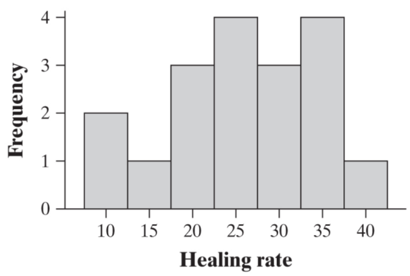
Section 8.3

Check Your Understanding, page 514:

1. df = 22 – 1 = 21, t^* = 2.189. *Using technology*: invT(area: 0.02, df = 21) = –2.189, so t^* = 2.189.
2. df = 71 – 1 = 70, t^* = 2.660 (using df = 60). *Using technology*: invT(area: 0.005, df = 70) = –2.648, so t^* = 2.648.

Check Your Understanding, page 522:

State: We are trying to estimate μ = the true mean healing rate at a 95% confidence level. *Plan*: We should use a one-sample *t* interval for μ if the conditions are met. Random: The description says that the newts were randomly chosen. 10%: The sample size (18) is less than 10% of the population of newts. Normal/Large Sample: The histogram below shows no strong skewness or outliers, so this condition is met.



Do: From the data, we find that $\bar{x} = 25.67$ and $s_x = 8.32$. For a 95% confidence interval and a sample size of 18, $t^* = 2.110$. The confidence interval is:
$25.67 \pm 2.110\left(\frac{8.32}{\sqrt{18}}\right) = 25.67 \pm 4.14 = (21.53, 29.81)$. *Conclude*: We are 95% confident that the interval from 21.53 to 29.81 micrometers per hour captures the true mean healing rate for newts.

Check Your Understanding, page 524:

(new) The margin of error is defined to be $z^* \frac{\sigma}{\sqrt{n}}$. Using σ = 154 and z^* = 1.645 for 90% confidence, $30 \geq 1.645\frac{154}{\sqrt{n}}$. Thus, $n \geq \left(\frac{1.645(154)}{30}\right)^2 = 71.3$, so take a sample of 72 students.

Exercises, page 527:

8.55 (a) df = 9, t^* = 2.262. *Using technology*: invT(area: 0.025, df = 9) = –2.262, so t^* = 2.262.
(b) df = 19, t^* = 2.861. *Using technology*: invT(area: 0.005, df = 19) = –2.861, so t^* = 2.861.
(c) Use df = 60, t^* = 1.671. *Using technology*: invT(area: 0.05, df = 76) = –1.665, so t^* = 1.665.

8.56 (a) df = 11, t^* = 1.796. *Using technology*: invT(area: 0.05, df = 11) = –1.796, so t^* = 1.796.
(b) df = 29, t^* = 2.045. *Using technology*: invT(area: 0.025, df = 29) = –2.045, so t^* = 2.045.
(c) Use df = 50, t^* = 2.678. *Using technology*: invT(area: 0.005, df = 57) = –2.665, so t^* = 2.665.

8.57 Because the sample size is small (n = 20 < 30) and there are outliers in the data, it is not appropriate to use a t critical value to calculate the confidence interval. We cannot assume that the population is approximately Normal.

8.58 Because the sample size is small (n = 28 < 30) and there are outliers in the data, it is not appropriate to use a t critical value to calculate the confidence interval. We cannot assume that the population is approximately Normal.

8.59 (a) A t procedure would not be appropriate here because we are trying to estimate a population proportion, not a population mean.
(b) A t procedure would not be appropriate here because the 15 team members are not a random sample from the population (all male college students at this school).
(c) A t procedure would not be appropriate here because the sample size is small (n = 25 <30) and there are outliers in the sample. We cannot assume that the population is approximately Normal.

8.60 (a) A t procedure is not appropriate here because we have measurements on the entire population. We do not need to estimate the population mean with a confidence interval—we can calculate it exactly.
(b) The t procedure is not appropriate here because members of the AP® Statistics class are not a random sample of all students. These students may not be representative of the population of all students.
(c) The t procedure is appropriate here. Random: The words were selected at random. 10%: the sample size (100) is less than 10% of the number of words in the entire medical journal. Normal/Large Sample: The sample size is large ($n = 100 \geq 30$).

8.61 $\text{SE}_{\bar{x}} = \frac{s_x}{\sqrt{n}} = \frac{9.3}{\sqrt{27}} = 1.7898$. If we take many samples of size 27, the sample mean blood pressure will typically vary by about 1.7898 from the population mean blood pressure.

8.62 $\text{SE}_{\bar{x}} = \frac{s_x}{\sqrt{n}} = \frac{21.88}{\sqrt{20}} = 4.8925$. If we take many samples of size 20, the sample mean commute time will typically vary by about 4.8925 from the population mean commute time.

8.63 (a) Because $\text{SE}_{\bar{x}} = 19.03 = \frac{s_x}{\sqrt{n}} = \frac{s_x}{\sqrt{23}}$ we get $s_x = 19.03\sqrt{23} = 91.26\text{cm}$.
(b) Because the researchers only added 1 standard error, they are using a critical value of $t^* = 1$. With df = 23 – 1 = 22, the area between $t = -1$ and $t = 1$ is approximately tcdf(lower: –1, upper: 1, df: 22) = 0.67. So, the confidence level is 67%.

8.64 (a) Because $\text{SE}_{\bar{x}} = 45 = \frac{s_x}{\sqrt{n}} = \frac{s_x}{\sqrt{12}}$ we get $s_x = 45\sqrt{12} = 155.88$ milliseconds.
(b) Because the researchers only added 1 standard error, they are using a critical value of $t^* = 1$. With df = 12 – 1 = 11, the area between $t = -1$ and $t = 1$ is approximately tcdf(lower: –1, upper: 1, df: 11) = 0.66. So, the confidence level is 66%.

8.65 (a) *State*: We want to estimate μ = the true mean percent change in BMC for breast-feeding mothers at the 99% confidence level. *Plan*: We should use a one-sample t interval for μ

if the conditions are met. Random: the mothers were randomly selected. 10%: 47 is less than 10% of all breast-feeding mothers. Normal/Large Sample: $n = 47 \geq 30$. *Do*: Because $n = 47$, df = 46 and $t^* = 2.704$ (using Table B and df = 40). The confidence interval is

$-3.587 \pm 2.704\left(\frac{2.506}{\sqrt{47}}\right) = -3.587 \pm 0.988 = (-4.575, -2.599)$. *Using technology*: (–4.569,–2.605) with df = 46. *Conclude*: We are 99% confident that the interval from –4.569 to –2.605 captures the true mean percent change in BMC for breast-feeding mothers.
(b) Because all of the plausible values in the interval are negative (indicating bone loss), the data give convincing evidence that breast-feeding mothers lose bone mineral, on average.

8.66 (a) State: We want to estimate μ = the true mean score for all Atlanta eighth-graders at the 99% confidence level. *Plan*: We should use a one-sample *t* interval for μ if the conditions are met. Random: the students were randomly selected. 10%: 1470 is less than 10% of eighth-graders in Atlanta. Normal/Large Sample: $n = 1470 \geq 30$. *Do*: Because $n = 1470$, df = 1469 and $t^* = 2.581$ (using Table B and df = 1000). The confidence interval is

$240 \pm 2.581\left(\frac{42.17}{\sqrt{1470}}\right) = 240 \pm 2.84 = (237.16, 242.84)$. *Using technology*: (237.16, 242.84) with df = 1469. *Conclude*: We are 99% confident that the interval from 237.16 to 242.84 captures the true mean score for all Atlanta eighth-graders.
(b) Because all of the plausible values in the interval are below 243, there is convincing evidence that the mean score for all Atlanta eighth-graders is below the basic level.

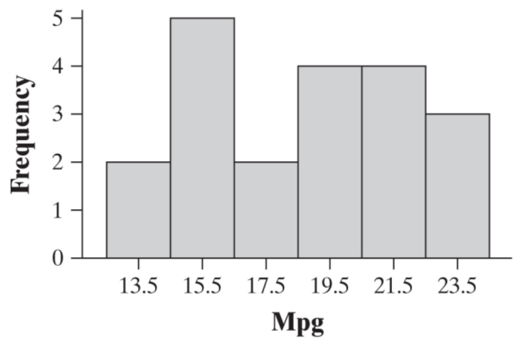
8.67 (a) State: We want to estimate μ = the true mean size of the muscle gap for the population of American and European young men at the 95% confidence level. *Plan*: We should use a one-sample *t* interval for μ if the conditions are met. Random: the young men were randomly selected. 10%: 200 is less than 10% of young men in America and Europe. Normal/Large Sample: $n = 200 \geq 30$. *Do*: Because $n = 200$, df = 199 and $t^* = 1.984$ (using Table B and df = 100). The confidence interval is $2.35 \pm 1.984\left(\frac{2.5}{\sqrt{200}}\right) = 2.35 \pm 0.351 = (1.999, 2.701)$. *Using technology*: (2.001, 2.699) with df = 199. *Conclude*: We are 95% confident that the interval from 2.001 to 2.699 captures the true mean size of the muscle gap for the population of American and European young men.
(b) Even though the population is unlikely to be approximately Normal, the large sample size ($n = 200 \geq 30$) allows us to use a *t* interval for μ.

8.68 (a) *State*: We want to estimate μ = the true mean angle of deformity in the population of all such patients at the 90% confidence level. *Plan*: We should construct a one-sample *t* interval for μ if the conditions are met. Random: The data come from a random sample. 10%: The sample size (37) is less than 10% of the population of people under age 21 with a bunion on the big toe. Normal/Large Sample: $n = 37 \geq 30$. *Do*: Because $n = 37$, df = 36 and $t^* = 1.697$ (using Table B and df = 30). The confidence interval is

$24.76 \pm 1.697\left(\frac{6.34}{\sqrt{37}}\right) = 24.76 \pm 1.769 = (22.991, 26.529)$. *Using technology*: (23.00, 26.52) with df = 36. *Conclude*: We are 90% confident that the interval from 23.00 to 26.52 degrees captures the true mean angle of deformity for the population of all such patients.

(b) Adding back the outlier would increase the mean, making the center of the interval larger. It would also increase the standard deviation, making the interval wider.

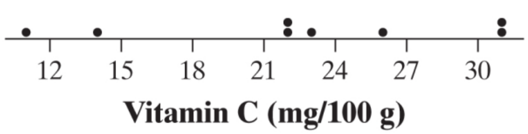
8.69 *State*: We want to estimate μ = the true mean fuel efficiency for this vehicle at a 95% confidence level. *Plan*: We should use a one-sample *t* interval for μ if the conditions are met. Random: the records were selected at random. 10%: it is reasonable to assume that 20 is less than 10% of all records for this vehicle. Normal/Large Sample: the histogram does not show any strong skewness or outliers so this condition is met.



Do: From the data we find that $\bar{x}$ = 18.48, s_x = 3.116, and *n* = 20. Thus, df = 19 and *t** = 2.093. The confidence interval is $18.48 \pm 2.093\left(\frac{3.116}{\sqrt{20}}\right) = 18.48 \pm 1.458 = (17.022, 19.938)$.

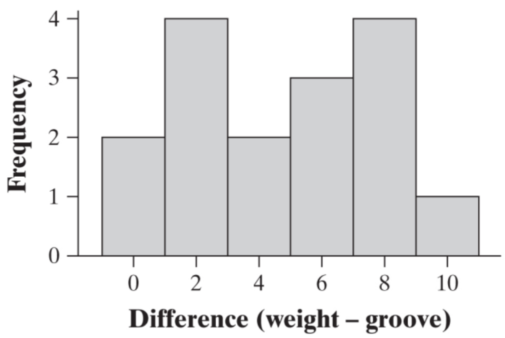
Conclude: We are 95% confident that the interval from 17.022 to 19.938 captures the true mean fuel efficiency for this vehicle.

8.70 *State*: We want to estimate μ = the true mean amount of vitamin C in the CSB from this production run at the 95% confidence level. *Plan*: We should construct a one-sample *t* interval for μ if the conditions are met. Random: The data come from a random sample. 10%: We assume the sample size (8) is less than 10% of the CSB from this run. Normal/Large Sample: The dotplot does not show any strong skewness or outliers so this condition is met.



Do: From the data we find that $\bar{x} = 22.5$, $s_x = 7.19$, and $n = 8$. Thus, df = 7 and $t^* = 2.365$. The confidence interval is $22.5 \pm 2.365\left(\frac{7.19}{\sqrt{8}}\right) = 22.5 \pm 6.01 = (16.49, 28.51)$. *Conclude*: We are 95% confident that the interval from 16.49 to 28.51 mg/100 g captures the true mean amount of vitamin C in this production run.

8.71 (a) *State*: We want to estimate μ = the true mean difference in the estimates from these two methods in the population of tires at the 95% confidence level. *Plan*: We should construct a one-sample *t* interval for μ if the conditions are met. Random: A random sample of tires was selected. 10%: the sample size (16) is less than 10% of all tires. Normal/Large Sample: The histogram of differences shows no strong skewness or outliers, so this condition is met.

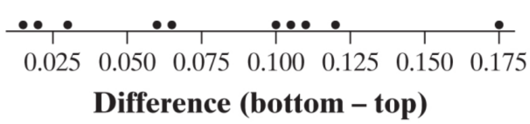


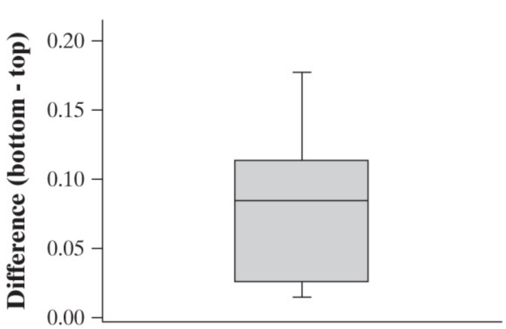
Do: We compute from the differences that $\bar{x} = 4.556$, $s_x = 3.226$, and $n = 16$. Thus, df = 15 and $t^* = 2.131$. The confidence interval is

$4.556 \pm 2.131\left(\frac{3.226}{\sqrt{16}}\right) = 4.556 \pm 1.719 = (2.837, 6.275)$. *Conclude*: We are 95% confident that the interval from 2.837 to 6.275 thousands of miles captures the true mean difference in the estimates from these two methods in the population of tires.

(b) Because 0 is not included in the confidence interval, there is convincing evidence of a difference in the two methods of estimating tire wear.

8.72 (a) *State*: We want to estimate μ = the true mean difference in the amount of zinc in the top and the bottom of wells in this large region at the 95% confidence level. *Plan*: We should construct a one-sample *t* interval for μ if the conditions are met. Random: A random sample of wells was selected. 10%: The sample size (10) is less than 10% of the population of wells in this large region. Normal/Large Sample: The dotplot and boxplot show no strong skewness or outliers, so this condition is met.





Do: We compute from the differences that $\bar{x} = 0.0804$, $s_x = 0.0523$, and $n = 10$. Thus df = 9 and $t^* = 2.262$. The confidence interval is

$0.0804 \pm 2.262\left(\frac{0.0523}{\sqrt{10}}\right) = 0.0804 \pm 0.0374 = (0.043, 0.1178)$. *Conclude*: We are 95% confident that the interval from 0.043 to 0.1178 mg/l captures the true mean difference in the amount of zinc between the bottom and the top of the wells in this large region.

(b) Because 0 is not included in the confidence interval, there is convincing evidence of a difference in zinc concentrations at the top and bottom of wells in this large region.

8.73 We need $2.576\frac{7.5}{\sqrt{n}} \le 1$ so $n \ge \left(\frac{2.576(7.5)}{1}\right)^2 = 373.26$. Take an SRS of 374 women.

8.74 We need $1.96\frac{50}{\sqrt{n}} \le 2$ so $n \ge \left(\frac{1.96(50)}{2}\right)^2 = 2401$. Take a sample of 2401 students who took the SAT a second time.

8.75 b.

8.76 e.

8.77 b.

8.78 a.

8.79 (a) We know that the sum of the probabilities must be 1, so

$$\begin{aligned} P(X=7) &= 1 - \left(P(X=0) + P(X=1) + P(X=2) + P(X=3) + P(X=4) + P(X=5) + P(X=6)\right) \\ &= 1 - (0.04 + 0.03 + 0.06 + 0.08 + 0.09 + 0.08 + 0.05) \\ &= 1 - 0.43 = 0.57 \end{aligned}$$

(b) The mean number of days that a randomly selected young person (aged 19 to 25) watched television is

$$\begin{aligned} \mu_X &= 0(0.04) + 1(0.03) + 2(0.06) + 3(0.08) + 4(0.09) + 5(0.08) + 6(0.05) + 7(0.57) \\ &= 0 + 0.03 + 0.12 + 0.24 + 0.36 + 0.40 + 0.30 + 3.99 \\ &= 5.44 \end{aligned}$$

If we were to randomly select many young people, the average number of days they watched television in the past 7 days would be about 5.44.

(c) **Step 1: State the distribution and values of interest.** First we need to find the standard deviation of *X*:

$$\sigma_X = \sqrt{(0-5.44)^2\,0.04 + \mathrm{L} \;\; + (7-5.44)^2\,0.57}$$
$$= \sqrt{4.5664}$$
$$= 2.14$$

Because the sample size is large ($n = 100 \geq 30$), we expect the mean number of days $\bar{x}$ for 100 randomly selected young people (aged 19 to 25) to be approximately Normally distributed with mean $\mu_{\bar{x}} = \mu = 5.44$. Because the sample size (100) is less than 10% of all young people aged 19 to 25, the standard deviation is $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{2.14}{\sqrt{100}} = 0.214$. We want to find $P(\bar{x} \leq 4.96)$. **Step 2: Perform calculations. Show your work.** The standardized score for a sample mean of 4.96 is $z = \dfrac{4.96 - 5.44}{0.214} = -2.24$ and the desired probability is $P(Z \leq -2.24) = 0.0125$. *Using technology:* normalcdf(lower: –1000, upper: 4.96, μ: 5.44, σ: 0.214) = 0.0124. **Step 3: Answer the question.** There is a 0.0124 probability of getting a sample mean of 4.96 or smaller. Because this probability is small, a sample mean of 4.96 or smaller would be surprising.

8.80 (a) There will be 8 treatment groups, with 25 people randomly assigned to each group. The treatments are:

Treatment 1: 25% of food on sale, 60% off
Treatment 2: 50% of food on sale, 60% off
Treatment 3: 75% of food on sale, 60% off
Treatment 4: 100% of food on sale, 60% off
Treatment 5: 25% of food on sale, 40-70% off
Treatment 6: 50% of food on sale, 40-70% off
Treatment 7: 75% of food on sale, 40-70% off
Treatment 8: 100% of food on sale, 40-70% off

Researchers will compare the mean attractiveness rating given by individuals in the eight groups.
(b) Because there are 200 subjects, we label the subjects 001, 002, …, 200. Moving from left to right through the table of random digits, select three-digit numbers. The labels 000 and 201 to 999 are not assigned to a subject, so we ignore them. We also ignore any repeats of a label, because that subject has already been assigned to a treatment group. Once we have 25 subjects for the first treatment, we select 25 subjects for the second treatment, and so on, until all subjects have been assigned to a treatment group. Using the digits provided: 457 (ignore), 404 (ignore), 180 (assign to treatment 1), 765 (ignore), 561 (ignore), 333 (ignore), 020 (assign to treatment 1), 705 (ignore), 193 (assign to treatment 1).
(c) The range "40% to 70% off" slowly decreases in attractiveness to customers as the percent of food on sale increases. However, the precise "60% off" grows increasingly attractive to customers as the percent of food on sale increases. When only 25% of food is on sale, customers rate the "40% to 70% off" range as more attractive than the precise "60% off." For all other percents of foods on sale, the precise "60% off" is more attractive to customers and becomes more and more attractive than the range "40% to 70% off" as the percent of food items on sale increases.