

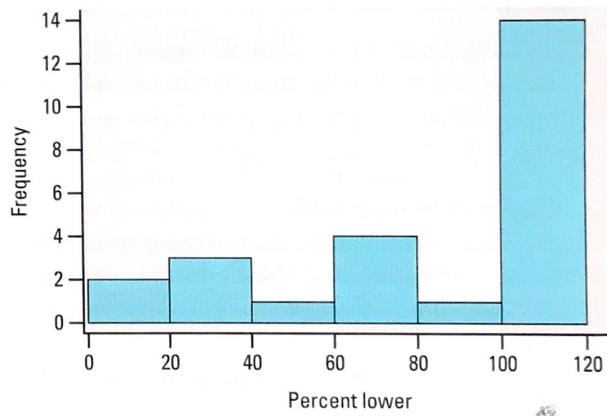
SECTION 9.3

Exercises

63. **Attitudes** The Survey of Study Habits and Attitudes (SSHA) is a psychological test that measures students' attitudes toward school and study habits. Scores range from 0 to 200. The mean score for U.S. college students is about 115. A teacher suspects that older students have better attitudes toward school. She gives the SSHA to an SRS of 45 of the over 1000 students at her college who are at least 30 years of age. Check the conditions for carrying out a significance test of the teacher's suspicion.
64. **Anemia** Hemoglobin is a protein in red blood cells that carries oxygen from the lungs to body tissues. People with fewer than 12 grams of hemoglobin per deciliter of blood (g/dl) are anemic. A public health official in Jordan suspects that Jordanian children are at risk of anemia. He measures a random sample of 50 children. Check the conditions for carrying out a significance test of the official's suspicion.
65. **Paying high prices?** A retailer entered into an exclusive agreement with a supplier who guaranteed to provide all products at competitive prices. The retailer eventually began to purchase supplies from other vendors who offered better prices. The original supplier filed a lawsuit claiming violation of the agreement. In defense, the retailer had an audit performed on a random sample of 25 invoices. For each audited invoice, all purchases made from other suppliers were

pg 566

examined and compared with those offered by the original supplier. The percent of purchases on each invoice for which an alternative supplier offered a lower price than the original supplier was recorded.²⁶ For example, a data value of 38 means that the price would be lower with a different supplier for 38% of the items on the invoice. A histogram and some computer output for these data are shown below. Explain why we should not carry out a one-sample *t* test in this setting.



Summary statistics

Column	n	Mean	Std. Dev.	Std. Err.	Median	Min	Max	Q1	Q3
pctlower	25	77.76	32.6768	6.5353603	100	0	100	68	100

- 66. Ancient air** The composition of the earth's atmosphere may have changed over time. To try to discover the nature of the atmosphere long ago, we can examine the gas in bubbles inside ancient amber. Amber is tree resin that has hardened and been trapped in rocks. The gas in bubbles within amber should be a sample of the atmosphere at the time the amber was formed. Measurements on 9 specimens of amber from the late Cretaceous era (75 to 95 million years ago) give these percents of nitrogen:²⁷
- 63.4 65.0 64.4 63.3 54.8 64.5 60.8 49.1 51.0
- Explain why we should not carry out a one-sample *t* test in this setting.
- 67. Attitudes** In the study of older students' attitudes from Exercise 63, the sample mean SSHA score was 125.7 and the sample standard deviation was 29.8.
- Calculate the test statistic.
 - Find the *P*-value using Table B. Then obtain a more precise *P*-value from your calculator.
- 68. Anemia** For the study of Jordanian children in Exercise 64, the sample mean hemoglobin level was 11.3 mg/dl and the sample standard deviation was 1.6 mg/dl.
- Calculate the test statistic.
 - Find the *P*-value using Table B. Then obtain a more precise *P*-value from your calculator.
- 69. One-sided test** Suppose you carry out a significance test of $H_0: \mu = 5$ versus $H_a: \mu > 5$ based on a sample of size $n = 20$ and obtain $t = 1.81$.
- Find the *P*-value for this test using (i) Table B and (ii) your calculator. What conclusion would you draw at the 5% significance level? At the 1% significance level?
 - Redo part (a) using an alternative hypothesis of $H_a: \mu \neq 5$.
- 70. Two-sided test** The one-sample *t* statistic from a sample of $n = 25$ observations for the two-sided test of $H_0: \mu = 64$
- $$H_a: \mu \neq 64$$
- has the value $t = -1.12$.
- Find the *P*-value for this test using (i) Table B and (ii) your calculator. What conclusion would you draw at the 5% significance level? At the 1% significance level?
 - Redo part (a) using an alternative hypothesis of $H_a: \mu < 64$.
- 71. Sweetening colas** Cola makers test new recipes for loss of sweetness during storage. Trained tasters rate the sweetness before and after storage. From experience, the population distribution of sweetness losses will be close to Normal. Here are the sweetness losses (sweetness before storage minus sweetness after storage) found by tasters from a random sample of 10 batches of a new cola recipe:
- 2.0 0.4 0.7 2.0 -0.4 2.2 -1.3 1.2 1.1 2.3
- Are these data good evidence that the cola lost sweetness? Carry out a test to help you answer this question.
- 72. Heat through the glass** How well materials conduct heat matters when designing houses, for example. Conductivity is measured in terms of watts of heat power transmitted per square meter of surface per degree Celsius of temperature difference on the two sides of the material. In these units, glass has conductivity about 1. The National Institute of Standards and Technology provides exact data on properties of materials. Here are measurements of the heat conductivity of 11 randomly selected pieces of a particular type of glass:²⁸
- 1.11 1.07 1.11 1.07 1.12 1.08 1.08 1.18 1.18 1.18 1.12
- Is there convincing evidence that the conductivity of this type of glass is greater than 1? Carry out a test to help you answer this question.
- 73. Healthy bones** The recommended daily allowance (RDA) of calcium for women between the ages of 18 and 24 years is 1200 milligrams (mg). Researchers who were involved in a large-scale study of women's bone health suspected that their participants had significantly lower calcium intakes than the RDA. To test this suspicion, the researchers measured the daily calcium intake of a random sample of 36 women from the study who fell in the desired age range. The Minitab output below displays descriptive statistics for these data, along with the results of a significance test.
- | Descriptive Statistics: Calcium intake (mg) | | | | | | | | | |
|---|----|-------|---------|-------|-------|-------|-------|--------|---------|
| Variable | N | Mean | SE Mean | StDev | Min | Q1 | Med | Q3 | Maximum |
| Calcium | 36 | 856.2 | 51.1 | 306.7 | 374.0 | 632.3 | 805.0 | 1090.5 | 1425.0 |
- | One-Sample T: Calcium intake (mg) | | | | | | | | |
|-----------------------------------|----|-------|-------|---------|-------|------|-------|---|
| Test of | mu | = | 1200 | vs | < | 1200 | T | P |
| Variable | N | Mean | StDev | SE Mean | | | | |
| Calcium | 36 | 856.2 | 306.7 | 51.1 | -6.73 | | 0.000 | |
- Determine whether there are any outliers. Show your work.
 - Interpret the *P*-value in context.
 - Do these data give convincing evidence to support the researchers' suspicion? Carry out a test to help you answer this question.

- 74. Taking stock** An investor with a stock portfolio worth several hundred thousand dollars sued his broker due to the low returns he got from the portfolio at a time when the stock market did well overall. The investor's lawyer wants to compare the broker's performance against the market as a whole. He collects data on the broker's returns for a random sample of 36 weeks. Over the 10-year period that the broker has managed portfolios, stocks in the Standard & Poor's 500 index gained an average of 0.95% per month. The Minitab output below displays descriptive statistics for these data, along with the results of a significance test.

Descriptive Statistics: Return (percent)									
Variable	N	Mean	SE Mean	StDev	Min	Q1	Median	Q3	Max
Return(%)	36	-1.441	0.802	4.810	-10.270	-3.418	-1.425	1.543	7.340

One-Sample T: Return (percent)						
Test of $\mu = 0.95$ vs < 0.95						
Variable	N	Mean	StDev	SE Mean	T	P
Return (percent)	36	-1.441	4.810	0.802	-2.98	0.003

- (a) Determine whether there are any outliers. Show your work.
 (b) Interpret the P-value in context.
 (c) Do these data give convincing evidence to support the lawyer's case? Carry out a test to help you answer this question.

- 75. Growing tomatoes** An agricultural field trial compares the yield of two varieties of tomatoes for commercial use. Researchers randomly select 10 Variety A and 10 Variety B tomato plants. Then the researchers divide in half each of 10 small plots of land in different locations. For each plot, a coin toss determines which half of the plot gets a Variety A plant; a Variety B plant goes in the other half. After harvest, they compare the yield in pounds for the plants at each location. The 10 differences (Variety A – Variety B) give $\bar{x} = 0.34$ and $s_x = 0.83$. A graph of the differences looks roughly symmetric and single-peaked with no outliers. Is there convincing evidence that Variety A has the higher mean yield? Perform a significance test using $\alpha = 0.05$ to answer the question.

- 76. Study more!** A student group claims that first-year students at a university study 2.5 hours per night during the school week. A skeptic suspects that they study less than that on average. He takes a random sample of 30 first-year students and finds that $\bar{x} = 137$ minutes and $s_x = 45$ minutes. A graph of the data shows no outliers but some skewness. Carry out an

appropriate significance test at the 5% significance level. What conclusion do you draw?

- 77. The power of tomatoes** The researchers who carried out the experiment in Exercise 75 suspect that the large P-value (0.114) is due to low power.
- (a) Describe a Type I and a Type II error in this setting. Which type of error could you have made in Exercise 75? Why?
 (b) Explain two ways that the researchers could have increased the power of the test to detect $\mu = 0.5$.

- 78. Study more!** The significance test in Exercise 76 yields a P-value of 0.0622.
- (a) Describe a Type I and a Type II error in this setting. Which type of error could you have made in Exercise 76? Why?
 (b) Which of the following changes would give the test a higher power to detect $\mu = 120$ minutes: using $\alpha = 0.01$ or $\alpha = 0.10$? Explain.

- 79. Pressing pills** A drug manufacturer forms tablets by compressing a granular material that contains the active ingredient and various fillers. The hardness of a sample from each batch of tablets produced is measured to control the compression process. The target value for the hardness is $\mu = 11.5$. The hardness data for a random sample of 20 tablets are

11.627	11.613	11.493	11.602	11.360
11.374	11.592	11.458	11.552	11.463
11.383	11.715	11.485	11.509	11.429
11.477	11.570	11.623	11.472	11.531

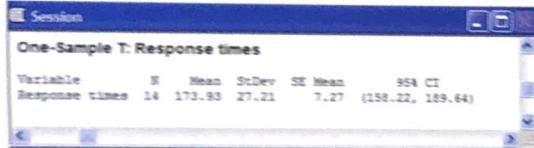
Is there significant evidence at the 5% level that the mean hardness of the tablets differs from the target value? Carry out an appropriate test to support your answer.

- 80. Filling cola bottles** Bottles of a popular cola are supposed to contain 300 milliliters (ml) of cola. There is some variation from bottle to bottle because the filling machinery is not perfectly precise. From experience, the distribution of the contents is approximately Normal. An inspector measures the contents of six randomly selected bottles from a single day's production. The results are

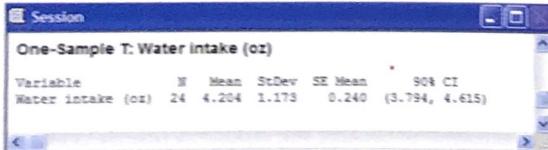
299.4 297.7 301.0 298.9 300.2 297.0

Do these data provide convincing evidence that the mean amount of cola in all the bottles filled that day differs from the target value of 300 ml? Carry out an appropriate test to support your answer.

- 81. Pressing pills** Refer to Exercise 79. Construct and interpret a 95% confidence interval for the population mean μ . What additional information does the confidence interval provide?
- 82. Filling cola bottles** Refer to Exercise 80. Construct and interpret a 95% confidence interval for the population mean μ . What additional information does the confidence interval provide?
- 83. Fast connection?** How long does it take for a chunk of information to travel from one server to another and back on the Internet? According to the site internettrafficreport.com, a typical response time is 200 milliseconds (about one-fifth of a second). Researchers collected data on response times of a random sample of 14 servers in Europe. A graph of the data reveals no strong skewness or outliers. The figure below displays Minitab output for a one-sample t interval for the population mean. Is there convincing evidence at the 5% significance level that the site's claim is incorrect? Use the confidence interval to justify your answer.



- 84. Water!** A blogger claims that U.S. adults drink an average of five 8-ounce glasses of water per day. Skeptical researchers ask a random sample of 24 U.S. adults about their daily water intake. A graph of the data shows a roughly symmetric shape with no outliers. The figure below displays Minitab output for a one-sample t interval for the population mean. Is there convincing evidence at the 10% significance level that the blogger's claim is incorrect? Use the confidence interval to justify your answer.



- 85. Tests and CIs** The P -value for a two-sided test of the null hypothesis $H_0: \mu = 10$ is 0.06.
- Does the 95% confidence interval for μ include 10? Why or why not?
 - Does the 90% confidence interval for μ include 10? Why or why not?
- 86. Tests and CIs** The P -value for a one-sided test of the null hypothesis $H_0: \mu = 15$ is 0.03.
- Does the 99% confidence interval for μ include 15? Why or why not?
 - Does the 95% confidence interval for μ include 15? Why or why not?

Exercises 87 and 88 refer to the following setting. An understanding of cockroach biology may lead to an effective control strategy for these annoying insects. Researchers studying the absorption of sugar by insects feed a random sample of cockroaches a diet containing measured amounts of sugar. After 10 hours, the cockroaches are killed and the concentration of the sugar in various body parts is determined by a chemical analysis. The paper that reports the research states that a 95% confidence interval for the mean amount (in milligrams) of sugar in the hindguts of cockroaches is 4.2 ± 2.3 .²⁹

- 87. Cockroaches** Does this paper give convincing evidence that the mean amount of sugar in the hindguts under these conditions is not equal to 7 mg? Justify your answer.
- 88. Cockroaches** Would the hypothesis that $\mu = 5$ mg be rejected at the 5% level in favor of a two-sided alternative? Justify your answer.
- 89. Right versus left** The design of controls and instruments affects how easily people can use them. A student project investigated this effect by asking 25 right-handed students to turn a knob (with their right hands) that moved an indicator. There were two identical instruments, one with a right-hand thread (the knob turns clockwise) and the other with a left-hand thread (the knob must be turned counterclockwise). Each of the 25 students used both instruments in a random order. The following table gives the times in seconds each subject took to move the indicator a fixed distance.³⁰

Subject	Right thread	Left thread
1	113	137
2	105	105
3	130	133
4	101	108
5	138	115
6	118	170
7	87	103
8	116	145
9	75	78
10	96	107
11	122	84
12	103	148
13	116	147
14	107	87
15	118	166
16	103	146
17	111	123
18	104	135
19	111	112

Subject	Right thread	Left thread
20	89	93
21	78	76
22	100	116
23	89	78
24	85	101
25	88	123

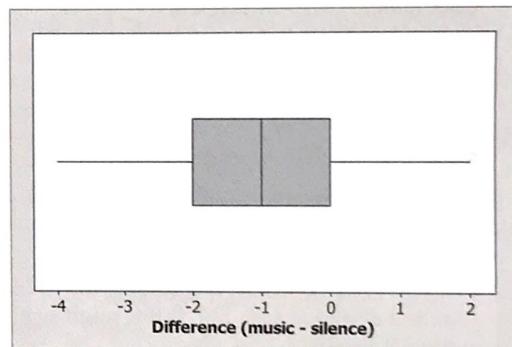
- (a) Explain why it was important to randomly assign the order in which each subject used the two knobs.
- (b) The project designers hoped to show that right-handed people find right-hand threads easier to use. Carry out a significance test at the 5% significance level to investigate this claim.
90. **Floral scents and learning** We hear that listening to Mozart improves students' performance on tests. Maybe pleasant odors have a similar effect. To test this idea, 21 subjects worked two different but roughly equivalent paper-and-pencil mazes while wearing a mask. The mask was either unscented or carried a floral scent. Each subject used both masks, in a random order. The table below gives the subjects' times with both masks.³¹

Subject	Unscented	Scented
1	30.60	37.97
2	48.43	51.57
3	60.77	56.67
4	36.07	40.47
5	68.47	49.00
6	32.43	43.23
7	43.70	44.57
8	37.10	28.40
9	31.17	28.23
10	51.23	68.47
11	65.40	51.10
12	58.93	83.50
13	54.47	38.30
14	43.53	51.37
15	37.93	29.33
16	43.50	54.27
17	87.70	62.73
18	53.53	58.00
19	64.30	52.40
20	47.37	53.63
21	53.67	47.00

- (a) Explain why it was important to randomly assign the order in which each subject used the two masks.

(b) Do these data provide convincing evidence that the floral scent improved performance? Carry out an appropriate test to support your answer.

91. **Music and memory** Does listening to music while studying help or hinder students' learning? Two AP Statistics students designed an experiment to find out. They selected a random sample of 30 students from their medium-sized high school to participate. Each subject was given 10 minutes to memorize two different lists of 20 words, once while listening to music and once in silence. The order of the two word lists was determined at random; so was the order of the treatments. A boxplot of the differences in the number of words recalled ($\text{music} - \text{silent}$) is shown below, along with some Minitab output from a one-sample t test. Perform a complete analysis of the students' data. Include a confidence interval.



One-Sample T: Difference (music - silence)						
Test of mu = 0 vs not = 0		N	Mean	StDev	SE Mean	T
Difference	30	-0.933	1.701	0.310	-3.01	0.005

92. **Darwin's plants** Charles Darwin, author of *On the Origin of Species* (1859), designed an experiment to compare the effects of cross-fertilization and self-fertilization on the size of plants. He planted pairs of very similar seedling plants, one self-fertilized and one cross-fertilized, in each of 15 pots at the same time. After a period of time, Darwin measured the heights (in inches) of all the plants. Here are the data:

Pair	Cross	Self	Pair	Cross	Self
1	23.5	17.4	9	18.3	16.5
2	12.0	20.4	10	21.6	18.0
3	21.0	20.0	11	23.3	16.3
4	22.0	20.0	12	21.0	18.0
5	19.1	18.4	13	22.1	12.8
6	21.5	18.6	14	23.0	15.5
7	22.1	18.6	15	12.0	18.0
8	20.4	15.3			

- (a) Explain why it is not appropriate to perform a paired t test in this setting.
- (b) A hasty student generates the Minitab output shown below. What conclusion should he draw at the $\alpha = 0.05$ significance level? Explain.

One-Sample T: Diff (cross - self)						
Test of mu = 0 vs not = 0						
Variable	N	Mean	StDev	SE Mean	T	P
Diff	15	2.61	4.71	1.22	2.14	0.050

93. **Is it significant?** For students without special preparation, SAT Math scores in recent years have varied Normally with mean $\mu = 518$. One hundred students go through a rigorous training program designed to raise their SAT Math scores by improving their mathematics skills. Use your calculator to carry out a test of

$$H_0: \mu = 518$$

$$H_a: \mu > 518$$

in each of the following situations.

- (a) The students' scores have mean $\bar{x} = 536.7$ and standard deviation $s_x = 114$. Is this result significant at the 5% level?
- (b) The students' scores have mean $\bar{x} = 537.0$ and standard deviation $s_x = 114$. Is this result significant at the 5% level?
- (c) When looked at together, what is the intended lesson of (a) and (b)?
94. **Significance and sample size** A study with 5000 subjects reported a result that was statistically significant at the 5% level. Explain why this result might not be particularly large or important.

95. **Sampling shoppers** A marketing consultant observes 50 consecutive shoppers at a supermarket, recording how much each shopper spends in the store. Explain why it would not be wise to use these data to carry out a significance test about the mean amount spent by all shoppers at this supermarket.

96. **Ages of presidents** Joe is writing a report on the backgrounds of American presidents. He looks up the ages of all the presidents when they entered office. Because Joe took a statistics course, he uses these numbers to perform a significance test about the mean age of all U.S. presidents. Explain why this makes no sense.

97. **Do you have ESP?** A researcher looking for evidence of extrasensory perception (ESP) tests 500 subjects. Four of these subjects do significantly better ($P < 0.01$) than random guessing.

- (a) Is it proper to conclude that these four people have ESP? Explain your answer.

- (b) What should the researcher now do to test whether any of these four subjects have ESP?

98. **What is significance good for?** Which of the following questions does a significance test answer? Justify your answer.

- (a) Is the sample or experiment properly designed?
 (b) Is the observed effect due to chance?
 (c) Is the observed effect important?

Multiple choice: Select the best answer for Exercises 99 to 104.

99. The reason we use t procedures instead of z procedures when carrying out a test about a population mean is that

- (a) z can be used only for large samples.
 (b) z requires that you know the population standard deviation σ .
 (c) z requires you to regard your data as an SRS from the population.
 (d) z applies only if the population distribution is perfectly Normal.
 (e) z can be used only for confidence intervals.

100. You are testing $H_0: \mu = 10$ against $H_a: \mu < 10$ based on an SRS of 20 observations from a Normal population. The t statistic is $t = -2.25$. The P -value

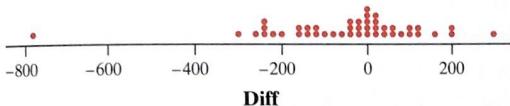
- (a) falls between 0.01 and 0.02.
 (b) falls between 0.02 and 0.04.
 (c) falls between 0.04 and 0.05.
 (d) falls between 0.05 and 0.25.
 (e) is greater than 0.25.

101. You are testing $H_0: \mu = 10$ against $H_a: \mu \neq 10$ based on an SRS of 15 observations from a Normal population. What values of the t statistic are statistically significant at the $\alpha = 0.005$ level?

- (a) $t > 3.326$ (d) $t < -3.326$ or $t > 3.326$
 (b) $t > 3.286$ (e) $t < -3.286$ or $t > 3.286$
 (c) $t > 2.977$

102. After checking that conditions are met, you perform a significance test of $H_0: \mu = 1$ versus $H_a: \mu \neq 1$. You obtain a P -value of 0.022. Which of the following is true?

- (a) A 95% confidence interval for μ will include the value 1.
 (b) A 95% confidence interval for μ will include the value 0.
 (c) A 99% confidence interval for μ will include the value 1.

- (d) A 99% confidence interval for μ will include the value 0.
 (e) None of these is necessarily true.
103. Does Friday the 13th have an effect on people's behavior? Researchers collected data on the numbers of shoppers at a sample of 45 different grocery stores on Friday the 6th and Friday the 13th in the same month. The dotplot and computer output below summarize the data on the difference in the number of shoppers at each store on these two days (subtracting in the order 6th minus 13th).³²
- 
- | N | Mean | Median | TrMean | StDev | SEMean | Min | Max | Q1 | Q3 |
|----|-------|--------|--------|-------|--------|--------|-------|--------|------|
| 45 | -46.5 | -11.0 | -37.4 | 178.0 | 26.1 | -774.0 | 302.0 | -141.0 | 53.5 |
- Researchers would like to carry out a test of $H_0: \mu_d = 0$ versus $H_a: \mu_d \neq 0$, where μ_d is the true mean difference in the number of grocery shoppers on these two days. Which of the following conditions for performing a paired t test is *not* met?
- I. Random II. Normal III. Independent
 (a) I only (c) III only (e) I, II, and III
 (b) II only (d) I and II only
104. The most important condition for sound conclusions from statistical inference is that
 (a) the data come from a well-designed random sample or randomized experiment.
 (b) the population distribution be exactly Normal.
 (c) the data contain no outliers.
 (d) the sample size be no more than 10% of the population size.
 (e) the sample size be at least 30.
105. Is your food safe? (8.1) "Do you feel confident or not confident that the food available at most grocery stores is safe to eat?" When a Gallup Poll asked this question, 87% of the sample said they were confident.³³ Gallup announced the poll's margin of error for 95% confidence as ± 3 percentage points. Which of the following sources of error are included in this margin of error? Explain.
- (a) Gallup dialed landline telephone numbers at random and so missed all people without landline phones, including people whose only phone is a cell phone.
 (b) Some people whose numbers were chosen never answered the phone in several calls or answered but refused to participate in the poll.
 (c) There is chance variation in the random selection of telephone numbers.
106. Spinning for apples (6.3 or 7.3) In the "Ask Marilyn" column of *Parade* magazine, a reader posed this question: "Say that a slot machine has five wheels, and each wheel has five symbols: an apple, a grape, a peach, a pear, and a plum. I pull the lever five times. What are the chances that I'll get at least one apple?" Suppose that the wheels spin independently and that the five symbols are equally likely to appear on each wheel in a given spin.
- (a) Find the probability that the slot player gets at least one apple in one pull of the lever. Show your method clearly.
 (b) Now answer the reader's question. Show your method clearly.
- Exercises 107 and 108 refer to the Case Closed about "normal" body temperature on page 585.*
107. Normal body temperature (8.3) Check that the conditions are met for performing inference about the mean body temperature in the population of interest.
108. Normal body temperature (8.2) If "normal" body temperature really is 98.6°F, we would expect the proportion p of all healthy 18- to 40-year-olds who have body temperatures less than this value to be 0.5. Construct and interpret a 95% confidence interval for p . What conclusion would you draw?