

Here are the scores on the Survey of Study Habits and Attitudes (SSHA) for 18 first-year college women:

154	109	137	115	152	140	154	178	101
103	126	126	137	165	165	129	200	148

- 1.44 The SSHA data for women given in the previous exercise contain one high outlier. Calculate the mean  $\bar{x}$  and the median  $M$  for these data with and without the outlier. How does removing the outlier affect  $\bar{x}$ ? How does it affect  $M$ ? Your results illustrate the greater resistance of the median.
- 1.48 The NASDAQ Composite Index describes the average price of common stock traded over the counter, that is, not on one of the stock exchanges. In 1991, the mean capitalization of the companies in the NASDAQ index was \$80 million and the median capitalization was \$20 million. (A company's capitalization is the total market value of its stock.) Explain why the mean capitalization is much higher than the median.
- 1.64 Figure 1.12 (page 28) shows the distribution of the mean SAT verbal scores in the 50 states and the District of Columbia for 1996. In 1995, the SATs were "recentered" to locate the mean scores for individuals close to the center of the 200 to 800 range of possible scores. To see the effect of recentering, we enter the state results for the mathematics and verbal SATs for both 1990 and 1996 into a statistical software package. Here are the results. (This output is from the Minitab statistical software; other software produces similar results.)

SATV-1990							
N	MEAN	MEDIAN	STDEV	MIN	MAX	Q1	Q3
51	448.16	443.00	30.82	397.00	511.00	422.00	476.00
SATM-1990							
N	MEAN	MEDIAN	STDEV	MIN	MAX	Q1	Q3
51	497.39	490.00	34.57	437.00	577.00	470.00	523.00
SATV-1996							
N	MEAN	MEDIAN	STDEV	MIN	MAX	Q1	Q3
51	531.90	525.00	33.76	480.00	596.00	501.00	565.00
SATM-1996							
N	MEAN	MEDIAN	STDEV	MIN	MAX	Q1	Q3
51	529.30	521.00	34.83	473.00	600.00	500.00	557.00

- (a) Use the output to make side-by-side boxplots of the four distributions. What was the effect of recentering on the distribution of verbal SAT

scores in the states? How did the distributions of state verbal and mathematics scores compare before recentering? How do they compare after recentering?

- (b) Look again at Figure 1.12. What important feature of the distribution of 1996 verbal scores does the boxplot fail to display?

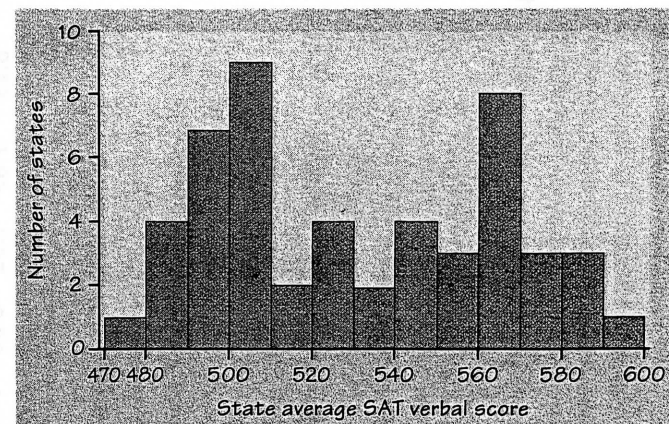


FIGURE 1.12 Histogram of the mean scores on the verbal part of the SAT for the 50 states and the District of Columbia, for Exercise 1.17.

- 1.82 Using either Table A or your calculator or software, find the relative frequency of each of the following events in a standard normal distribution. In each case, sketch a standard normal curve with the area representing the relative frequency shaded.
- (a)  $Z \leq -2.25$
  - (b)  $Z \geq -2.25$
  - (c)  $Z > 1.77$
  - (d)  $-2.25 < Z < 1.77$
- 1.84 The variable  $Z$  has a standard normal distribution.
- (a) Find the number  $z$  such that the event  $Z < z$  has relative frequency 0.8
  - (b) Find the number  $z$  such that the event  $Z > z$  has relative frequency 0.35.

**1.86** The rate of return on stock indexes (which combine many individual stocks) is approximately normal. Since 1945, the Standard & Poor's 500 index has had a mean yearly return of about 12%, with a standard deviation of about 16.5%. Take this normal distribution to be the distribution of yearly returns over a long period.

- (a) In what range do the middle 95% of all yearly returns lie?
- (b) The market is down for the year if the return on the index is less than zero. In what percent of years is the market down?
- (c) In what percent of years does the index gain 25% or more?

**1.88** The length of human pregnancies from conception to birth varies according to a distribution that is approximately normal with mean 266 days and standard deviation 16 days.

- (a) What percent of pregnancies last less than 240 days (that's about 8 months)?
- (b) What percent of pregnancies last between 240 and 270 days (roughly between 8 months and 9 months)?
- (c) How long do the longest 20% of pregnancies last?

The scores of a reference population on the Wechsler Intelligence Scale for Children (WISC) are normally distributed with  $\mu = 100$  and  $\sigma = 15$ .

**1.94** The distribution of scores on the WISC is described in the previous exercise. What score will place a child in the top 5% of the population? In the top 1%?

**1.110** Deborah is a student at a midwestern college who lives off campus. She records the time she takes to drive to school each morning during the fall term. Here are the times (in minutes) for 42 consecutive weekdays, with the dates in order along the rows:

8.25	7.83	8.30	8.42	8.50	8.67	8.17	9.00	9.00	8.17	7.92
9.00	8.50	9.00	7.75	7.92	8.00	8.08	8.42	8.75	8.08	9.75
8.33	7.83	7.92	8.58	7.83	8.42	7.75	7.42	6.75	7.42	8.50
8.67	10.17	8.75	8.58	8.67	9.17	9.08	8.83	8.67		

- (a) Make a graph of these drive times. Is the distribution roughly symmetric, clearly skewed, or neither? Are there any clear outliers?
- (b) The data show three unusual situations: the day after Thanksgiving (no traffic on campus); a delay due to an accident; and a day with icy roads. Identify and remove these three observations, and give a numerical summary of the remaining 39 drive times.
- (c) Are these data reasonably close to having a normal distribution when the unusual days are removed? How do you know?
- (d) Make a time plot of the drive times. Do the data show any trend? For example, do drive times increase over time because the weather gets worse later in the term?

**1.114** The Chapin Social Insight Test evaluates how accurately the subject appraises other people. In the reference population used to develop the test, scores are approximately normally distributed with mean 25 and standard deviation 5. The range of possible scores is 0 to 41.

- (a) What proportion of the population has scores below 20 on the Chapin test?
- (b) What proportion has scores below 10?
- (c) How high a score must you have in order to be in the top quarter of the population in social insight?