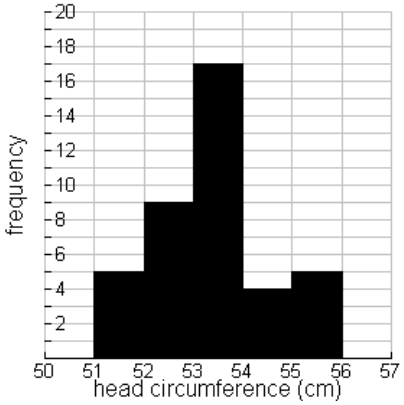
Statistics: Chapter 5 Solutions

Lesson 5.1.1

5-1. a. $\bar{x} = 53.3$ cm, $s = 1.12$ cm

b. See histogram at right.



c. 9 women

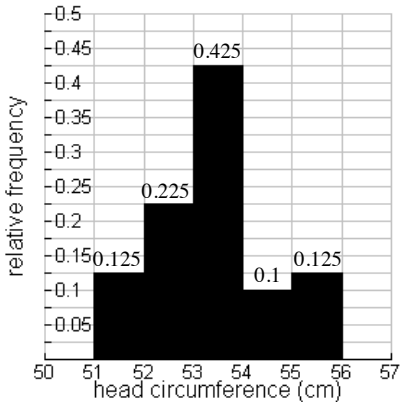
d. $\frac{5+9+17}{40} = 77.5\%$

5-2. a.

Women's hat sizes

size interval	frequency (# of hats)	relative frequency (proportion of hats)
50 to 51	0	0
51 to 52	5	0.125
52 to 53	9	0.225
53 to 54	17	**0.425**
54 to 55	**4**	**0.1**
55 to 56	**5**	**0.125**
56 to 57	**0**	**0**

b. See histogram at right; it is the same shape as the original histogram.

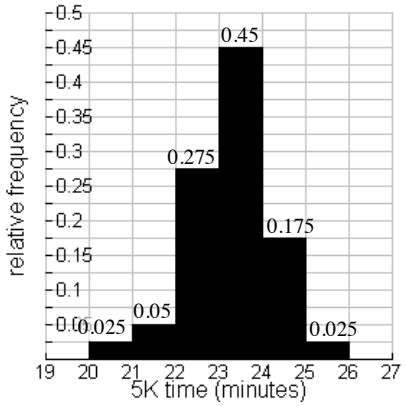


c. $0.225 + 0.425 + 0.1 = 75\%$

d. $0.125 + 0.225 + 0.425 + 0.1 + 0.125 = 100\%$

5-3. a. $\bar{x} = 23.2175$ minutes; Since this is a random sample of 5K racers, not the times for all racers, use the *sample* standard deviation, $s = 0.9729$ minutes.

b.



c. $0.05 + 0.025 = 7.5\%$ of racers

d. $0.275 + 0.45 + 0.175 = 90\%$. The area of the bars is the percentage of the sample.

5-4. a. *i*. $P(X < 35) = 0.28$

ii. $P(X > 65) = 0.05$

iii. $P(30 \le X \le 60) = 0.90 - 0.18 = 0.72$

iv. $P(0 \le X \le 90) = 1.00$

b. $X = 44$: $P(X \le 44) = 0.50$

5-5. a. z = 1; Caspar is 1 standard deviation above the mean.

b. z = 2; Ollie is 2 standard deviations above the mean.

c. Ollie has the best relative score. His z-score is greater than Caspar's z-score.

d. z = –1. Jasper is 1 standard deviation below the mean.

5-6. a. $P(X > 70) \approx 0.32$

b. $P(X < 60) \approx 0.50$

c. $P(80 < X < 120) \approx 0.20$

d. $P(30 < X < 130) \approx 1.00$

5-7. a. This should be an observational study. It would not be fair to ask students to arrive late for class if you really believed that there was a relationship between being tardy and lower grades.

b. School records should have information on how many times each student has been late to class and the grades they received on tests. If it is possible to hide the identity of the students, Lana should be able to take a census for her study. She would not need to randomize anything.

5-8. a. Yes, the residual plot shows random scatter so the linear model is appropriate.

b. The residual point appears to be at (16, 0.45) or thereabouts. The expected value according to the regression output is 0.2026(16) – 0.1845 = \$3.05, so the actual cost of the 16-nugget value is about \$3.50. (Note: the actual cost in the original data is \$3.49, but any value between about \$3.45 and \$3.60 is reasonable given the graph.)

c. According to the LSRL of the current data, the expected cost for 30 nuggets is \$5.90. The new point is clearly an outlier from the pattern established by the other data points, and so will decrease *r*.

5-9. $\frac{1}{10!} = \frac{1}{3628800} \approx 0.0000003$

