Chapter 5 review exercises

2. The printout does not look reasonable. Since these test scores are converted to standard units (i.e., Z scores from a normal table), we know from the empirical rule that over 99% of the data should lie between -3 and 3 under a standard normal curve. However, seven of the ten scores shown on the printout are greater than 3 or less than -3. There are 100 data points total, so this indicates that at least 7% of the data is more than 3 standard deviations from the mean—which contradicts the empirical rule.

3. (a) To convert a score of 700 in 1967 to standard units, we take the Z score:

$$Z = \frac{x - \mu}{\sigma} = \frac{700 - 543}{110} \approx 1.43.$$

   To find the percentage of students who scored at least 700 in 1967, we must find the area to the right of 1.43 standard units under the normal curve. Using a Z table (p. A104 in the back of the book), we see that about 84.7% of students scored within 1.43 standard units of the mean. We can then calculate the area to the right of 1.43 under the normal curve as $\frac{100-84.7}{2}$. Therefore, 7.7% of students scored at least 700 in 1967.

   (b) Using the same formula, the Z score for a score of 700 in 1994 is:

$$Z = \frac{x - \mu}{\sigma} = \frac{700 - 499}{110} \approx 1.83.$$

   To find the percentage of students who scored at least 700 in 1994, we must find the area to the right of 1.83 standard units under the normal curve. Using a Z table (p. A104 in the back of the book), we see that about 93.3% of students scored within 1.83 standard units of the mean. We can then calculate the area to the right of 1.83 under the normal curve as $\frac{100-93.3}{2}$. Therefore, 3.4% of students scored at least 700 in 1994.

4. (a) As in #3, we will use Z scores to determine the percentage in each group that scored at least 700. The Z score for an SAT score of 700 among men is

$$Z = \frac{x - \mu}{\sigma} = \frac{700 - 538}{120} \approx 1.35.$$

   This means about 8.9% of men scored at least 700 in 2005.

(b) The Z score for 700 in the group of women is

$$Z = \frac{x-\mu}{\sigma} = \frac{700-504}{120} \approx 1.63.$$

This means about 5.1% of women scored at least 700 in 2005.

7. (a) We want to find the percentage of students who scored 400 or less, given that the average score was 550 and the SD was 100. The Z value for a score of 400 is

$$Z = \frac{x-\mu}{\sigma} = \frac{400-550}{100} = -1.5.$$

This means we want to find the percentage of values from the standard normal distribution that are at least 1.5 SDs below the mean. Since the normal distribution is symmetric, this will be equal to the percent that are at least 1.5 SDs above the mean. From looking at a Z table, we can see that this is about 6.7%.

(b) We want to find the SAT score (call it x) such that 75% of the scores are below x. First we must find the Z score such that 75% of a standard normal distribution is less than Z. Because 50% of a standard normal distribution is less than 0, we must find a Z score such that 25% of the area under the normal distribution curve is between 0 and Z. Because the normal distribution is symmetric, that means 50% of the area should be between $-Z$ and Z. Looking at the Z table (p. A104), we see that 50% of the area of the curve falls between Z and $-Z$ when Z is somewhere between 0.65 and 0.70. To convert this from standardized units into SAT scores, we want to solve for X in the Z formula:

$$Z = \frac{x-\mu}{\sigma} \qquad \Leftrightarrow \qquad x = Z \cdot \sigma + \mu = 0.675 \cdot 100 + 550 \approx 618.$$

Therefore, a student would need to score about 618 to be at the 75th percentile of the distribution.

8. (a) True.
   (b) False.
   (c) True.
   (d) True.
   (e) True.

(f) False.

For explanations, see the section on "Change of Scale" in Lecture 7 (posted on the course website).

9. (a) False. For example, consider these two lists: (1, 1, 1, 1, 1) and (1, 1, 1, 1, 1000). In the first list, the median and average are the same. In the second, the median is 1 and the average is 200.8.

   (b) False. For example, consider the list (1, 1000, 1000, 1000, 1000). The average is 800.2, and only one value is below the average.

   (c) False. Even with a large, representative sample, the histogram may not follow the normal curve. For example, family income in the U.S. is not normally distributed; it is skewed right. (see p. 88)

   (d) False. If the two lists of numbers are not both distributed normally and instead, for example, one has a left-skewed distribution, then we would not expect the same percentage of entries between 40 and 60 for both lists.