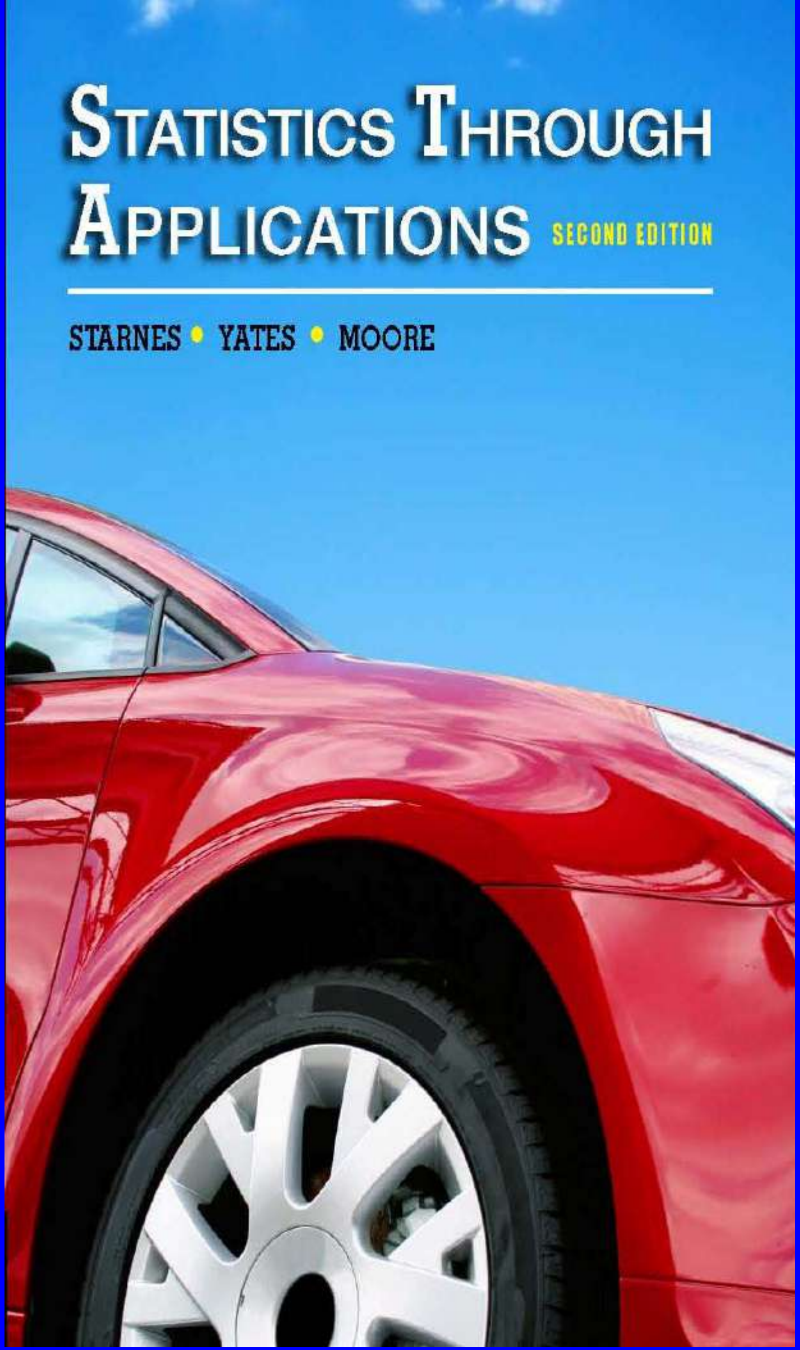
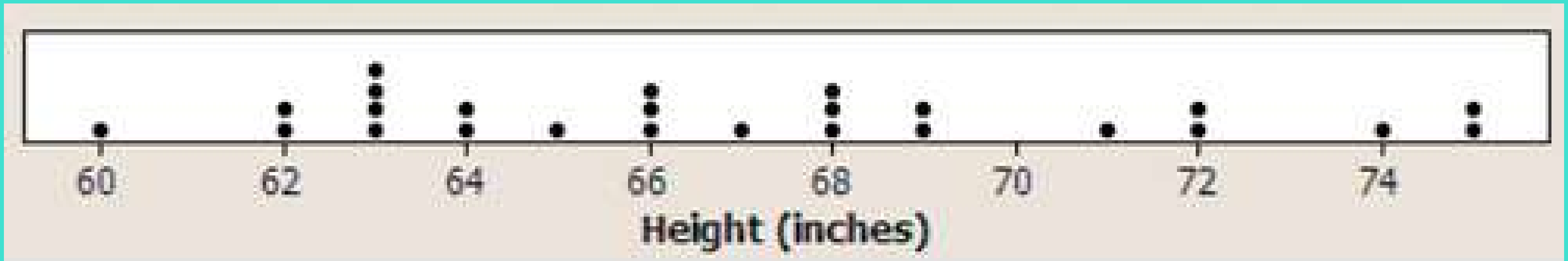


Chapter 3

Modeling Distributions of Data

Section 3.1
Measuring Location in a Distribution

Where do I stand?



Variable	n	Mean	Std. dev.	Min.	Q_1	M	Q_3	Max.
Height	25	67	4.29	60	63	66	69	75

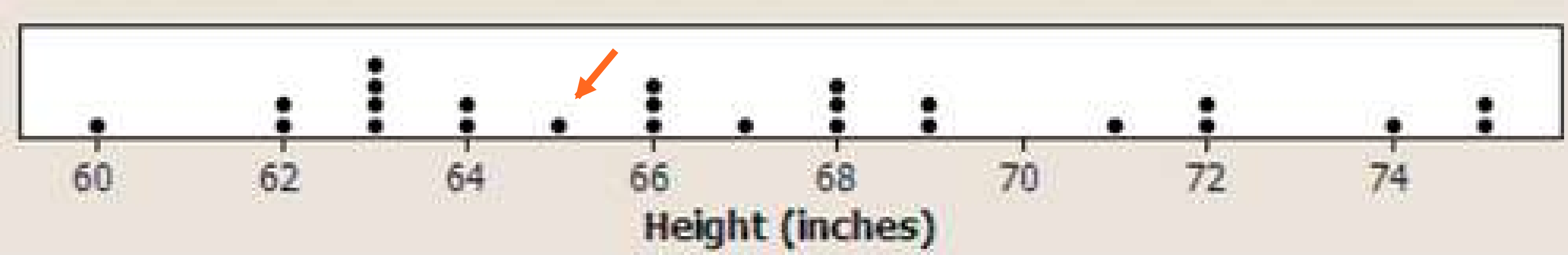
Figure 3.1
Dotplot and summary statistics for the heights of Mrs. Navard's statistics students.

Lynette, a student in the class, is 65 inches tall. Is she tall or short relative to her classmates?

Measuring location: percentiles

One way to describe Lynette's location within the distribution of heights is to tell what percent of students in the class are her height or shorter.

The pth percentile of a distribution is the value with p percent of the observations less than or equal to it.



Variable	n	Mean	Std. dev.	Min.	Q_1	M	Q_3	Max.
Height	25	67	4.29	60	63	66	69	75

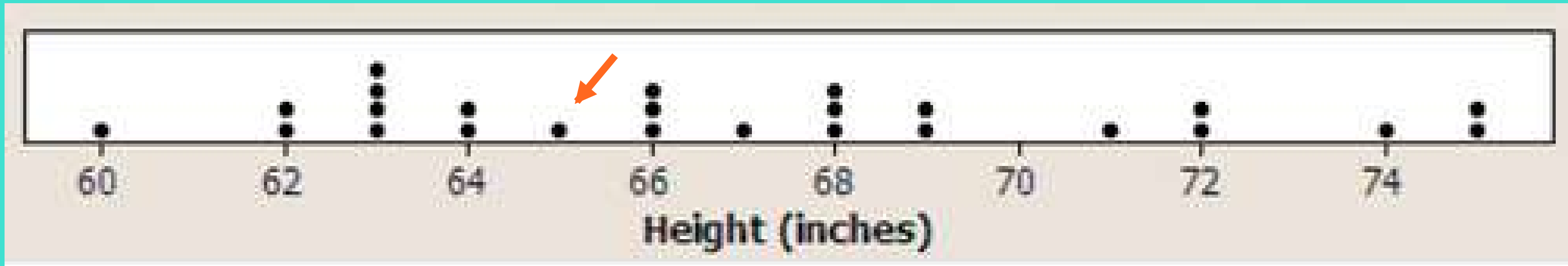
Lynnette's 65 inch height is 10th from the bottom.

Since 10 of 25 observations (40%) are at or below her height, Lynette is at the 40th percentile in the class's height distribution.

*** Note that some define the pth percentile of a distribution as the value with p percent of observations below it. That way you never go above the 99th percentile!

Measuring location: z-scores

Where does Lynette's height fall relative to the mean of this distribution?



Variable	n	Mean	Std. dev.	Min.	Q_1	M	Q_3	Max.
Height	25	67	4.29	60	63	66	69	75

The mean is 67 inches and the standard deviation is a little over 4 inches, her height is about one-half standard deviation below the mean. This conversion is known as standardizing.

Standardized values and z-scores

If x is an observation from a distribution that has known mean and standard deviation, the standardized value of x is

$$z = \frac{x - \text{mean}}{\text{standard deviation}}$$

A standardized value is often called a z-score.

For Lynette,

$$z = \frac{65 - 67}{4.29} = -0.47$$

Sophia scores 660 on the SAT Math test. The distribution of SAT scores has a mean of 500 and standard deviation of 100.

Jim takes the ACT Math test and scores 26. ACT scores have a mean of 18 and standard deviation of 6.

Assuming that both tests measure the same kind of ability, who did better?

Find the z-score from each to compare...

Density curves

Sometimes the overall pattern of a large number of observations is so regular that we can describe it by a smooth curve.

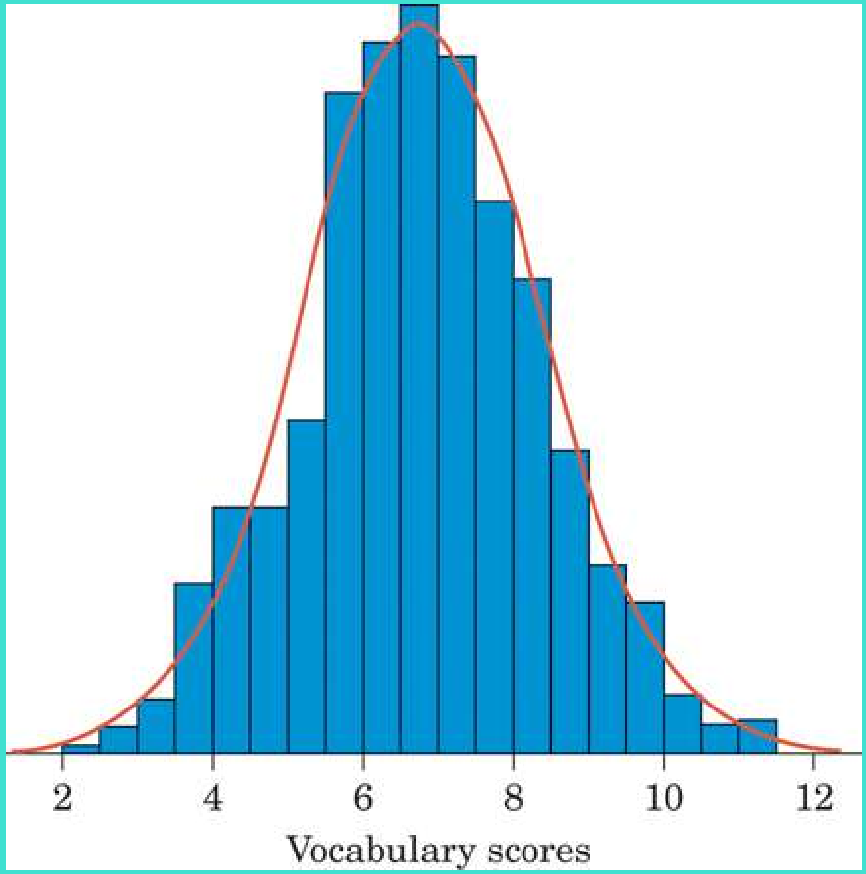
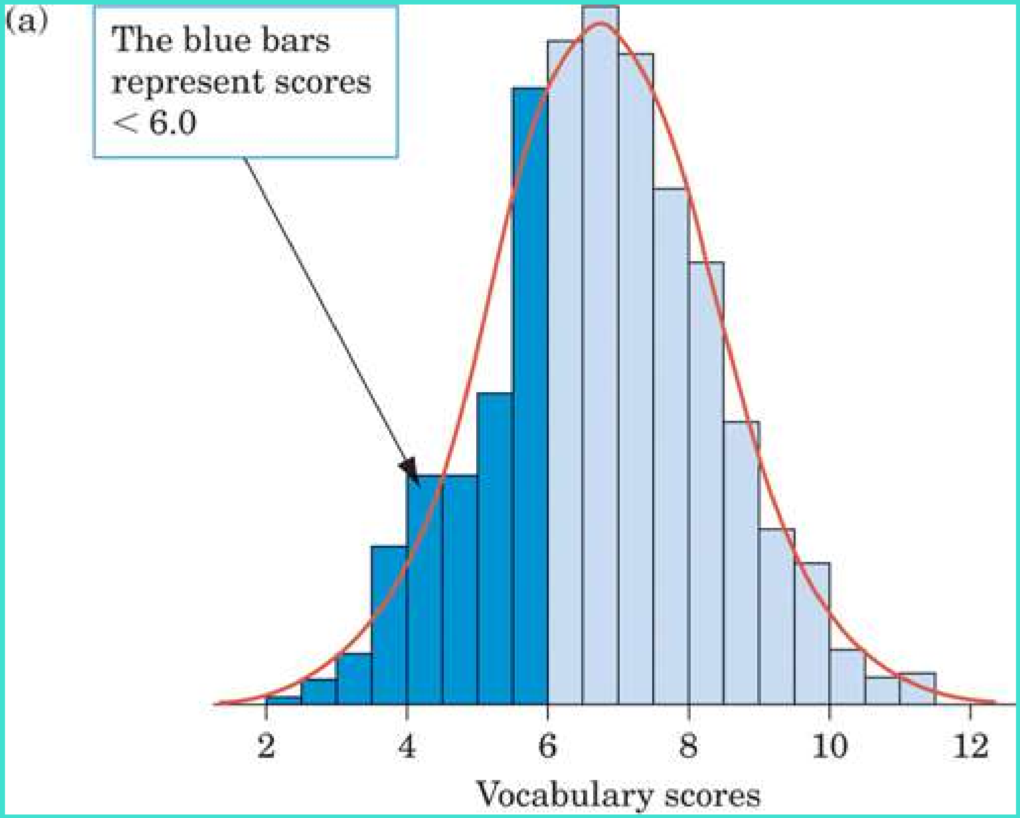


Figure 3.3
Histogram of the vocabulary scores of all seventh-grade students in Gary, Indiana.



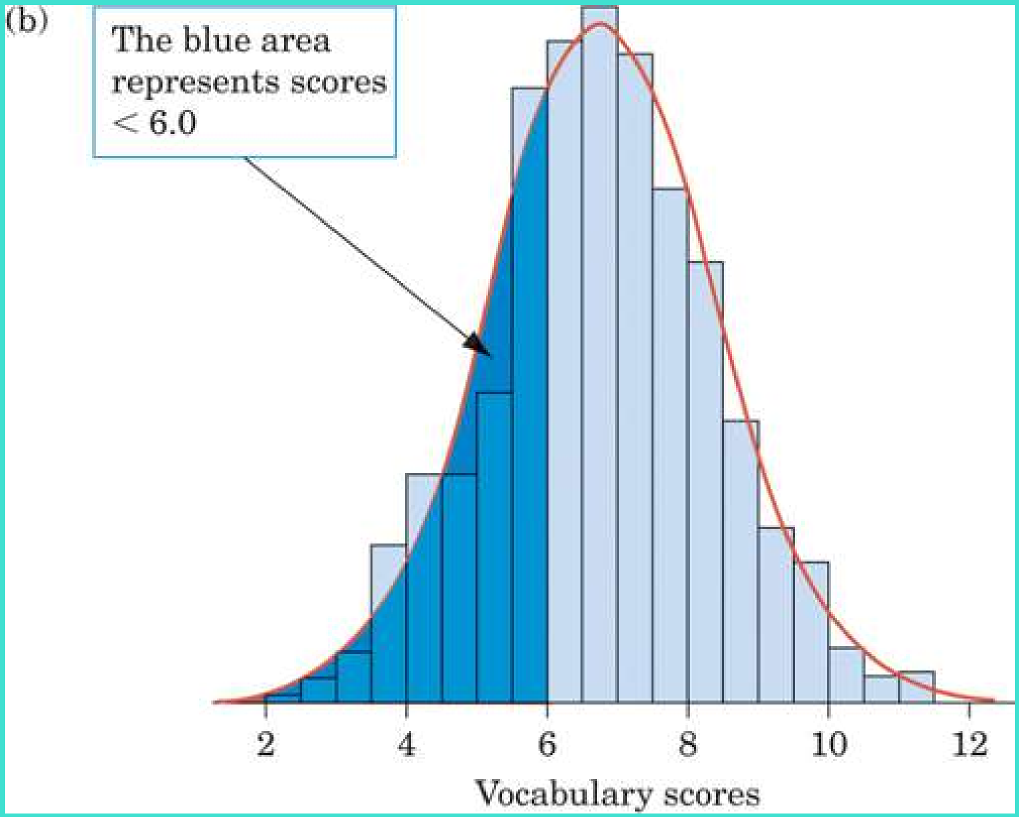
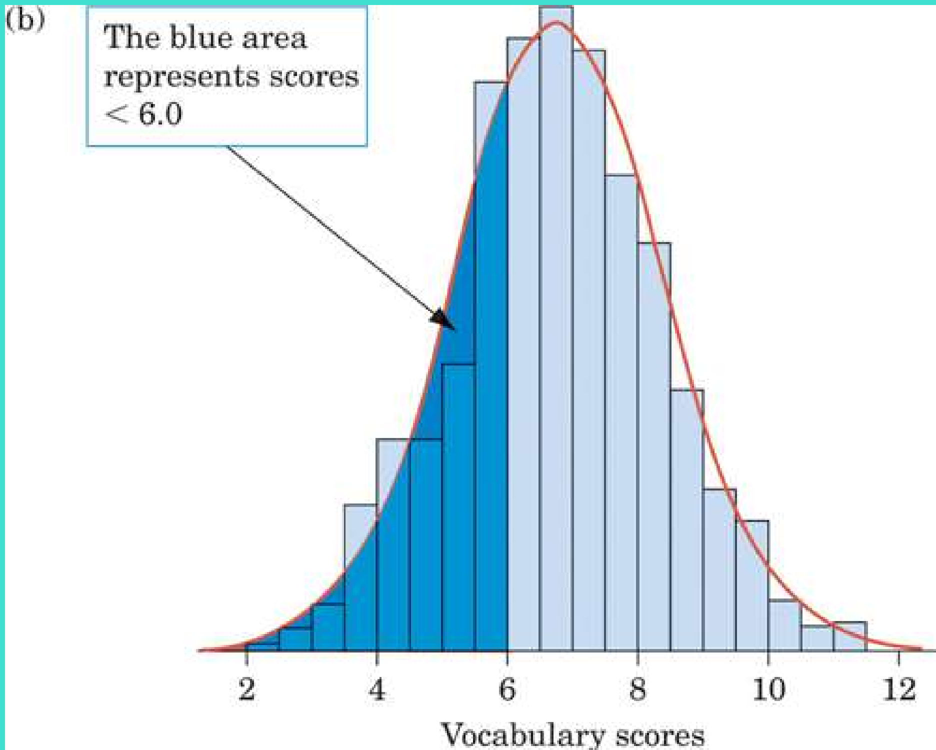


Figure 3.4

(a) The proportion of scores less than 6.0 from the histogram is 0.303.

(b) The proportion of scores less than 6.0 from the density curve is 0.293.



The total area under the curve is exactly 1.

The curve is called a density curve.

Areas under the density curve are quite good approximations of areas given by the histogram.

Normal Curve

The Normal curve has a distinctive, symmetric, single-peaked bell shape.

Normal curves have special properties

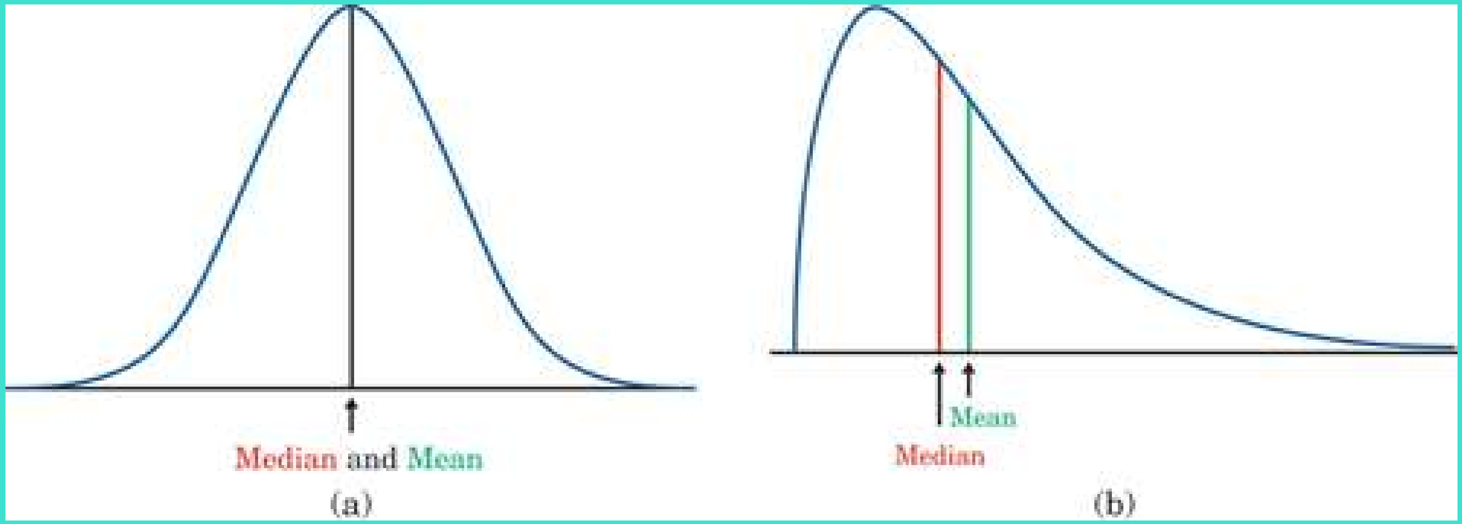
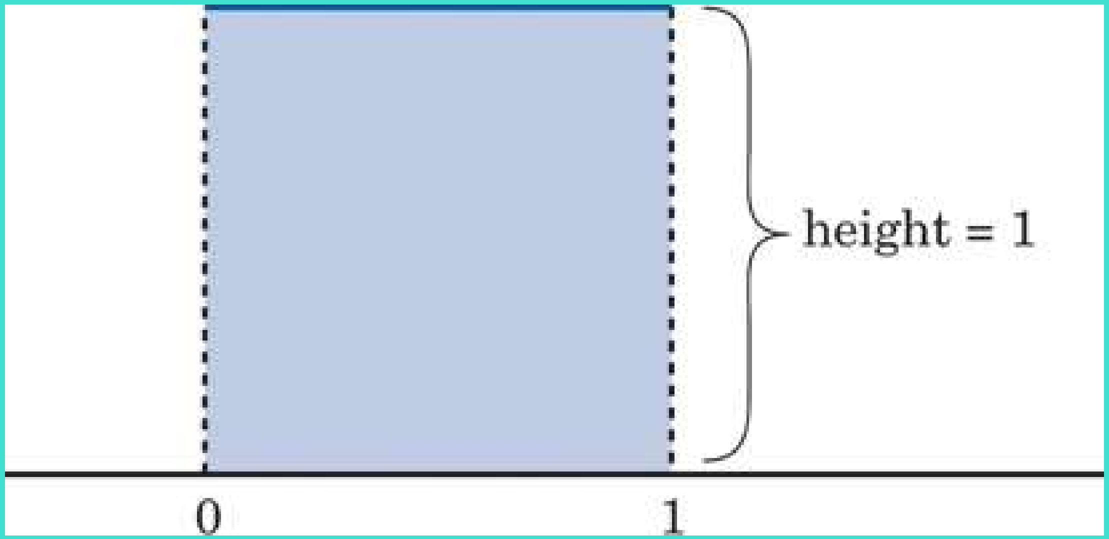


Figure 3.5

The median and mean for two density curves: a symmetric Normal curve and a curve that is skewed to the right.

The median of the Normal curve is the equal-areas point.
The mean of the Normal curve is the balance point.
The mean and median are the same for a symmetric curve.
The mean of a skewed curve is pulled away from the median in the direction of the long tail.



This figure shows the density curve for a uniform distribution. This curve has height 1 over the interval from 0 to 1 and is zero outside that range.

Practice 3-1

Practice 3.1.xps

Homework

Exercises Page 26-28

1.31 - 1.36

Attachments

Practice 3.1.xps