2.E: Descriptive Statistics (Exercises)

These are homework exercises to accompany the Textmap created for "Introductory Statistics" by Shafer and Zhang.

2.1: Three popular data displays

Basic

1. Describe one difference between a frequency histogram and a relative frequency histogram.
2. Describe one advantage of a stem and leaf diagram over a frequency histogram.
3. Construct a stem and leaf diagram, a frequency histogram, and a relative frequency histogram for the following data set. For the histograms use classes $51-60$, $61-70$, and so on.

$$\begin{array}{ccccc} 9 & 92 & 68 & 77 & 80 \\ 70 & 85 & 88 & 85 & 96 \\ 93 & 75 & 76 & 82 & 100 \\ 53 & 70 & 70 & 82 & 85 \end{array} \tag{2.E.1}$$

4. Construct a stem and leaf diagram, a frequency histogram, and a relative frequency histogram for the following data set. For the histograms use classes $6.0-6.9$, $7.0-7.9$, and so on.

$$\begin{array}{ccccc} .5 & 8.2 & 7.0 & 7.0 & 4.9 \\ 6.5 & 8.2 & 7.6 & 1.5 & 9.3 \\ 9.6 & 8.5 & 8.8 & 8.5 & 8.7 \\ 8.0 & 7.7 & 2.9 & 9.2 & 6.9 \end{array} \tag{2.E.2}$$

5. A data set contains $n=10$ observations. The values x and their frequencies f are summarized in the following data frequency table.

$$\begin{array}{c|cccc} x & -1 & 0 & 1 & 2 \\ \hline f & 3 & 4 & 2 & 1 \end{array} \tag{2.E.3}$$

   Construct a frequency histogram and a relative frequency histogram for the data set.
6. A data set contains the $n=20$ observations The values x and their frequencies f are summarized in the following data frequency table.

$$\begin{array}{c|cccc} x & -1 & 0 & 1 & 2 \\ \hline f & 3 & a & 2 & 1 \end{array} \tag{2.E.4}$$

   The frequency of the value 0 is missing. Find a and then sketch a frequency histogram and a relative frequency histogram for the data set.
7. A data set has the following frequency distribution table:

$$\begin{array}{c|cccc} x & 1 & 2 & 3 & 4 \\ \hline f & 3 & a & 2 & 1 \end{array} \tag{2.E.5}$$

   The number a is unknown. Can you construct a frequency histogram? If so, construct it. If not, say why not.
8. A table of some of the relative frequencies computed from a data set is

$$\begin{array}{c|cccc} x & 1 & 2 & 3 & 4 \\ \hline f/n & 0.3 & p & 0.2 & 0.1 \end{array} \tag{2.E.6}$$

   The number p is yet to be computed. Finish the table and construct the relative frequency histogram for the data set.

Applications

9. The IQ scores of ten students randomly selected from an elementary school are given.

$$\begin{array}{ccccc} 08 & 100 & 99 & 125 & 87 \\ 105 & 107 & 105 & 119 & 118 \end{array} \tag{2.E.7}$$

   Grouping the measures in the $80s$, the $90s$, and so on, construct a stem and leaf diagram, a frequency histogram, and a relative frequency histogram.
10. The IQ scores of ten students randomly selected from an elementary school for academically gifted students are given.

$$\begin{array}{ccccc} 33 & 140 & 152 & 142 & 137 \\ 145 & 160 & 138 & 139 & 138 \end{array} \tag{2.E.8}$$

    Grouping the measures by their common hundreds and tens digits, construct a stem and leaf diagram, a frequency histogram, and a relative frequency histogram.

11. During a one-day blood drive 300 people donated blood at a mobile donation center. The blood types of these 300 donors are summarized in the table.

$Blood\,Type$	O	A	B	AB
$Frequency$	136	120	32	12

(2.E.9)

Construct a relative frequency histogram for the data set.

12. In a particular kitchen appliance store an electric automatic rice cooker is a popular item. The weekly sales for the last 20weeks are shown.

$$\begin{array}{ccccc} 0 & 15 & 14 & 14 & 18 \\ 15 & 17 & 16 & 16 & 18 \\ 15 & 19 & 12 & 13 & 9 \\ 19 & 15 & 15 & 16 & 15 \end{array} \tag{2.E.10}$$

Construct a relative frequency histogram with classes $6-10$, $11-15$, and $16-20$.

Additional Exercises

13. Random samples, each of size $n=10$, were taken of the lengths in centimeters of three kinds of commercial fish, with the following results:

$$\begin{array}{lcccccccccc} Sample\,1: & 108 & 100 & 99 & 125 & 87 & 105 & 107 & 105 & 119 & 118 \\ Sample\,2: & 133 & 140 & 152 & 142 & 137 & 145 & 160 & 138 & 139 & 138 \\ Sample\,3: & 82 & 60 & 83 & 82 & 82 & 74 & 79 & 82 & 80 & 80 \end{array} \tag{2.E.11}$$

Grouping the measures by their common hundreds and tens digits, construct a stem and leaf diagram, a frequency histogram, and a relative frequency histogram for each of the samples. Compare the histograms and describe any patterns they exhibit.

14. During a one-day blood drive 300 people donated blood at a mobile donation center. The blood types of these 300 donors are summarized below.

$Blood\,Type$	O	A	B	AB
$Frequency$	136	120	32	12

(2.E.12)

Identify the blood type that has the highest relative frequency for these 300 people. Can you conclude that the blood type you identified is also most common for all people in the population at large? Explain.

15. In a particular kitchen appliance store, the weekly sales of an electric automatic rice cooker for the last 20 weeks are as follows.

$$\begin{array}{ccccc} 0 & 15 & 14 & 14 & 18 \\ 15 & 17 & 16 & 16 & 18 \\ 15 & 19 & 12 & 13 & 9 \\ 19 & 15 & 15 & 16 & 15 \end{array} \tag{2.E.13}$$

In retail sales, too large an inventory ties up capital, while too small an inventory costs lost sales and customer satisfaction. Using the relative frequency histogram for these data, find approximately how many rice cookers must be in stock at the beginning of each week if

a. the store is not to run out of stock by the end of a week for more than 15% of the weeks; and

b. the store is not to run out of stock by the end of a week for more than 5% of the weeks.

16. In retail sales, too large an inventory ties up capital, while too small an inventory costs lost sales and customer satisfaction. Using the relative frequency histogram for these data, find approximately how many rice cookers must be in stock at the beginning of each week if the store is not to run out of stock by the end of a week for more than 15% of the weeks; and the store is not to run out of stock by the end of a week for more than 5% of the weeks.

Answers

1. The vertical scale on one is the frequencies and on the other is the relative frequencies.
2.
3.

5	3						
6	8	9					
7	0	0	0	5	6	7	
8	0	2	3	5	5	5	8
9	2	3	6				
10	0						

(2.E.14)

4.
5. Noting that $n=10$ the relative frequency table is:

x	-1	0	1	2
f/n	0.3	0.4	0.2	0.1

(2.E.15)

6.

7. Since n is unknown, a is unknown, so the histogram cannot be constructed.

8.

9.

8	7
9	9
10	$0\quad 5\quad 5\quad 7\quad 8$
11	$8\quad 9$
12	5

(2.E.16)

Frequency and relative frequency histograms are similarly generated.

10.

11. Noting $n = 300$, the relative frequency table is therefore:

$Blood\,Type$	O	A	B	AB
f/n	0.4533	0.4	0.1067	0.04

(2.E.17)

A relative frequency histogram is then generated.

12.

13. The stem and leaf diagrams listed for Samples $1,\ 2,$ and 3 in that order:

6	
7	
8	7
9	9
10	$0\quad 5\quad 5\quad 7\quad 8$
11	$8\quad 9$
12	5
13	
14	
15	
16	

6	
7	
8	
9	
10	
11	
12	
13	$3\quad 7\quad 8\quad 8\quad 9$
14	$0\quad 2\quad 5$
15	2
16	0

(2.E.18)

(2.E.19)

6	0
7	4 9
8	0 0 2 2 2 2 3
9	
10	
11	
12	
13	
14	
15	
16	

(2.E.20)

The frequency tables are given below in the same order:

$Length$	$80 \sim 89$	$90 \sim 99$	$100 \sim 109$
f	1	1	5

(2.E.21)

$Length$	$110 \sim 119$	$120 \sim 129$
f	2	1

(2.E.22)

$Length$	$130 \sim 139$	$140 \sim 149$	$150 \sim 159$
f	5	3	1

(2.E.23)

$Length$	$160 \sim 169$
f	1

(2.E.24)

$Length$	$60 \sim 69$	$70 \sim 79$	$80 \sim 89$
f	1	2	7

(2.E.25)

The relative frequency tables are also given below in the same order:

$Length$	$80 \sim 89$	$90 \sim 99$	$100 \sim 109$
f / n	0.1	0.1	0.5

(2.E.26)

$Length$	$110 \sim 119$	$120 \sim 129$
f / n	0.2	0.1

(2.E.27)

$Length$	$130 \sim 139$	$140 \sim 149$	$150 \sim 159$
f / n	0.5	0.3	0.1

(2.E.28)

$Length$	$160 \sim 169$
f / n	0.1

(2.E.29)

$Length$	$60 \sim 69$	$70 \sim 79$	$80 \sim 89$
f / n	0.1	0.2	0.7

(2.E.30)

14.

15. a. 19
    b. 20

16.

2.2: Measures of Central Location

Basic

1. For the sample data set $\{1, 2, 6\}$ find
    1. $\sum x$
    2. $\sum x^2$
    3. $\sum (x - 3)$
    4. $\sum (x - 3)^2$
2. For the sample data set $\{-1, 0, 1, 4\}$ find

1. $\sum x$
2. $\sum x^2$
3. $\sum (x-1)$
4. $\sum (x-1)^2$

3. Find the mean, the median, and the mode for the sample

$$1\ 2\ 3\ 4 \tag{2.E.31}$$

4. Find the mean, the median, and the mode for the sample

$$3\ 3\ 4\ 4 \tag{2.E.32}$$

5. Find the mean, the median, and the mode for the sample

$$2\ 1\ 2\ 7 \tag{2.E.33}$$

6. Find the mean, the median, and the mode for the sample

$$-1\ 0\ 1\ 4\ 1\ 1 \tag{2.E.34}$$

7. Find the mean, the median, and the mode for the sample data represented by the table

x	1	2	7
f	1	2	1

(2.E.35)

8. Find the mean, the median, and the mode for the sample data represented by the table

x	-1	0	1	4
f	1	1	3	1

(2.E.36)

9. Create a sample data set of size $n = 3$ for which the mean $\bar{x}$ is greater than the median $\tilde{x}$.
10. Create a sample data set of size $n = 3$ for which the mean $\bar{x}$ is less than the median $\tilde{x}$.
11. Create a sample data set of size $n = 4$ for which the mean $\bar{x}$, the median $\tilde{x}$, and the mode are all identical.
12. Create a sample data set of size $n = 4$ for which the median $\tilde{x}$ and the mode are identical but the mean $\bar{x}$ is different.

Applications

13. Find the mean and the median for the LDL cholesterol level in a sample of ten heart patients.

$$\begin{array}{ccccc} 132 & 162 & 133 & 145 & 148 \\ 139 & 147 & 160 & 150 & 153 \end{array} \tag{2.E.37}$$

14. Find the mean and the median, for the LDL cholesterol level in a sample of ten heart patients on a special diet.

$$\begin{array}{ccccc} 127 & 152 & 138 & 110 & 152 \\ 113 & 131 & 148 & 135 & 158 \end{array} \tag{2.E.38}$$

15. Find the mean, the median, and the mode for the number of vehicles owned in a survey of 52 households.

x	0	1	2	3	4	5	6	7
f	2	12	15	11	6	3	1	2

(2.E.39)

16. The number of passengers in each of 120 randomly observed vehicles during morning rush hour was recorded, with the following results.

x	1	2	3	4	5
f	84	29	3	3	1

(2.E.40)

Find the mean, the median, and the mode of this data set.

17. Twenty-five $1-lb$ boxes of $16d$ nails were randomly selected and the number of nails in each box was counted, with the following results.

x	47	48	49	50	51
f	1	3	18	2	1

(2.E.41)

Find the mean, the median, and the mode of this data set.

Additional Exercises

18. Five laboratory mice with thymus leukemia are observed for a predetermined period of 500 days. After 500 days, four mice have died but the fifth one survives. The recorded survival times for the five mice are

$$493 \quad 421 \quad 222 \quad 378 \quad 500^* \tag{2.E.42}$$

where 500^* indicates that the fifth mouse survived for at least 500 days but the survival time (i.e., the exact value of the observation) is unknown.

a. Can you find the sample mean for the data set? If so, find it. If not, why not?
b. Can you find the sample median for the data set? If so, find it. If not, why not?

19. Five laboratory mice with thymus leukemia are observed for a predetermined period of 500 days. After 450 days, three mice have died, and one of the remaining mice is sacrificed for analysis. By the end of the observational period, the last remaining mouse still survives. The recorded survival times for the five mice are

$$222 \quad 421 \quad 378 \quad 450^* \quad 500^* \tag{2.E.43}$$

where * indicates that the mouse survived for at least the given number of days but the exact value of the observation is unknown.

a. Can you find the sample mean for the data set? If so, find it. If not, explain why not.
b. Can you find the sample median for the data set? If so, find it. If not, explain why not.

20. A player keeps track of all the rolls of a pair of dice when playing a board game and obtains the following data.

x	2	3	4	5	6	7
f	10	29	40	56	68	77

(2.E.44)

x	8	9	10	11	12
f	67	55	39	28	11

(2.E.45)

Find the mean, the median, and the mode.

21. Cordelia records her daily commute time to work each day, to the nearest minute, for two months, and obtains the following data.

x	26	27	28	29	30	31	32
f	3	4	16	12	6	2	1

(2.E.46)

a. Based on the frequencies, do you expect the mean and the median to be about the same or markedly different, and why?
b. Compute the mean, the median, and the mode.

22. An ordered stem and leaf diagram gives the scores of 71 students on an exam.

10	0	0																
9	1	1	1	1	2	3												
8	0	1	1	2	2	3	4	5	7	8	8	9						
7	0	0	0	1	1	2	4	4	5	6	6	6	7	7	7	8	8	9
6	0	1	2	2	2	3	4	4	5	7	7	7	7	8	8			
5	0	2	3	3	4	4	6	7	7	8	9							
4	2	5	6	8	8													
3	9	9																

(2.E.47)

a. Based on the shape of the display, do you expect the mean and the median to be about the same or markedly different, and why?
b. Compute the mean, the median, and the mode.

23. A man tosses a coin repeatedly until it lands heads and records the number of tosses required. (For example, if it lands heads on the first toss he records a 1; if it lands tails on the first two tosses and heads on the third he records a 3.) The data are shown.

x	1	2	3	4	5	6	7	8	9	10
f	384	208	98	56	28	12	8	2	3	1

(2.E.48)

a. Find the mean of the data.
b. Find the median of the data.

24. a. Construct a data set consisting of ten numbers, all but one of which is above average, where the average is the mean.
b. Is it possible to construct a data set as in part (a) when the average is the median? Explain.

25. Show that no matter what kind of average is used (mean, median, or mode) it is impossible for all members of a data set to be above average.

26. a. Twenty sacks of grain weigh a total of $1,003$ lb What is the mean weight per sack?
b. Can the median weight per sack be calculated based on the information given? If not, construct two data sets with the same total but different medians.

27. Begin with the following set of data, call it Data Set I.

$$5 \quad -2 \quad 6 \quad 14 \quad -3 \quad 0 \quad 1 \quad 4 \quad 3 \quad 2 \quad 5 \tag{2.E.49}$$

a. Compute the mean, median, and mode.

b. Form a new data set, Data Set II, by adding 3 to each number in Data Set I. Calculate the mean, median, and mode of Data Set II.
c. Form a new data set, Data Set III, by subtracting 6 from each number in Data Set I. Calculate the mean, median, and mode of Data Set III.
d. Comparing the answers to parts (a), (b), and (c), can you guess the pattern? State the general principle that you expect to be true.

Large Data Set Exercises

Note: For Large Data Set Exercises below, all of the data sets associated with these questions are missing, but the questions themselves are included here for reference.

28. Large Data Set 1 lists the SAT scores and GPAs of $1,000$ students.
    a. Compute the mean and median of the $1,000$ SAT scores.
    b. Compute the mean and median of the $1,000$ GPAs.
29. Large Data Set 1 lists the SAT scores of $1,000$ students.
    a. Regard the data as arising from a census of all students at a high school, in which the SAT score of every student was measured. Compute the population mean μ.
    b. Regard the first 25 observations as a random sample drawn from this population. Compute the sample mean $\bar{x}$ and compare it to μ.
    c. Regard the next 25 observations as a random sample drawn from this population. Compute the sample mean $\bar{x}$ and compare it to μ.
30. Large Data Set 1 lists the GPAs of $1,000$ students.
    a. Regard the data as arising from a census of all freshman at a small college at the end of their first academic year of college study, in which the GPA of every such person was measured. Compute the population mean μ.
    b. Regard the first 25 observations as a random sample drawn from this population. Compute the sample mean $\bar{x}$ and compare it to μ.
    c. Regard the next 25 observations as a random sample drawn from this population. Compute the sample mean $\bar{x}$ and compare it to μ.
31. Large Data Sets 7, $7A$, and $7B$ list the survival times in days of 140 laboratory mice with thymic leukemia from onset to death.
    a. Compute the mean and median survival time for all mice, without regard to gender.
    b. Compute the mean and median survival time for the 65 male mice (separately recorded in Large Data Set 7A).
    c. Compute the mean and median survival time for the 75 female mice (separately recorded in Large Data Set 7B).

Answers

1. a. 9
   b. 41
   c. 0
   d. 14
2.
3. $\bar{x} = 2.5,\ \tilde{x} = 2.5,\ \text{mode} = \{1, 2, 3, 4\}$
4.
5. $\bar{x} = 3,\ \tilde{x} = 2,\ \text{mode} = 2$
6.
7. $\bar{x} = 3,\ \tilde{x} = 2,\ \text{mode} = 2$
8.
9. $\{0, 0, 3\}$
10.
11. $\{0, 1, 1, 2\}$
12.
13. $\bar{x} = 146.9,\ \tilde{x} = 147.5$
14.
15. $\bar{x} = 2.6,\ \tilde{x} = 2,\ \text{mode} = 2$
16.
17. $\bar{x} = 48.96,\ \tilde{x} = 49,\ \text{mode} = 49$
18.
19. a. No, the survival times of the fourth and fifth mice are unknown.
    b. Yes, $\tilde{x} = 421$.
20.
21. $\bar{x} = 28.55,\ \tilde{x} = 28,\ \text{mode} = 28$
22.
23. $\bar{x} = 2.05,\ \tilde{x} = 2,\ \text{mode} = 1$
24.
25. **Mean**: $nx_{min} \leq \sum x$ so dividing by n yields $x_{min} \leq \bar{x}$, so the minimum value is not above average. **Median**: the middle measurement, or average of the two middle measurements, $\tilde{x}$, is at least as large as x_{min}, so the minimum value is not above average. **Mode**: the mode is one of

the measurements, and is not greater than itself

26.

27. a. $\bar{x} = 3.18,\ \tilde{x} = 3,\ \text{mode} = 5$
    b. $\bar{x} = 6.18,\ \tilde{x} = 6,\ \text{mode} = 8$
    c. $\bar{x} = -2.81,\ \tilde{x} = -3,\ \text{mode} = -1$
    d. If a number is added to every measurement in a data set, then the mean, median, and mode all change by that number.

28.

29. a. $\mu = 1528.74$
    b. $\bar{x} = 1502.8$
    c. $\bar{x} = 1532.2$

30.

31. a. $\bar{x} = 553.4286,\ \tilde{x} = 552.5$
    b. $\bar{x} = 665.9692,\ \tilde{x} = 667$
    c. $\bar{x} = 455.8933,\ \tilde{x} = 448$

2.3 Measures of Variability

Basic

1. Find the range, the variance, and the standard deviation for the following sample.

$$1\ 2\ 3\ 4 \tag{2.E.50}$$

2. Find the range, the variance, and the standard deviation for the following sample.

$$2\ -3\ 6\ 0\ 3\ 1 \tag{2.E.51}$$

3. Find the range, the variance, and the standard deviation for the following sample.

$$2\ 1\ 2\ 7 \tag{2.E.52}$$

4. Find the range, the variance, and the standard deviation for the following sample.

$$-1\ 0\ 1\ 4\ 1\ 1 \tag{2.E.53}$$

5. Find the range, the variance, and the standard deviation for the sample represented by the data frequency table.

x	1	2	7
f	1	2	1

(2.E.54)

6. Find the range, the variance, and the standard deviation for the sample represented by the data frequency table.

x	-1	0	1	4
f	1	1	3	1

(2.E.55)

Applications

7. Find the range, the variance, and the standard deviation for the sample of ten IQ scores randomly selected from a school for academically gifted students.

$$\begin{array}{ccccc} 132 & 162 & 133 & 145 & 148 \\ 139 & 147 & 160 & 150 & 153 \end{array} \tag{2.E.56}$$

8. Find the range, the variance and the standard deviation for the sample of ten IQ scores randomly selected from a school for academically gifted students.

$$\begin{array}{ccccc} 142 & 152 & 138 & 145 & 148 \\ 139 & 147 & 155 & 150 & 153 \end{array} \tag{2.E.57}$$

Additional Exercises

9. Consider the data set represented by the table

x	26	27	28	29	30	31	32
f	3	4	16	12	6	2	1

(2.E.58)

    a. Use the frequency table to find that $\sum x = 1256$ and $\sum x^2 = 35,926$.
    b. Use the information in part (a) to compute the sample mean and the sample standard deviation.

10. Find the sample standard deviation for the data

x	1	2	3	4	5
f	384	208	98	56	28

(2.E.59)

x	6	7	8	9	10
f	12	8	2	3	1

(2.E.60)

11. A random sample of 49 invoices for repairs at an automotive body shop is taken. The data are arrayed in the stem and leaf diagram shown. (Stems are thousands of dollars, leaves are hundreds, so that for example the largest observation is $3,800$.)

3	5	6	8							
3	0	0	1	1	2	4				
2	5	6	6	7	7	8	8	9	9	
2	0	0	0	0	1	2	2	4		
1	5	5	5	6	6	7	7	7	8	8	9
1	0	0	1	3	4	4	4			
0	5	6	8	8						
0	4									

(2.E.61)

For these data, $\sum x = 101$, $\sum x^2 = 244,830,000$.

a. Compute the mean, median, and mode.
b. Compute the range.
c. Compute the sample standard deviation.

12. What must be true of a data set if its standard deviation is 0?

13. A data set consisting of 25 measurements has standard deviation 0. One of the measurements has value 17. What are the other 24 measurements?

14. Create a sample data set of size $n = 3$ for which the range is 0 and the sample mean is 2.

15. Create a sample data set of size $n = 3$ for which the sample variance is 0 and the sample mean is 1.

16. The sample $\{-1, 0, 1\}$ has mean $\bar{x} = 0$ and standard deviation $\bar{x} = 0$. Create a sample data set of size $n = 3$ for which $\bar{x} = 0$ and s is greater than 1.

17. The sample $\{-1, 0, 1\}$ has mean $\bar{x} = 0$ and standard deviation $\bar{x} = 0$. Create a sample data set of size $n = 3$ for which $\bar{x} = 0$ and the standard deviation s is less than 1.

18. Begin with the following set of data, call it Data Set I.

$$5 \ -2 \ 6 \ 1 \ 4 \ -3 \ 0 \ 1 \ 4 \ 3 \ 2 \ 5 \tag{2.E.62}$$

a. Compute the sample standard deviation of Data Set I.
b. Form a new data set, Data Set II, by adding 3 to each number in Data Set I. Calculate the sample standard deviation of Data Set II.
c. Form a new data set, Data Set III, by subtracting 6 from each number in Data Set I. Calculate the sample standard deviation of Data Set III.
d. Comparing the answers to parts (a), (b), and (c), can you guess the pattern? State the general principle that you expect to be true.

Large Data Set Exercises

Note: For Large Data Set Exercises below, all of the data sets associated with these questions are missing, but the questions themselves are included here for reference.

19. Large Data Set 1 lists the SAT scores and GPAs of $1,000$ students.

a. Compute the range and sample standard deviation of the $1,000$ SAT scores.
b. Compute the range and sample standard deviation of the $1,000$ GPAs.

20. Large Data Set 1 lists the SAT scores of $1,000$ students.

a. Regard the data as arising from a census of all students at a high school, in which the SAT score of every student was measured. Compute the population range and population standard deviation σ.
b. Regard the first 25 observations as a random sample drawn from this population. Compute the sample range and sample standard deviation s and compare them to the population range and σ.
c. Regard the next 25 observations as a random sample drawn from this population. Compute the sample range and sample standard deviation s and compare them to the population range and σ.

21. Large Data Set 1 lists the GPAs of $1,000$ students.

a. Regard the data as arising from a census of all freshman at a small college at the end of their first academic year of college study, in which the GPA of every such person was measured. Compute the population range and population standard deviation σ.

b. Regard the first 25 observations as a random sample drawn from this population. Compute the sample range and sample standard deviation s and compare them to the population range and σ.
c. Regard the next 25 observations as a random sample drawn from this population. Compute the sample range and sample standard deviation s and compare them to the population range and σ.

22. Large Data Set 7, 7A, and 7B list the survival times in days of 140 laboratory mice with thymic leukemia from onset to death.

a. Compute the range and sample standard deviation of survival time for all mice, without regard to gender.
b. Compute the range and sample standard deviation of survival time for the 65 male mice (separately recorded in Large Data Set 7A).
c. Compute the range and sample standard deviation of survival time for the 75 female mice (separately recorded in Large Data Set 7B). Do you see a difference in the results for male and female mice? Does it appear to be significant?

Answers

1. $R = 3,\ s^2 = 1.7,\ s = 1.3$.
2.
3. $R = 6,\ s^2 = 7.\bar{3},\ s = 2.7$.
4.
5. $R = 6,\ s^2 = 7.3,\ s = 2.7$.
6.
7. $R = 30,\ s^2 = 103.2,\ s = 10.2$.
8.
9. $\bar{x} = 28.55,\ s = 1.3$.
10.
11. a. $\bar{x} = 2063,\ \tilde{x} = 2000,\ \text{mode} = 2000$.
    b. $R = 3400$.
    c. $s = 869$.
12.
13. All are 17.
14.
15. $\{1, 1, 1\}$
16.
17. One example is $\{-.5, 0, .5\}$.
18.
19. a. $R = 1350$ and $s = 212.5455$
    b. $R = 4.00$ and $s = 0.7407$
20.
21. a. $R = 4.00$ and $\sigma = 0.740375$
    b. $R = 3.04$ and $s = 0.808045$
    c. $R = 2.49$ and $s = 0.657843$

2.4 Relative Position of Data

Basic

1. Consider the data set

$$\begin{array}{ccccc} 69 & 92 & 68 & 77 & 80 \\ 93 & 75 & 76 & 82 & 100 \\ 70 & 85 & 88 & 85 & 96 \\ 53 & 70 & 70 & 82 & 85 \end{array} \tag{2.E.63}$$

a. Find the percentile rank of 82.
b. Find the percentile rank of 68.

2. Consider the data set

$$\begin{array}{ccccc} 8.5 & 8.2 & 7.0 & 7.0 & 4.9 \\ 9.6 & 8.5 & 8.8 & 8.5 & 8.7 \\ 6.5 & 8.2 & 7.6 & 1.5 & 9.3 \\ 8.0 & 7.7 & 2.9 & 9.2 & 6.9 \end{array} \tag{2.E.64}$$

a. Find the percentile rank of 6.5.
b. Find the percentile rank of 7.7.

3. Consider the data set represented by the ordered stem and leaf diagram

10	0	0																
9	1	1	1	1	2	3												
8	0	1	1	2	2	3	4	5	7	8	8	9						
7	0	0	0	1	1	2	4	4	5	6	6	6	7	7	7	8	8	9
6	0	1	2	2	2	3	4	4	5	7	7	7	7	8	8			
5	0	2	3	3	4	4	6	7	7	8	9							
4	2	5	6	8	8													
3	9	9																

(2.E.65)

   a. Find the percentile rank of the grade 75.
   b. Find the percentile rank of the grade 57.
4. Is the 90^{th} percentile of a data set always equal to 90%? Why or why not?
5. The 29^{th} percentile in a large data set is 5.
   a. Approximately what percentage of the observations are less than 5?
   b. Approximately what percentage of the observations are greater than 5?
6. The 54^{th} percentile in a large data set is 98.6.
   a. Approximately what percentage of the observations are less than 98.6?
   b. Approximately what percentage of the observations are greater than 98.6?
7. In a large data set the 29^{th} percentile is 5 and the 79^{th} percentile is 10. Approximately what percentage of observations lie between 5 and 10?
8. In a large data set the 40^{th} percentile is 125 and the 82^{nd} percentile is 158. Approximately what percentage of observations lie between 125 and 158?
9. Find the five-number summary and the IQR and sketch the box plot for the sample represented by the stem and leaf diagram in Figure 2.1.2 "Ordered Stem and Leaf Diagram".
10. Find the five-number summary and the IQR and sketch the box plot for the sample explicitly displayed in "Example 2.2.7" in Section 2.2.
11. Find the five-number summary and the IQR and sketch the box plot for the sample represented by the data frequency table

x	1	2	5	8	9
f	5	2	3	6	4

(2.E.66)

12. Find the five-number summary and the IQR and sketch the box plot for the sample represented by the data frequency table

x	-5	-3	-2	-1	0	1	3	4	5
f	2	1	3	2	4	1	1	2	1

(2.E.67)

13. Find the z-score of each measurement in the following sample data set.

$$-5 \;\; 6 \;\; 2 \;\; -1 \;\; 0 \tag{2.E.68}$$

14. Find the z-score of each measurement in the following sample data set.

$$1.6 \;\; 5.2 \;\; 2.8 \;\; 3.7 \;\; 4.0 \tag{2.E.69}$$

15. The sample with data frequency table

x	1	2	7
f	1	2	1

(2.E.70)

has mean $\bar{x} = 3$ and standard deviation $s \approx 2.71$. Find the z-score for every value in the sample.
16. The sample with data frequency table

x	-1	0	1	4
f	1	1	3	1

(2.E.71)

has mean $\bar{x} = 1$ and standard deviation $s \approx 1.67$. Find the z-score for every value in the sample.
17. For the population

$$0 \;\; 0 \;\; 2 \;\; 2 \tag{2.E.72}$$

compute each of the following.
   a. The population mean μ.
   b. The population variance σ^2.
   c. The population standard deviation σ.

d. The z-score for every value in the population data set.

18. For the population

$$0.5 \quad 2.1 \quad 4.4 \quad 1.0 \tag{2.E.73}$$

compute each of the following.

a. The population mean μ.
b. The population variance σ^2.
c. The population standard deviation σ.
d. The z-score for every value in the population data set.

19. A measurement x in a sample with mean $\bar{x} = 10$ and standard deviation $s = 3$ has z-score $z = 2$. Find x.
20. A measurement x in a sample with mean $\bar{x} = 10$ and standard deviation $s = 3$ has z-score $z = -1$. Find x.
21. A measurement x in a population with mean $\mu = 2.3$ and standard deviation $\sigma = 1.3$ has z-score $z = 2$. Find x.
22. A measurement x in a sample with mean $\mu = 2.3$ and standard deviation $\sigma = 1.3$ has z-score $z = -1.2$. Find x.

Applications

23. The weekly sales for the last 20 weeks in a kitchen appliance store for an electric automatic rice cooker are

$$\begin{array}{ccccc} 20 & 15 & 14 & 14 & 18 \\ 15 & 19 & 12 & 13 & 9 \\ 15 & 17 & 16 & 16 & 18 \\ 19 & 15 & 15 & 16 & 15 \end{array} \tag{2.E.74}$$

a. Find the percentile rank of 15.
b. If the sample accurately reflects the population, then what percentage of weeks would an inventory of 15 rice cookers be adequate?

24. The table shows the number of vehicles owned in a survey of 52 households.

x	0	1	2	3	4	5	6	7
f	2	12	15	11	6	3	1	2

(2.E.75)

a. Find the percentile rank of 2.
b. If the sample accurately reflects the population, then what percentage of households have at most two vehicles?

25. For two months Cordelia records her daily commute time to work each day to the nearest minute and obtains the following data:

x	26	27	28	29	30	31	32
f	3	4	16	12	6	2	1

(2.E.76)

Cordelia is supposed to be at work at $8:00\ a.m.$ but refuses to leave her house before $7:30\ a.m.$

a. Find the percentile rank of 30, the time she has to get to work.
b. Assuming that the sample accurately reflects the population of *all* of Cordelia's commute times, use your answer to part (a) to predict the proportion of the work days she is late for work.

26. The mean score on a standardized grammar exam is 49.6; the standard deviation is 1.35. Dromio is told that the z-score of his exam score is -1.19.

a. Is Dromio's score above average or below average?
b. What was Dromio's actual score on the exam?

27. A random sample of 49 invoices for repairs at an automotive body shop is taken. The data are arrayed in the stem and leaf diagram shown. (Stems are thousands of dollars, leaves are hundreds, so that for example the largest observation is $3,800$.)

3	5	6	8							
3	0	0	1	1	2	4				
2	5	6	6	7	7	8	8	9	9	
2	0	0	0	0	1	2	2	4		
1	5	5	5	6	6	7	7	7	8	8	9
1	0	0	1	3	4	4	4			
0	5	6	8	8						
0	4									

(2.E.77)

For these data, $\sum x = 101,100$, $\sum x^2 = 244,830,000$.

a. Find the z-score of the repair that cost $\$1,100$.
b. Find the z-score of the repairs that cost $\$2,700$.

28. The stem and leaf diagram shows the time in seconds that callers to a telephone-order center were on hold before their call was taken.

0	0 0 0 0 0 0 1 1 1 1 1 1 1 2 2 2 2 2 3 3 3 3 3 3 3 4 4 4 4 4
0	5 5 5 5 5 5 5 5 5 6 6 6 6 6 6 6 6 6 6 7 7 7 7 7 7 8 8 8 9 9
1	0 0 1 1 1 1 2 2 2 2 4 4
1	5 6 6 8 9
2	2 4
2	5
3	0

(2.E.78)

a. Find the quartiles.
b. Give the five-number summary of the data.
c. Find the range and the IQR.

Additional Exercises

29. Consider the data set represented by the ordered stem and leaf diagram

10	0 0
9	1 1 1 1 2 3
8	0 1 1 2 2 3 4 5 7 8 8 9
7	0 0 0 1 1 2 4 4 5 6 6 6 7 7 7 8 8 9
6	0 1 2 2 2 3 4 4 5 7 7 7 7 8 8
5	0 2 3 3 4 4 6 7 7 8 9
4	2 5 6 8 8
3	9 9

(2.E.79)

a. Find the three quartiles.
b. Give the five-number summary of the data.
c. Find the range and the IQR.

30. For the following stem and leaf diagram the units on the stems are thousands and the units on the leaves are hundreds, so that for example the largest observation is $3,800$.

3	5 6 8
3	0 0 1 1 2 4
2	5 6 6 7 7 8 8 9 9
2	0 0 0 0 1 2 2 4
1	5 5 5 6 6 7 7 7 8 8 9
1	0 0 1 3 4 4 4
0	5 6 8 8
0	4

(2.E.80)

a. Find the percentile rank of 800.
b. Find the percentile rank of $3,200$.

31. Find the five-number summary for the following sample data.

x	26	27	28	29	30	31	32
f	3	4	16	12	6	2	1

(2.E.81)

32. Find the five-number summary for the following sample data.

x	1	2	3	4	5	6	7	8	9	10
f	384	208	98	56	28	12	8	2	3	1

(2.E.82)

33. For the following stem and leaf diagram the units on the stems are thousands and the units on the leaves are hundreds, so that for example the largest observation is $3,800$.

3	5 6 8
3	0 0 1 1 2 4
2	5 6 6 7 7 8 8 9 9
2	0 0 0 0 1 2 2 4
1	5 5 5 6 6 7 7 7 8 8 9
1	0 0 1 3 4 4 4
0	5 6 8 8
0	4

(2.E.83)

a. Find the three quartiles.
b. Find the IQR.
c. Give the five-number summary of the data.

34. Determine whether the following statement is true. "In any data set, if an observation x_1 is greater than another observation x_2, then the z-score of x_1 is greater than the z-score of x_2."
35. Emilia and Ferdinand took the same freshman chemistry course, Emilia in the fall, Ferdinand in the spring. Emilia made an 83 on the common final exam that she took, on which the mean was 76 and the standard deviation 8. Ferdinand made a 79 on the common final exam that he took, which was more difficult, since the mean was 65 and the standard deviation 12. The one who has a higher z-score did relatively better. Was it Emilia or Ferdinand?
36. Refer to the previous exercise. On the final exam in the same course the following semester, the mean is 68 and the standard deviation is 9. What grade on the exam matches Emilia's performance? Ferdinand's?
37. Rosencrantz and Guildenstern are on a weight-reducing diet. Rosencrantz, who weighs $178\ lb$, belongs to an age and body-type group for which the mean weight is $145\ lb$ and the standard deviation is $15\ lb$. Guildenstern, who weighs $204\ lb$, belongs to an age and body-type group for which the mean weight is $165\ lb$ and the standard deviation is $20\ lb$. Assuming z-scores are good measures for comparison in this context, who is more overweight for his age and body type?

Large Data Set Exercises

Note: For Large Data Set Exercises below, all of the data sets associated with these questions are missing, but the questions themselves are included here for reference.

38. Large Data Set 1 lists the SAT scores and GPAs of $1,000$ students.
    a. Compute the three quartiles and the interquartile range of the $1,000$ SAT scores.
    b. Compute the three quartiles and the interquartile range of the $1,000$ GPAs.
39. Large Data Set 10 records the scores of 72 students on a statistics exam.
    a. Compute the five-number summary of the data.
    b. Describe in words the performance of the class on the exam in the light of the result in part (a).
40. Large Data Sets 3 and 3A list the heights of 174 customers entering a shoe store.
    a. Compute the five-number summary of the heights, without regard to gender.
    b. Compute the five-number summary of the heights of the men in the sample.
    c. Compute the five-number summary of the heights of the women in the sample.
41. Large Data Sets 7, 7A, and 3B list the survival times in days of 140 laboratory mice with thymic leukemia from onset to death.
    a. Compute the three quartiles and the interquartile range of the survival times for all mice, without regard to gender.
    b. Compute the three quartiles and the interquartile range of the survival times for the 65 male mice (separately recorded in Data Set 7A).
    c. Compute the three quartiles and the interquartile range of the survival times for the 75 female mice (separately recorded in Data Sets 7B).

Answer

1. a. 60
   b. 10
2.
3. a. 59
   b. 23
4.
5. a. 29
   b. 71
6.
7. 50%
8.
9. $x_{min} = 25,\ \ Q_1 = 70,\ \ Q_2 = 77.5\ \ Q_3 = 90,\ \ x_{max} = 100,\ \ IQR = 20$

10.
11. $x_{min} = 1,\ \ Q_1 = 1.5,\ \ Q_2 = 6.5\ \ Q_3 = 8,\ \ x_{max} = 9,\ \ IQR = 6.5$
12.
13. $-1.3,\ 1.39,\ 0.4,\ -0.35,\ -0.11$
14.
15. $z = -0.74$ for $x = 1,\ z = -0.37$ for $x = 2,\ z = 1.48$ for $x = 7$
16.
17. a. 1
    b. 1
    c. 1
    d. $z = -1$ for $x = 0,\ z = 1$ for $x = 2$
18.
19. 16
20.
21. 4.9
22.
23. a. 55
    b. 55
24.
25. a. 93
    b. 0.07
26.
27. a. -1.11
    b. 0.73
28.
29. a. $Q_1 = 59,\ Q_2 = 70,\ Q_3 = 81$
    b. $x_{min} = 39,\ Q_1 = 59,\ Q_2 = 70,\ Q_3 = 81,\ x_{max} = 100$
    c. $R = 61,\ IQR = 22$
30.
31. $x_{min} = 26,\ Q_1 = 28,\ Q_2 = 28,\ Q_3 = 29,\ x_{max} = 32$
32.
33. a. $Q_1 = 1450,\ Q_2 = 2000,\ Q_3 = 2800$
    b. $IQR = 1350$
    c. $x_{min} = 400,\ Q_1 = 1450,\ Q_2 = 2000,\ Q_3 = 2800,\ x_{max} = 3800$
34.
35. Emilia: $z = 0.875$, Ferdinand: $z = 1.1\bar{6}$
36.
37. Rosencrantz: $z = 2.2$, Guildenstern: $z = 1.95$. Rosencrantz is more overweight for his age and body type.
38.
39. a. $x_{min} = 15,\ Q_1 = 51,\ Q_2 = 67,\ Q_3 = 82,\ x_{max} = 97$
    b. The data set appears to be skewed to the left.
40.
41. a. $Q_1 = 440,\ Q_2 = 552.5,\ Q_3 = 661$ and $IQR = 221$
    b. $Q_1 = 641,\ Q_2 = 667,\ Q_3 = 700$ and $IQR = 59$
    c. $Q_1 = 407,\ Q_2 = 448,\ Q_3 = 504$ and $IQR = 97$

2.5 The Empirical Rule and Chebyshev's Theorem

Basic

1. State the Empirical Rule.
2. Describe the conditions under which the Empirical Rule may be applied.
3. State Chebyshev's Theorem.
4. Describe the conditions under which Chebyshev's Theorem may be applied.
5. A sample data set with a bell-shaped distribution has mean $\bar{x} = 6$ and standard deviation $s = 2$. Find the approximate proportion of observations in the data set that lie:
    a. between 4 and 8;
    b. between 2 and 10;

c. between 0 and 12.

6. A population data set with a bell-shaped distribution has mean $\mu = 6$ and standard deviation $\sigma = 2$. Find the approximate proportion of observations in the data set that lie:
   a. between 4 and 8;
   b. between 2 and 10;
   c. between 0 and 12.
7. A population data set with a bell-shaped distribution has mean $\mu = 2$ and standard deviation $\sigma = 1.1$. Find the approximate proportion of observations in the data set that lie:
   a. above 2;
   b. above 3.1;
   c. between 2 and 3.1.
8. A sample data set with a bell-shaped distribution has mean $\bar{x} = 2$ and standard deviation $s = 1.1$. Find the approximate proportion of observations in the data set that lie:
   a. below -0.2;
   b. below 3.1;
   c. between -1.3 and 0.9.
9. A population data set with a bell-shaped distribution and size $N = 500$ has mean $\mu = 2$ and standard deviation $\sigma = 1.1$. Find the approximate number of observations in the data set that lie:
   a. above 2;
   b. above 3.1;
   c. between 2 and 3.1.
10. A sample data set with a bell-shaped distribution and size $n = 128$ has mean $\bar{x} = 2$ and standard deviation $s = 1.1$. Find the approximate number of observations in the data set that lie:
    a. below -0.2;
    b. below 3.1;
    c. between -1.3 and 0.9.
11. A sample data set has mean $\bar{x} = 6$ and standard deviation $s = 2$. Find the minimum proportion of observations in the data set that must lie:
    a. between 2 and 10;
    b. between 0 and 12;
    c. between 4 and 8.
12. A population data set has mean $\mu = 2$ and standard deviation $\sigma = 1.1$. Find the minimum proportion of observations in the data set that must lie:
    a. between -0.2 and 4.2;
    b. between -1.3 and 5.3.
13. A population data set of size $N = 500$ has mean $\mu = 5.2$ and standard deviation $\sigma = 1.1$. Find the minimum number of observations in the data set that must lie:
    a. between 3 and 7.4;
    b. between 1.9 and 8.5.
14. A sample data set of size $n = 128$ has mean $\bar{x} = 2$ and standard deviation $s = 2$. Find the minimum number of observations in the data set that must lie:
    a. between -2 and 6 (including -2 and 6);
    b. between -4 and 8 (including -4 and 8).
15. A sample data set of size $n = 30$ has mean $\bar{x} = 6$ and standard deviation $s = 2$.
    a. What is the maximum proportion of observations in the data set that can lie outside the interval $(2, 10)$?
    b. What can be said about the proportion of observations in the data set that are below 2?
    c. What can be said about the proportion of observations in the data set that are above 10?
    d. What can be said about the number of observations in the data set that are above 10?
16. A population data set has mean $\mu = 2$ and standard deviation $\sigma = 1.1$.
    a. What is the maximum proportion of observations in the data set that can lie outside the interval $(-1.3, 5.3)$?
    b. What can be said about the proportion of observations in the data set that are below -1.3?
    c. What can be said about the proportion of observations in the data set that are above 5.3?

Applications

17. Scores on a final exam taken by $1,200$ students have a bell-shaped distribution with mean 72 and standard deviation 9.
    a. What is the median score on the exam?
    b. About how many students scored between 63 and 81?

c. About how many students scored between 72 and 90?
d. About how many students scored below 54?

18. Lengths of fish caught by a commercial fishing boat have a bell-shaped distribution with mean 23 inches and standard deviation 1.5 inches.
    a. About what proportion of all fish caught are between 20 inches and 26 inches long?
    b. About what proportion of all fish caught are between 20 inches and 23 inches long?
    c. About how long is the longest fish caught (only a small fraction of a percent are longer)?
19. Hockey pucks used in professional hockey games must weigh between 5.5 and 6 ounces. If the weight of pucks manufactured by a particular process is bell-shaped, has mean 5.75 ounces and standard deviation 0.125 ounce, what proportion of the pucks will be usable in professional games?
20. Hockey pucks used in professional hockey games must weigh between 5.5 and 6 ounces. If the weight of pucks manufactured by a particular process is bell-shaped and has mean 5.75 ounces, how large can the standard deviation be if 99.7% of the pucks are to be usable in professional games?
21. Speeds of vehicles on a section of highway have a bell-shaped distribution with mean 60 mph and standard deviation 2.5 mph.
    a. If the speed limit is 55 mph, about what proportion of vehicles are speeding?
    b. What is the median speed for vehicles on this highway?
    c. What is the percentile rank of the speed 65 mph?
    d. What speed corresponds to the 16_{th} percentile?
22. Suppose that, as in the previous exercise, speeds of vehicles on a section of highway have mean 60 mph and standard deviation 2.5 mph, but now the distribution of speeds is unknown.
    a. If the speed limit is 55 mph, at least what proportion of vehicles must speeding?
    b. What can be said about the proportion of vehicles going 65 mph or faster?
23. An instructor announces to the class that the scores on a recent exam had a bell-shaped distribution with mean 75 and standard deviation 5.
    a. What is the median score?
    b. Approximately what proportion of students in the class scored between 70 and 80?
    c. Approximately what proportion of students in the class scored above 85?
    d. What is the percentile rank of the score 85?
24. The GPAs of all currently registered students at a large university have a bell-shaped distribution with mean 2.7 and standard deviation 0.6. Students with a GPA below 1.5 are placed on academic probation. Approximately what percentage of currently registered students at the university are on academic probation?
25. Thirty-six students took an exam on which the average was 80 and the standard deviation was 6. A rumor says that five students had scores 61 or below. Can the rumor be true? Why or why not?

Additional Exercises

26. For the sample data

x	26	27	28	29	30	31	32
f	3	4	16	12	6	2	1

(2.E.84)

$$\sum x = 1,256 \text{ and } \sum x^2 = 35,926 \tag{2.E.85}$$

    a. Compute the mean and the standard deviation.
    b. About how many of the measurements does the Empirical Rule predict will be in the interval $(\bar{x}-s, \bar{x}+s)$, the interval $(\bar{x}-2s, \bar{x}+2s)$, and the interval $(\bar{x}-3s, \bar{x}+3s)$?
    c. Compute the number of measurements that are actually in each of the intervals listed in part (a), and compare to the predicted numbers.
27. A sample of size $n = 80$ has mean 139 and standard deviation 13, but nothing else is known about it.
    a. What can be said about the number of observations that lie in the interval $(126, 152)$?
    b. What can be said about the number of observations that lie in the interval $(113, 165)$?
    c. What can be said about the number of observations that exceed 165?
    d. What can be said about the number of observations that either exceed 165 or are less than 113?
28. For the sample data

x	1	2	3	4	5
f	84	29	3	3	1

(2.E.86)

$$\sum x = 168 \text{ and } \sum x^2 = 300 \tag{2.E.87}$$

    a. Compute the sample mean and the sample standard deviation.
    b. Considering the shape of the data set, do you expect the Empirical Rule to apply? Count the number of measurements within one standard deviation of the mean and compare it to the number predicted by the Empirical Rule.

c. What does Chebyshev's Rule say about the number of measurements within one standard deviation of the mean?
d. Count the number of measurements within two standard deviations of the mean and compare it to the minimum number guaranteed by Chebyshev's Theorem to lie in that interval.

29. For the sample data set

x	47	48	49	50	51
f	1	3	18	2	1

(2.E.88)

$$\sum x = 1224 \text{ and } \sum x^2 = 59,940 \tag{2.E.89}$$

a. Compute the sample mean and the sample standard deviation.
b. Considering the shape of the data set, do you expect the Empirical Rule to apply? Count the number of measurements within one standard deviation of the mean and compare it to the number predicted by the Empirical Rule.
c. What does Chebyshev's Rule say about the number of measurements within one standard deviation of the mean?
d. Count the number of measurements within two standard deviations of the mean and compare it to the minimum number guaranteed by Chebyshev's Theorem to lie in that interval.

Answers

1. See the displayed statement in the text.
2.
3. See the displayed statement in the text.
4.
5. a. 0.68
   b. 0.95
   c. 0.997
6.
7. a. 0.5
   b. 0.16
   c. 0.34
8.
9. a. 250
   b. 80
   c. 170
10.
11. a. $3/4$
    b. $8/9$
    c. 0
12.
13. a. 375
    b. 445
14.
15. a. At most 0.25.
    b. At most 0.25.
    c. At most 0.25.
    d. At most 7.
16.
17. a. 72
    b. 816
    c. 570
    d. 30
18.
19. 0.95
20.
21. a. 0.975
    b. 60
    c. 97.5
    d. 57.5

22.

23. a. 75
    b. 0.68
    c. 0.025
    d. 0.975

24.

25. By Chebyshev's Theorem at most $1/9$ of the scores can be below 62, so the rumor is impossible.

26.

27. a. Nothing.
    b. It is at least 60.
    c. It is at most 20.
    d. It is at most 20.

28.

29. a. $\bar{x} = 48.96$, $s = 0.7348$.
    b. Roughly bell-shaped, the Empirical Rule should apply. **True count**: 18, **Predicted**: 17.
    c. Nothing.
    d. **True count**: 23, **Guaranteed**: at least 18.75, hence at least 19.

Contributor

- Anonymous
