up! The graph below is a cumulative relative frequency graph

showing the lifetimes (in hours) of 200 lamps.[5]
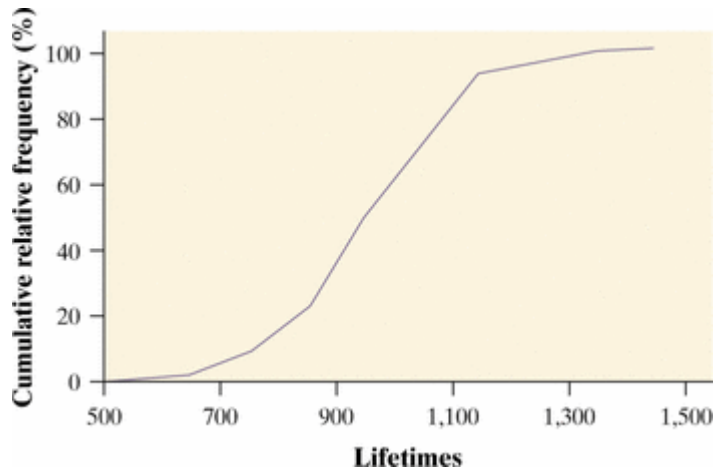
- (a) Estimate the 60th percentile of this distribution. Show your method.

- (b) What is the percentile for a lamp that lasted 900 hours?

11.

SAT versus ACT Eleanor scores 680 on the SAT Mathematics test. The distribution of SAT scores is symmetric and single-peaked, with mean 500 and standard deviation 100. Gerald takes the American College Testing (ACT) Mathematics test and scores 27. ACT scores also follow a symmetric, single-peaked distribution—but with mean 18 and standard deviation 6. Find the standardized scores for both students. Assuming that both tests measure the same kind of ability, who has the higher score?

Hide Answer
Correct Answer

Eleanor's standardized score, $z = 1.8$, is higher than Gerald's standardized score, $z = 1.5$.

12.

Comparing batting averages Three landmarks of baseball achievement are Ty Cobb's batting average of .420 in 1911, Ted Williams's .406 in 1941, and George Brett's .390 in 1980. These batting averages cannot be compared directly because the distribution of major league batting averages has changed over the years. The distributions are quite symmetric, except for outliers such as Cobb, Williams, and Brett. While the mean batting average has been held roughly constant by rule changes and the balance between hitting and pitching, the standard deviation has dropped over time. Here are the facts:

Decade	Mean	Standard deviation
1910s	0.266	0.0371
1940s	0.267	0.0326
1970s	0.261	0.0317

Compute the standardized batting averages for Cobb, Williams, and Brett to compare how far each stood above his peers.[6]

13.

Measuring bone density Individuals with low bone density have a high risk of broken bones (fractures). Physicians who are concerned about low bone density (osteoporosis) in patients can refer them for specialized testing. Currently, the most common method for testing bone density is dual-energy X-ray absorptiometry (DEXA). A patient who undergoes a DEXA test usually gets bone density results in grams per square centimeter (g/cm^2) and in standardized units.

Judy, who is 25 years old, has her bone density measured using DEXA. Her results indicate a bone density in the hip of 948 g/cm^2 and a standardized score of $z = -1.45$. In the reference population of 25-year-old women like Judy, the mean bone density in the hip is 956 g/cm^2.[7]

- (a) Judy has not taken a statistics class in a few years. Explain to her in simple language what the standardized score tells her about her bone density.

- (b) Use the information provided to calculate the standard deviation of bone density in the reference population.

Hide Answer
Correct Answer

(a) Judy's bone density score is about one and a half standard deviations below the average score for all women her age. The fact that your standardized score is negative indicates that your bone density is below the average for your peer group. The magnitude of the standardized score tells us how many standard deviations you are below the average (about 1.5). **(b)** $\sigma = 5.52$ grams/cm^2.

14.

Comparing bone density Refer to the previous exercise. One of Judy's friends, Mary, has the bone density in her hip measured using DEXA. Mary is 35 years old. Her bone density is also reported as 948 g/cm^2, but her standardized score is $z = 0.50$. The mean bone density in the hip for the reference population of 35-year-old women is 944 grams/cm^2.

- (a) Whose bones are healthier—Judy's or Mary's? Justify your answer.

- (b) Calculate the standard deviation of the bone density in Mary's reference population. How does this compare with your answer to

Exercises 15 and 16 refer to the dotplot and summary statistics of salaries for players on the World Champion 2008 Philadelphia Phillies baseball team.[8]

Salary (millions)
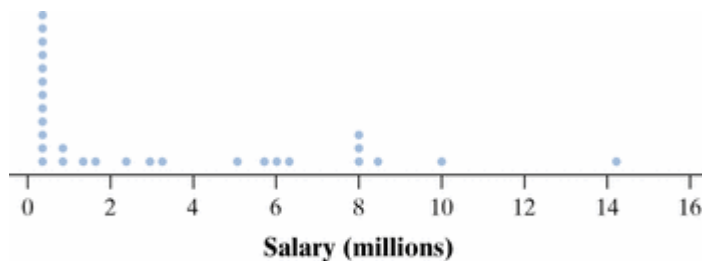
Variable	n	Mean	Std. dev.	Min	Q_1	M	Q_3	Max
Salary	29	3388617	3767484	390000	440000	1400000	6000000	14250000

15.

Baseball salaries Brad Lidge played a crucial role as the Phillies' "closer," pitching the end of many games throughout the season. Lidge's salary for the 2008 season was $6,350,000.

- (a) Find the percentile corresponding to Lidge's salary. Explain what this value means.

- (b) Find the z-score corresponding to Lidge's salary. Explain what this value means.

Hide Answer

Correct Answer

(a) Since 22 salaries were less than Lidge's salary, his salary is at the 75.86 percentile. **(b)** $z = 0.79$. Lidge's salary was 0.79 standard deviations above the mean salary of $3,388,617.

16.

Baseball salaries Did Ryan Madson, who was paid $1,400,000, have a high salary or a low salary compared with the rest of the team? Justify your answer by calculating and interpreting Madson's percentile and z-score.

Exercises 17 and 18 refer to the following setting. Each year, about 1.5 million college-bound high school juniors take the PSAT. In a recent year, the mean score on the Critical Reading test was 46.9 and the standard deviation was 10.9. Nationally, 5.2% of test takers earned a score of 65 or higher on the Critical Reading test's 20 to 80 scale.[9]

17.

PSAT scores Scott was one of 50 junior boys to take the PSAT at his school. He scored 64 on the Critical Reading test. This placed Scott at the 68th percentile within the group of boys. Looking at all 50 boys' Critical

Reading scores, the mean was 58.2 and the standard deviation was 9.4.

- (a) Write a sentence or two comparing Scott's percentile among the national group of test takers and among the 50 boys at his school.

- (b) Calculate and compare Scott's *z*-score among these same two groups of test takers.

Hide Answer
Correct Answer

(a) In the national group, about 94.8% of the test takers scored below 65. Scott's percentiles, 94.8th among the national group and 68th within the school, indicate that he did better among all test takers than he did among the 50 boys at his school. **(b)** Scott's *z*-scores are $z = 1.57$ among the national group and $z = 0.62$ among the 50 boys at his school.

18.

PSAT scores How well did the boys at Scott's school perform on the PSAT? Give appropriate evidence to support your answer.

19.

pg 96

Tall or short? Mr. Walker measures the heights (in inches) of the students in one of his classes. He uses a computer to calculate the following numerical summaries:

Mean	Std.dev.	Min	Q_1	M	Q_3	Max
69.188	3.20	61.5	67.75	69.5	71	74.5

Next, Mr. Walker has his entire class stand on their chairs, which are 18 inches off the ground. Then he measures the distance from the top of each student's head to the floor.

- (a) Find the mean and median of these measurements. Show your work.

- (b) Find the standard deviation and *IQR* of these measurements. Show your work.

Hide Answer
Correct Answer

(a) The mean and the median both increase by 18 so the mean is 87.188 and the median is 87.5. The distribution of heights just shifts by 18 inches. **(b)** The standard deviation and *IQR* do not change. For the standard deviation, note that although the mean increased by 18, the observations each increased by 18 as well so that the deviations did not change. For the *IQR*, Q_1 and Q_3 both increase by 18 so that their difference remains the same as in the original data set.

20.

Teacher raises A school system employs teachers at salaries between

$28,000 and $60,000. The teachers' union and the school board are negotiating the form of next year's increase in the salary schedule.

- (a) If every teacher is given a flat $1000 raise, what will this do to the mean salary? To the median salary? Explain your answers.

- (b) What would a flat $1000 raise do to the extremes and quartiles of the salary distribution? To the standard deviation of teachers' salaries? Explain your answers.

21.

Tall or short? Refer to **Exercise 19**. Mr. Walker converts his students' original heights from inches to feet.

- (a) Find the mean and median of the students' heights in feet. Show your work.

- (b) Find the standard deviation and *IQR* of the students' heights in feet. Show your work.

Hide Answer

Correct Answer

(a) To give the heights in feet, not inches, we would divide each observation by 12 (12 inches = 1 foot). Thus the mean and median are divided by 12. The new mean is 5.77 feet and the new median is 5.79 feet. **(b)** To find the standard deviation in feet, note that each deviation in terms of feet is found by dividing the original deviation by 12.

$$\text{standard deviation}_{New} = \sqrt{\frac{(\text{first deviation (ft)})^2 + \cdots + ((\text{last deviation (ft)})^2)}{n-1}}$$

$$= \sqrt{\frac{\left(\frac{\text{first deviation (in)}}{12}\right)^2 + \cdots + \left(\frac{\text{last deviation (in)}}{12}\right)^2}{n-1}}$$

$$= \frac{1}{12} \cdot \text{standard deviation}_{Old} = \frac{3.2}{12} = 0.27 \text{ feet}$$

The first and third quartiles are still the medians of the first and second halves of the data, these values must simply be converted to feet. To do this, divide the first and third quartiles of the original data set by 12: $Q_1 = \frac{67.75}{12} = 5.65$ feet and $Q_3 = \frac{71}{12} = 5.92$ feet. So the interquartile range is $IQR = 5.92 - 5.65 = 0.27$ feet.

22.

Teacher raises Refer to **Exercise 20**. If each teacher receives a 5% raise instead of a flat $1000 raise, the amount of the raise will vary from $1400 to $3000, depending on the present salary.

- (a) What will this do to the mean salary? To the median salary? Explain your answers.

- (b) Will a 5% raise increase the *IQR?* Will it increase the standard

deviation? Explain your answers.

23.

Cool pool? Coach Ferguson uses a thermometer to measure the temperature (in degrees Celsius) at 20 different locations in the school swimming pool. An analysis of the data yields a mean of 25°C and a standard deviation of 2°C. Find the mean and standard deviation of the temperature readings in degrees Fahrenheit (recall that °F = (9/5)°C + 32).

Hide Answer
Correct Answer

Mean in degrees Fahrenheit is 77. Standard deviation in degrees Fahrenheit is 3.6.

24.

Measure up Clarence measures the diameter of each tennis ball in a bag with a standard ruler. Unfortunately, he uses the ruler incorrectly so that each of his measurements is 0.2 inches too large. Clarence's data had a mean of 3.2 inches and a standard deviation of 0.1 inches. Find the mean and standard deviation of the corrected measurements in centimeters (recall that 1 inch = 2.54 cm).

25.

Density curves Sketch a density curve that might describe a distribution that is symmetric but has two peaks.

Hide Answer
Correct Answer

Sketches will vary.

26.

Density curves Sketch a density curve that might describe a distribution that has a single peak and is skewed to the left.

*Exercises 27 to 30 involve a special type of density curve–one that takes constant height (looks like a horizontal line) over some interval of values. This density curve describes a variable whose values are distributed evenly (uniformly) over some interval of values. We say that such a variable has a **uniform distribution**.*

27.

Biking accidents Accidents on a level, 3-mile bike path occur uniformly along the length of the path. The figure below displays the density curve that describes the uniform distribution of accidents.
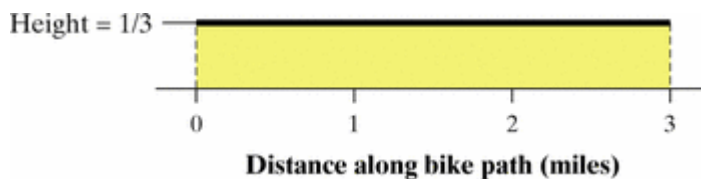
- (a) Explain why this curve satisfies the two requirements for a density curve.

- (b) The proportion of accidents that occur in the first mile of the path is the area under the density curve between 0 miles and 1 mile. What is this area?

- (c) Sue's property adjoins the bike path between the 0.8 mile mark and the 1.1 mile mark. What proportion of accidents happen in front of Sue's property? Explain.

Correct Answer

(a) It is on or above the horizontal axis everywhere and the area beneath the curve is 1. **(b)** $\frac{1}{3}$ **(c)** Since $(1.1 - 0.8) . \frac{1}{3} = 0.1$, one-tenth of accidents occur next to Sue's property.

28.

A uniform distribution The figure below displays the density curve of a uniform distribution. The curve takes the constant value 1 over the interval from 0 to 1 and is 0 outside the range of values. This means that data described by this distribution take values that are uniformly spread between 0 and 1.
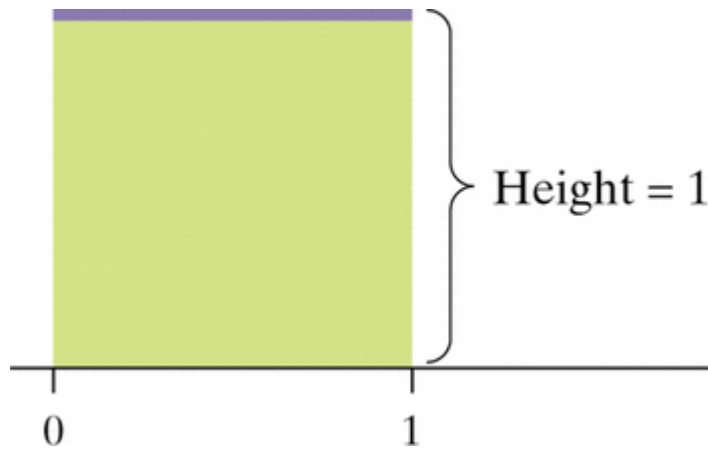
- (a) Explain why this curve satisfies the two requirements for a density curve.

- (b) What percent of the observations are greater than 0.8?

- (c) What percent of the observations lie between 0.25 and 0.75?

29.

Biking accidents What is the mean μ of the density curve pictured in **Exercise 27**? (That is, where would the curve balance?) What is the median? (That is, where is the point with area 0.5 on either side?)
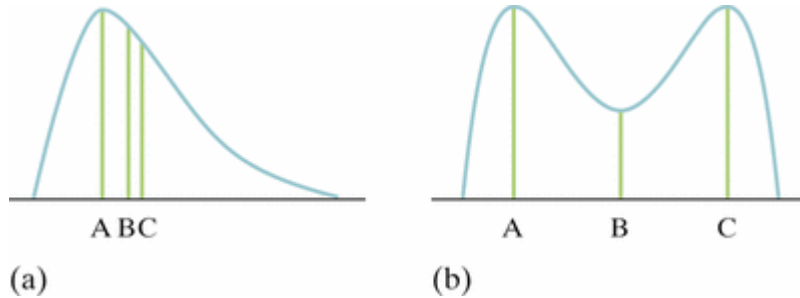
Correct Answer

Both are 1.5.

30.

A uniform distribution What is the mean μ of the density curve pictured in **Exercise 28**? What is the median?

31.

Mean and median The figure below displays two density curves, each with three points marked. At which of these points on each curve do the mean and the median fall?

(a)   (b)

Hide Answer
Correct Answer

(a) Mean is C, median is B. **(b)** Mean is B, median is B.

32.

Mean and median The figure below displays two density curves, each with three points marked. At which of these points on each curve do the mean and the median fall?
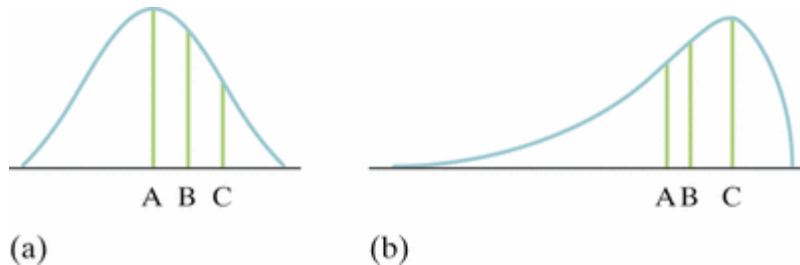

(a)   (b)

Multiple choice: Select the best answer for Exercises **33** *to* **38.**

33.

Jorge's score on Exam 1 in his statistics class was at the 64th percentile of the scores for all students. His score falls

- (a) between the minimum and the first quartile.

- (b) between the first quartile and the median.

- (c) between the median and the third quartile.

- (d) between the third quartile and the maximum.
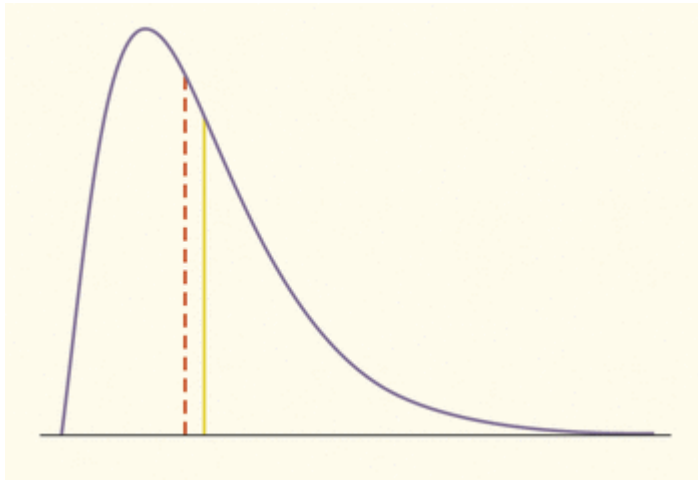
- (e) at the mean score for all students.

Correct Answer

c

34.

Two measures of center are marked on the density curve shown.

- (a) The median is at the yellow line and the mean is at the red line.

- (b) The median is at the red line and the mean is at the yellow line.

- (c) The mode is at the red line and the median is at the yellow line.

- (d) The mode is at the yellow line and the median is at the red line.

- (e) The mode is at the red line and the mean is at the yellow line.

35.

Scores on the ACT college entrance exam follow a bell-shaped distribution with mean 18 and standard deviation 6. Wayne's standardized score on the ACT was −0.7. What was Wayne's actual ACT score?

- (a) 4.2

- (b) −4.2

- (c) 13.8

- (d) 17.3

- (e) 22.2

Correct Answer

c

36.

George has an average bowling score of 180 and bowls in a league where the average for all bowlers is 150 and the standard deviation is 20. Bill has an average bowling score of 190 and bowls in a league where the average is 160 and the standard deviation is 15. Who ranks higher in his own league, George or Bill?

- (a) Bill, because his 190 is higher than George's 180.

- (b) Bill, because his standardized score is higher than George's.

- (c) Bill and George have the same rank in their leagues, because both are 30 pins above the mean.

- (d) George, because his standardized score is higher than Bill's.

- (e) George, because the standard deviation of bowling scores is higher in his league.

37.

If 30 is added to every observation in a data set, the only one of the following that is *not* changed is

- (a) the mean.

- (b) the 75th percentile.

- (c) the median.

- (d) the standard deviation.

- (e) the minimum.

Correct Answer

d

38.

If every observation in a data set is multiplied by 10, the only one of the following that is *not* multiplied by 10 is

- (a) the mean.

- (b) the median.

- (c) the *IQR*.

- (d) the standard deviation.

- (e) the variance.

Exercises 39 and 40 refer to the following setting. We used CensusAtSchool's Random Data Selector to choose a sample of 50 Canadian students who completed a survey in 2007–2008.

39.

Travel time (1.2) The dotplot below displays data on students' responses to the question "How long does it usually take you to travel to school?" Describe the shape, center, and spread of the distribution. Are there any outliers?

Travel time (minutes)
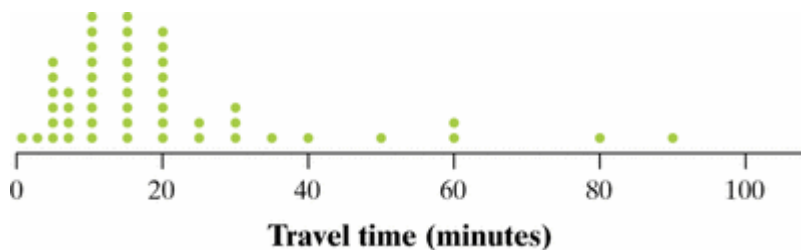
Hide Answer
Correct Answer

The distribution is skewed to the right since most of the values are 25 minutes or less, but the values stretch out up to about 90 minutes. The data are centered roughly around 20 minutes and the range of the distribution is close to 90 minutes. The two largest values appear to be outliers.

40.

Lefties (1.1) Students were asked, "Are you right-handed, left-handed, or ambidextrous?" The responses are shown below (R = right-handed; L = left-handed; A = ambidextrous).

R	R	R	R	R	R	R	R	R	R	R	L	R	R
R	R	R	R	R	R	R	R	R	R	R	R	R	A
R	R	R	R	A	R	R	L	R	R	R	R	L	A
R	R	R	R	R	R	R	R						

- (a) Make an appropriate graph to display these data.

- (b) Over 10,000 Canadian high school students took the CensusAtSchool survey in 2007–2008. What percent of this population would you estimate is left-handed? Justify your answer.

SECTION 2.1 Exercises