Inferences Based on a Single Sample
Estimation with Confidence Intervals

Chapter 5

5.2 The confidence coefficient in a confidence interval is the probability that an interval estimator encloses the population parameter. For a 90% confidence interval, the probability that an interval will enclose the population parameter is .90. In other words, if one took repeated samples and formed 90% confidence intervals for μ, 90% of the intervals will contain μ and 10% will not.

5.4 If we were to repeatedly draw samples of size n from the population and form the interval $\bar{x} \pm 1.96\sigma_{\bar{x}}$ each time, approximately 95% of the intervals would contain μ. We have no way of knowing whether our interval estimate is one of the 95% that contain μ or one of the 5% that do not.

5.6 The conditions necessary to form a valid large-sample confidence interval for μ are:

1. A random sample is selected from a target population.
2. The sample size n is large, i.e., $n \geq 30$

5.8 a. $z_{\alpha/2} = 1.96$, using Table III, Appendix A, $P(0 \leq z \leq 1.96) = .4750$. Thus, $\alpha/2 = .5000 - .4750 = .025$, $\alpha = 2(.025) = .05$, and $1-\alpha = 1-.05 = .95$. The confidence level is $100\% \times .95 = 95\%$.

b. $z_{\alpha/2} = 1.645$, using Table III, Appendix A, $P(0 \leq z \leq 1.645) = .45$. Thus, $\alpha/2 = .50 - .45 = .05$, $\alpha = 2(.05) = .10$, and $1-\alpha = 1-.10 = .90$. The confidence level is $100\% \times .90 = 90\%$.

c. $z_{\alpha/2} = 2.575$, using Table III, Appendix A, $P(0 \leq z \leq 2.575) = .495$. Thus, $\alpha/2 = .500 - .495 = .005$, $\alpha = 2(.005) = .01$, and $1-\alpha = 1-.01 = .99$. The confidence level is $100\% \times .99 = 99\%$.

d. $z_{\alpha/2} = 1.28$, using Table III, Appendix A, $P(0 \leq z \leq 1.28) = .4$. Thus, $\alpha/2 = .50 - .40 = .10$, $\alpha = 2(.10) = .20$, and $1-\alpha = 1-.20 = .80$. The confidence level is $100\% \times .80 = 80\%$.

e. $z_{\alpha/2} = .99$, using Table III, Appendix A, $P(0 \leq z \leq .99) = .3389$. Thus, $\alpha/2 = .5000 - .3389 = .1611$, $\alpha = 2(.1611) = .3222$, and $1-\alpha = 1-.3222 = .6778$. The confidence level is $100\% \times .6778 = 67.78\%$.

5.10 a. For confidence coefficient .95, $\alpha = .05$ and $\alpha/2 = .05/2 = .025$. From Table III, Appendix A, $z_{.025} = 1.96$. The confidence interval is:

$$\bar{x} \pm z_{.025}\frac{s}{\sqrt{n}} \Rightarrow 25.9 \pm 1.96\frac{2.7}{\sqrt{90}} \Rightarrow 25.9 \pm .56 \Rightarrow (25.34, 26.46)$$

b. For confidence coefficient .90, $\alpha = .10$ and $\alpha / 2 = .10 / 2 = .05$. From Table III, Appendix A, $z_{.05}$ = 1.645. The confidence interval is:

$$\bar{x} \pm z_{.05} \frac{s}{\sqrt{n}} \Rightarrow 25.9 \pm 1.645 \frac{2.7}{\sqrt{90}} \Rightarrow 25.9 \pm .47 \Rightarrow (25.43, 26.37)$$

c. For confidence coefficient .99, $\alpha = .01$ and $\alpha / 2 = .01 / 2 = .005$. From Table III, Appendix A, $z_{.005}$ = 2.58. The confidence interval is:

$$\bar{x} \pm z_{.005} \frac{s}{\sqrt{n}} \Rightarrow 25.9 \pm 2.58 \frac{2.7}{\sqrt{90}} \Rightarrow 25.9 \pm .73 \Rightarrow (25.17, 26.63)$$

5.12 a. For confidence coefficient .95, $\alpha = .05$ and $\alpha / 2 = .05 / 2 = .025$. From Table III, Appendix A, $z_{.025}$ = 1.96. The confidence interval is:

$$\bar{x} \pm z_{.025} \frac{s}{\sqrt{n}} \Rightarrow 33.9 \pm 1.96 \frac{3.3}{\sqrt{100}} \Rightarrow 33.9 \pm .647 \Rightarrow (33.253, 34.547)$$

b. $$\bar{x} \pm z_{.025} \frac{s}{\sqrt{n}} \Rightarrow 33.9 \pm 1.96 \frac{3.3}{\sqrt{400}} \Rightarrow 33.9 \pm .323 \Rightarrow (33.577, 34.223)$$

c. For part **a**, the width of the interval is 2(.647) = 1.294. For part **b**, the width of the interval is 2(.323) = .646. When the sample size is quadrupled, the width of the confidence interval is halved.

5.14 a. From the printout, the 90% confidence interval is (93.53, 109.71). We are 90% confident that the true mean fasting blood sugar level in hypertensive patients is between 93.53 and 109.71.

b. From the printout, the 90% confidence interval is (1.92771, 1.95429). We are 90% confident that the true mean magnesium level in hypertensive patients is between 1.92771 and 1.95429.

c. If the confidence level is raised to 95%, the width of the interval will increase. With more confidence, we must include more number.

d. If the sample size is increased from 50 to 100, the width of the interval should decrease. The standard deviation of $\bar{x}$ will be $\frac{s}{\sqrt{100}}$ instead of $\frac{s}{\sqrt{50}}$.

5.16 a. The target parameter is μ = average amount of time (in minutes) per day laptops are used for taking notes for all middle school students across the country.

b. The standard deviation of the sampling distribution of $\bar{x}$ is estimated with $\frac{s}{\sqrt{n}}$, not just *s*.

c. For confidence coefficient .90, $\alpha = .10$ and $\alpha / 2 = .05$. From Table III, Appendix A, $z_{.05}$ = 1.645. The confidence interval is:

$$\bar{x} \pm z_{.05}\frac{s}{\sqrt{n}} \Rightarrow 13.2 \pm 1.645\frac{19.5}{\sqrt{106}} \Rightarrow 13.2 \pm 3.116 \Rightarrow (10.084,\ 16.316)$$

We are 90% confident that the true average amount of time per day laptops are used for taking notes for all middle school students across the country is between 10.084 and 16.316 minutes.

d. "90% confidence" means that in repeated sampling, 90% of all confidence intervals constructed in this manner will contain the true mean.

e. No, this is not a problem. The Central Limit Theorem says that the sampling distribution of $\bar{x}$ will be approximately normal, regardless of the shape of the population being sampled from, as long as the sample size is relatively large. The sample size for this distribution is 106, which is sufficiently large.

5.18 a. The target parameter is the mean egg length for all New Zealand birds, μ.

b. Using MINITAB, 50 uniform random numbers were generated between 1 and 132. Those random numbers were:

3, 9, 12, 13, 14, 15, 17, 22, 23, 24, 25, 26, 27, 29, 37, 39, 45, 49, 50, 51, 52, 54, 60, 65, 68, 70, 71, 75, 93, 95, 96, 97, 98, 99, 101, 105, 108, 109, 110, 112, 118, 119, 120, 123, 126, 127, 128, 129, 130, 131

The egg lengths corresponding to these numbers were then selected. The 50 egg lengths are:

57, 44, 31, 36, 61, 67, 69, 48, 58, 61, 59, 65, 66, 46, 56, 64, 49, 40, 58, 30, 28, 43, 45, 42, 52, 40, 46, 74, 26, 23, 23, 19.5, 20, 23.5, 17, 19, 40, 35, 29, 45, 110, 124, 160, 205, 192, 125, 195, 218, 236, 94

c. Using MINITAB, the descriptive statistics are:

Descriptive Statistics: Egg Length

Variable	N	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Egg Length	50	8.28	55.65	17.00	34.00	48.50	67.50	236.00

The mean is $\bar{x} = 68.28$ and the standard deviation is $s = 55.65$.

d. For confidence coefficient .99, $\alpha = .01$ and $\alpha / 2 = .01/2 = .005$. From Table III, Appendix A, $z_{.005} = 2.58$. The 99% confidence interval is:

$$\bar{x} \pm z_{.005}\sigma_{\bar{x}} \Rightarrow \bar{x} \pm 2.58\frac{\sigma}{\sqrt{n}} \Rightarrow 68.28 \pm 2.58\frac{55.65}{\sqrt{50}} \Rightarrow 68.28 \pm 20.30 \Rightarrow (47.98,\ \ 88.58)$$

e. We are 99% confident that the mean egg length of the bird species of New Zealand is between 47.98 and 88.58 mm.

5.20 a. The parameter of interest is μ = the true mean shell length of all green sea turtles in the lagoon.

b. Using MINITAB, the descriptive statistics are:

Descriptive Statistics: Length

Variable	N	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Length	76	55.47	11.34	30.37	49.20	56.79	64.70	81.63

The point estimate of μ is $\bar{x} = 55.47$.

c. For confidence coefficient .95, $\alpha = .05$ and $\alpha / 2 = .025$. From Table III, Appendix A, $z_{.025}$ = 1.96. The confidence interval is:

$$\bar{x} \pm z_{.025} \frac{s}{\sqrt{n}} \Rightarrow 55.47 \pm 1.96 \frac{11.34}{\sqrt{76}} \Rightarrow 55.47 \pm 2.55 \Rightarrow (52.92,\ 58.02)$$

We are 95% confident that the true mean shell length of all green sea turtles in the lagoon is between 52.92 and 58.02 centimeters.

d. Since 60 is not in the 95% confidence interval, it is not a likely value for the true mean. Thus, we would be very suspicious of the claim.

5.22 a. Using MINITAB, the descriptive statistics are:

Descriptive Statistics: Times

Variable	N	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Times	38	0.1879	0.1814	-0.0100	0.0500	0.1400	0.3000	0.9000

For confidence coefficient .95, $\alpha = .05$ and $\alpha / 2 = .025$. From Table III, Appendix A, $z_{.025}$ = 1.96. The confidence interval is:

$$\bar{x} \pm z_{.025} \frac{s}{\sqrt{n}} \Rightarrow .1879 \pm 1.96 \frac{.1814}{\sqrt{38}} \Rightarrow .1879 \pm .0577 \Rightarrow (.1302,\ .2456)$$

We are 95% confident that the true mean decrease in sprint times for the population of all football players who participate in the speed training program is between .1302 and .2456 seconds.

b. Yes, the training program really is effective. If the training program was not effective, then the mean decrease in sprint times would be 0. We note that 0 is not in the 95% confidence interval. Therefore, it is not a likely value for the true mean.

5.24 Using MINITAB, the descriptive statistics are:

Descriptive Statistics: ATTIMES

Variable	N	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
ATTIMES	50	20.85	13.41	0.800	10.35	19.65	30.18	48.20

For confidence coefficient .90, $\alpha = .10$ and $\alpha / 2 = .10 / 2 = .05$. From Table III, Appendix A, $z_{.05}$ = 1.645. The 90% confidence interval is:

$$\bar{x} \pm z_{.05}\sigma_{\bar{x}} \Rightarrow \bar{x} \pm 1.645\frac{\sigma}{\sqrt{n}} \Rightarrow 20.85 \pm 1.645\frac{13.41}{\sqrt{50}} \Rightarrow 20.85 \pm 3.12 \Rightarrow (17.73,\ \ 23.97)$$

We are 90% confident that the mean attention time given to all twin boys by their parents is between 17.73 and 23.97 hours.

5.26 $\bar{x} = \dfrac{11,298}{5,000} = 2.26$

For confidence coefficient, .95, $\alpha = .05$ and $\alpha / 2 = .05 / 2 = .025$. From Table III, Appendix A, $z_{.025}$ = 1.96. The confidence interval is:

$$\bar{x} \pm z_{\alpha/2}\frac{s}{\sqrt{n}} \Rightarrow 2.26 \pm 1.96\frac{1.5}{\sqrt{5000}} \Rightarrow 2.26 \pm .04 \Rightarrow (2.22,\ 2.30)$$

We are 95% confident the mean number of roaches produced per roach per week is between 2.22 and 2.30.

5.28 Both the z-distribution and the t-distribution are mound-shaped and symmetric with mean 0. The primary difference between the z- and t-distributions is that the t-distribution is more spread out than the z-distribution.

5.30 a. For confidence coefficient .80, $\alpha = 1 - .80 = .20$ and $\alpha / 2 = .20 / 2 = .10$. From Table III, Appendix A, $z_{.10}$ = 1.28. From Table IV, with df = n – 1 = 7 – 1 = 6, $t_{.10}$ = 1.440.

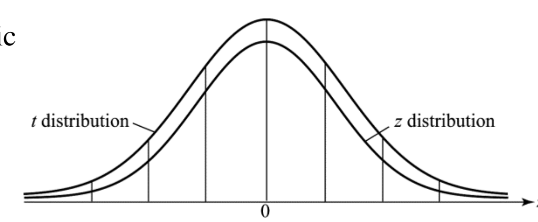
b. For confidence coefficient .90, $\alpha = 1 - .90 = .10$ and $\alpha / 2 = .10 / 2 = .05$. From Table III, Appendix A, $z_{.05}$ = 1.645. From Table IV, with df = n – 1 = 7 – 1 = 6, $t_{.05}$ = 1.943.

c. For confidence coefficient .95, $\alpha = 1 - .95 = .05$ and $\alpha / 2 = .05 / 2 = .025$. From Table III, Appendix A, $z_{.025}$ = 1.96. From Table IV, with df = n – 1 = 7 – 1 = 6, $t_{.025}$ = 2.447.

d. For confidence coefficient .98, $\alpha = 1 - .98 = .02$ and $\alpha / 2 = .02 / 2 = .01$. From Table III, Appendix A, $z_{.01}$ = 2.33. From Table IV, with df = n – 1 = 7 – 1 = 6, $t_{.01}$ = 3.143.

e. For confidence coefficient .99, $\alpha = 1 - .99 = .01$ and $\alpha / 2 = .01 / 2 = .005$. From Table III, Appendix A, $z_{.005}$ = 2.58. From Table IV, with df = n – 1 = 7– 1 = 6, $t_{.005}$ = 3.707.

f. Both the t and z-distributions are symmetric around 0 and mound-shaped. The t-distribution is more spread out than the z-distribution.



5.32 a. $P(-t_0 < t < t_0) = .95$ where df = 16

Because of symmetry, the statement can be written

$P(0 < t < t_0) = .475$ where df = 16

$\Rightarrow P(t \geq t_0) = .5 - .475 = .025$

$t_0 = 2.120$

b. $P(t \leq -t_0 \text{ or } t \geq t_0) = .05$ where df = 16

$\Rightarrow 2P(t \geq t_0) = .05$

$\Rightarrow P(t \geq t_0) = .025$ where df = 16

$t_0 = 2.120$

c. $P(t \leq t_0) = .05$ where df = 16

Because of symmetry, the statement can be written

$P(t \geq -t_0) = .05$ where df = 16

$t_0 = -1.746$

d. $P(t \leq -t_0 \text{ or } t \geq t_0) = .10$ where df = 12

$\Rightarrow 2P(t \geq t_0) = .10$

$\Rightarrow P(t \geq t_0) = .05$ where df = 12

$t_0 = 1.782$

e. $P(t \leq -t_0 \text{ or } t \geq t_0) = .01$ where df = 8

$\Rightarrow 2P(t \geq t_0) = .01$

$\Rightarrow P(t \geq t_0) = .005$ where df = 8

$t_0 = 3.355$

5.34 For this sample,

$$\bar{x} = \frac{\sum x}{n} = \frac{1567}{16} = 97.9375 \qquad s^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1} = \frac{155{,}867 - \frac{1567^2}{16}}{16-1} = 159.9292$$

$$s = \sqrt{s^2} = 12.6463$$

a. For confidence coefficient, .80, $\alpha = 1 - .80 = .20$ and $\alpha / 2 = .20 / 2 = .10$. From Table IV, Appendix A, with df = n − 1 = 16 − 1 = 15, $t_{.10} = 1.341$. The 80% confidence interval for μ is:

$$\bar{x} \pm t_{.10}\frac{s}{\sqrt{n}} \Rightarrow 97.94 \pm 1.341\frac{12.6463}{\sqrt{16}} \Rightarrow 97.94 \pm 4.240 \Rightarrow (93.700,\ 102.180)$$

b. For confidence coefficient, .95, $\alpha = 1 - .95 = .05$ and $\alpha / 2 = .05 / 2 = .025$. From Table IV, Appendix A, with df = n – 1 = 24 – 1 = 23, $t_{.025}$ = 2.131. The 95% confidence interval for μ is:

$$\bar{x} \pm t_{.250} \frac{s}{\sqrt{n}} \Rightarrow 97.94 \pm 2.131 \frac{12.6463}{\sqrt{16}} \Rightarrow 97.94 \pm 6.737 \Rightarrow (91.203,\ 104.677)$$

The 95% confidence interval for μ is wider than the 80% confidence interval for μ found in part **a**.

c. For part **a**:

We are 80% confident that the true population mean lies in the interval 93.700 to 102.180.

For part **b**:

We are 95% confident that the true population mean lies in the interval 91.203 to 104.677.

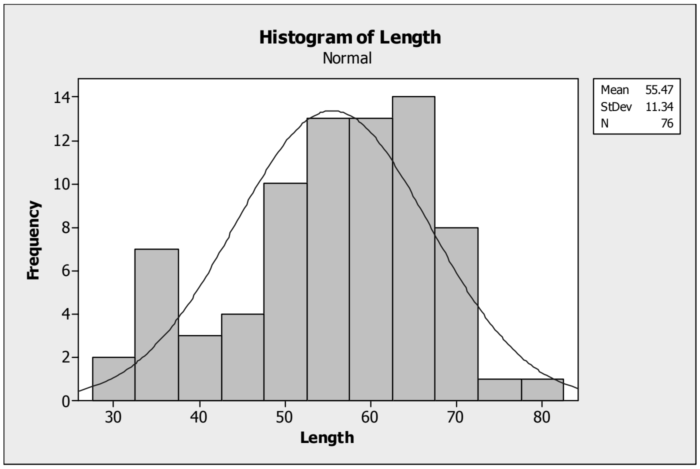
The 95% confidence interval is wider than the 80% confidence interval because the more confident you want to be that μ lies in an interval, the wider the range of possible values.

5.36 a. For confidence coefficient .99, $\alpha = .01$ and $\alpha / 2 = .005$. From Table IV, Appendix A, with df = n – 1 = 6 – 1 = 5, $t_{.005}$ = 4.032. The confidence interval is:

$$\bar{x} \pm t_{.005,5} \frac{s}{\sqrt{n}} \Rightarrow 52.9 \pm 4.032 \frac{6.8}{\sqrt{6}} \Rightarrow 52.9 \pm 11.91 \Rightarrow (41.71,\ 64.09)$$

We are 99% confident that the true mean shell length of all green sea turtles in the lagoon is between 41.71 and 64.09 cm.

b. We must assume that the data are sampled from a normal distribution and that the sample was randomly selected. A histogram of the data is:



These data do not look that normal. The confidence interval may not be valid.

5.38 For confidence coefficient .90, $\alpha = .10$ and $\alpha/2 = .10/2 = .05$. From Table IV, Appendix A, with df = n – 1 = 25 – 1 = 24, $t_{.05}$ = 1.711. The 90% confidence interval is:

$$\bar{x} \pm t_{.05,24} \frac{s}{\sqrt{n}} \Rightarrow 75.4 \pm 1.711 \frac{10.9}{\sqrt{25}} \Rightarrow 75.4 \pm 3.73 \Rightarrow (71.67, \quad 79.13)$$

We are 90% confident that the true mean breaking strength of white wood is between 71.67 and 79.13 MPa’s.

5.40 a. The point estimate for the average annual rainfall amount at ant sites in the Dry Steppe region of Central Asia is $\bar{x}$ =183.4 milliliters.

b. For confidence coefficient .90, $\alpha = .10$ and $\alpha/2 = .10/2 = .05$. From Table IV, Appendix A, with df = n – 1 = 5 – 1 = 4, $t_{.05}$ = 2.132.

c. The 90% confidence interval is:

$$\bar{x} \pm t_{.05} \frac{s}{\sqrt{n}} \Rightarrow 183.4 \pm 2.132 \frac{20.6470}{\sqrt{5}} \Rightarrow 183.4 \pm 19.686 \Rightarrow (163.714, \quad 203.086)$$

d. We are 90% confident that the average annual rainfall amount at ant sites in the Dry Steppe region of Central Asia is between 163.714 and 203.086 milliliters.

e. Using MINITAB, the 90% confidence interval is:

One-Sample T: DS Rain

```
Variable  N     Mean   StDev  SE Mean        90% CI
DS Rain   5  183.400  20.647    9.234  (163.715, 203.085)
```

The 90% confidence interval is (163.715, 203.085). This is very similar to the confidence interval calculated in part c.

f. The point estimate for the average annual rainfall amount at ant sites in the Gobi Desert region of Central Asia is $\bar{x} = 110.0$ milliliters.

For confidence coefficient .90, $\alpha = .10$ and $\alpha/2 = .10/2 = .05$. From Table IV, Appendix A, with df = n – 1 = 6 – 1 = 5, $t_{.05}$ = 2.015.

The 90% confidence interval is:

$$\bar{x} \pm t_{.05}\frac{s}{\sqrt{n}} \Rightarrow 110.0 \pm 2.015\frac{15.975}{\sqrt{6}} \Rightarrow 110.0 \pm 13.141 \Rightarrow (96.859,\ \ 123.141)$$

We are 90% confident that the average annual rainfall amount at ant sites in the Gobi Desert region of Central Asia is between 96.859 and 123.141 milliliters.

Using MINITAB, the 90% confidence interval is:

One-Sample T: GD Rain

```
Variable  N     Mean   StDev  SE Mean         90% CI
GD Rain   6  110.000  15.975    6.522  (96.858, 123.142)
```

The 90% confidence interval is (96.858, 123.142). This is very similar to the confidence interval calculated above.

5.42 Some preliminary calculations are:

$$\bar{x} = \frac{\sum x}{n} = \frac{247}{13} = 19 \qquad s^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1} = \frac{4751 - \frac{247^2}{13}}{13-1} = \frac{58}{12} = 4.8333$$

$$s = \sqrt{4.8333} = 2.198$$

For confidence coefficient .99, $\alpha = .01$ and $\alpha/2 = .01/2 = .005$. From Table IV, Appendix A, with df = n – 1 = 13 – 1 = 12, $t_{.005}$ = 3.055. The 99% confidence interval is:

$$\bar{x} \pm t_{.005,12}\frac{s}{\sqrt{n}} \Rightarrow 19 \pm 3.055\frac{2.198}{\sqrt{13}} \Rightarrow 19 \pm 1.86 \Rightarrow (17.14,\ \ 20.86)$$

We are 99% confident that the true mean quality of all studies on the treatment of Alzheimer's disease is between 17.14 and 20.86.

5.44 a. Some preliminary calculations are:

$$\bar{x} = \frac{\sum x}{n} = \frac{7,169}{20} = 358.45$$

$$s^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1} = \frac{2,833,465 - \frac{7,169^2}{20}}{20-1} = \frac{263,736.95}{19} = 13,880.89211$$

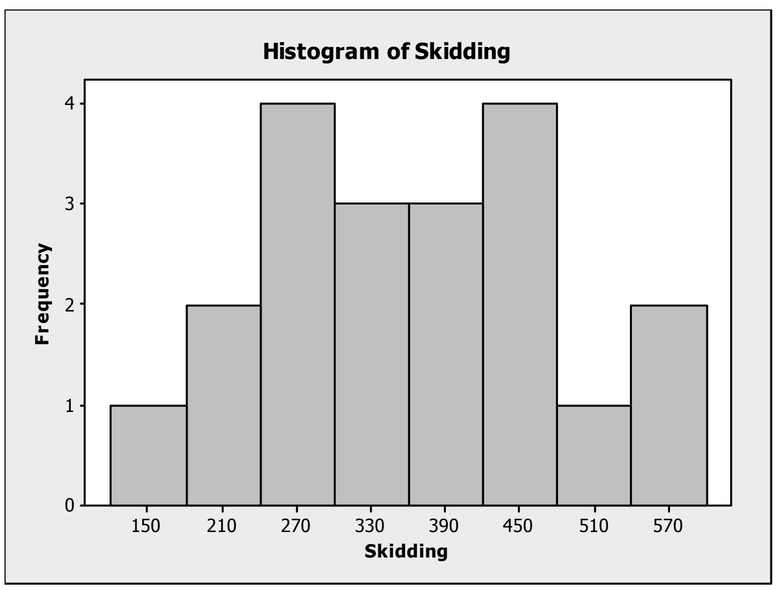
$$s = \sqrt{13,880.89211} = 117.8172$$

For confidence coefficient .95, $\alpha = .05$ and $\alpha / 2 = .05 / 2 = .025$. From Table IV, Appendix A, with df = n – 1 = 20 – 1 = 19, $t_{.025}$ = 2.093. The 95% confidence interval is:

$$\bar{x} \pm t_{.025} \frac{s}{\sqrt{n}} \Rightarrow 358.45 \pm 2.093 \frac{117.8172}{\sqrt{20}} \Rightarrow 358.45 \pm 55.140 \Rightarrow (303.310, \quad 413.590)$$

b. We are 95% confident that the true mean skidding distance for the road is between 303.310 and 413.590 meters.

c. We must assume that the population being sampled from is approximately normal. Using MINITAB, a histogram of the data is:



The data look fairly mound-shaped. This assumption appears to be satisfied.

d. The 95% confidence interval is (303.310, 413.590). Since the value of 425 is not in this interval, it is not a likely value for the true mean skidding distance. We would not agree with the logger.

5.46 a. Some preliminary calculations are:

$$\bar{x} = \frac{\sum x}{n} = \frac{196}{11} = 17.82 \qquad s^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1} = \frac{3{,}734 - \frac{196^2}{11}}{11-1} = 24.1636$$

$$s = \sqrt{26.1636} = 4.92$$

For confidence coefficient .95, $\alpha = .05$ and $\alpha / 2 = .05/2 = .025$. From Table IV, with df = n – 1 = 11 – 1 = 10, $t_{.025}$ = 2.228. The 95% confidence interval is:

$$\bar{x} \pm t_{\alpha/2}\frac{s}{\sqrt{n}} \Rightarrow 17.82 \pm 2.228\frac{4.92}{\sqrt{11}} \Rightarrow 17.82 \pm 3.31 \Rightarrow (14.51,\ 21.13)$$

We are 95% confident that the mean FNE score of the population of bulimic female students is between 14.51 and 21.13.

b. Some preliminary calculations are:

$$\bar{x} = \frac{\sum x}{n} = \frac{198}{14} = 14.14 \qquad s^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1} = \frac{3{,}164 - \frac{198^2}{14}}{14-1} = 27.9780$$

$$s = \sqrt{27.9780} = 5.29$$

For confidence coefficient .95, $\alpha = .05$ and $\alpha / 2 = .05/2 = .025$. From Table IV, with df = n – 1 = 14 – 1 = 13, $t_{.025}$ = 2.160. The 95% confidence interval is:

$$\bar{x} \pm t_{\alpha/2}\frac{s}{\sqrt{n}} \Rightarrow 14.14 \pm 2.160\frac{5.29}{\sqrt{14}} \Rightarrow 14.14 \pm 3.05 \Rightarrow (11.09,\ 17.19)$$

We are 95% confident that the mean FNE score of the population of normal female students is between 11.09 and 17.19.

c. We must assume that the populations of FNE scores for both the bulimic and normal female students are normally distributed.

Stem-and-leaf displays for the two groups are below:

```
Stem-and-leaf of Bulimia   N  = 11
Leaf Unit = 1.0

     1    1 0
     3    1 33
     4    1 4
     5    1 6
    (1)   1 9
     5    2 011
     2    2
     2    2 45


Stem-and-leaf of Normal    N  = 14
Leaf Unit = 1.0

     2    0 67
     3    0 8
     5    1 01
     7    1 33
     7    1 5
     6    1 6
     5    1 899
     2    2 0
     1    2 3
```

From both of these plots, the assumption of normality is questionable for both groups. Neither of the plots looks mound-shaped. However, it is hard to decide with such small sample sizes.

5.48 By the Central Limit Theorem, the sampling distribution of $\hat{p}$ is approximately normal with mean $\mu_{\hat{p}} = p$ and standard deviation $\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$.

5.50 If p is near 0 or 1, an *extremely* large sample size is required. For example if $p = .01$, then $np \geq 15$ implies that $n \geq \frac{15}{p} = \frac{15}{.01} = 1500$.

5.52 a. The sample size is large enough if both $n\hat{p} \geq 15$ and $n\hat{q} \geq 15$.

$n\hat{p} = 144(.76) = 109.44$ and $n\hat{q} = 144(.24) = 34.56$. Since both of these numbers are greater than or equal to 15, the sample size is sufficiently large to conclude the normal approximation is reasonable.

b. For confidence coefficient .90, $\alpha = .10$ and $\alpha / 2 = .10 / 2 = .05$. From Table III, Appendix A, $z_{.05} = 1.645$. The 90% confidence interval is:

$$\hat{p} \pm z_{.05}\sqrt{\frac{pq}{n}} \approx \hat{p} \pm 1.645\sqrt{\frac{\hat{p}\hat{q}}{n}} \Rightarrow .76 \pm 1.645\sqrt{\frac{.76(.24)}{144}} \Rightarrow .76 \pm .059 \Rightarrow (.701, .819)$$

c. We must assume the sample was randomly selected from the population of interest. We must also assume our sample size is sufficiently large to ensure the sampling distribution is approximately normal. From the results of part **a**, this appears to be a reasonable assumption.

5.54 a. Of the 50 observations, 15 like the product $\Rightarrow \hat{p} = \frac{15}{50} = .30$.

The sample size is large enough if both $n\hat{p} \geq 15$ and $n\hat{q} \geq 15$.

$n\hat{p} = 50(.3) = 15$ and $n\hat{q} = 50(.7) = 35$. Since both of these numbers are greater than or equal to 15, the sample size is sufficiently large to conclude the normal approximation is reasonable.

For the confidence coefficient .80, $\alpha = .20$ and $\alpha / 2 = .20 / 2 = .10$. From Table III, Appendix A, $z_{.10}$ = 1.28. The confidence interval is:

$$\hat{p} \pm z_{.10}\sqrt{\frac{\hat{p}\hat{q}}{n}} \Rightarrow .3 \pm 1.28\sqrt{\frac{.3(.7)}{50}} \Rightarrow .3 \pm .083 \Rightarrow (.217,\ .383)$$

b. We are 80% confident the proportion of all consumers who like the new snack food is between .217 and .383.

5.56 a. The estimate of the true proportion of satellite radio subscribers who have a satellite radio receiver in their car is $\hat{p} = \frac{396}{501} = .79$.

b. The sample size is large enough if both $n\hat{p} \geq 15$ and $n\hat{q} \geq 15$. For this problem, $\hat{p} = .79$. $n\hat{p} = 501(.79) = 395.79$ and $n\hat{q} = 501(.21) = 105.21$. Since both of these numbers are greater than or equal to 15, the sample size is sufficiently large to conclude the normal approximation is reasonable.

For confidence coefficient .90, $\alpha = .10$ and $\alpha / 2 = .10 / 2 = .05$. From Table III, Appendix A, $z_{.05}$ = 1.645. The 90% confidence interval is:

$$\hat{p} \pm z_{.05}\sqrt{\frac{pq}{n}} \approx \hat{p} \pm 1.645\sqrt{\frac{\hat{p}\hat{q}}{n}} \Rightarrow .79 \pm 1.645\sqrt{\frac{.79(.21)}{501}} \Rightarrow .79 \pm .030 \Rightarrow (.760,\ .820)$$

c. We are 90% confident that the true proportion of all of satellite radio subscribers who have a satellite radio receiver in their car is between .760 and .820.

5.58 a. The parameter of interest is the true proportion of fillets that are really red snapper.

b. The sample size is large enough if both $n\hat{p} \geq 15$ and $n\hat{q} \geq 15$.

$n\hat{p} = 22(.23) = 5.06$ and $n\hat{q} = 22(.77) = 16.94$. Since the first number is not greater than or equal to 15, the sample size is not sufficiently large to conclude the normal approximation is reasonable.

c. The Wilson adjusted sample proportion is

$$\tilde{p} = \frac{x+2}{n+4} = \frac{5+2}{22+4} = \frac{7}{26} = .269$$

For confidence coefficient .95, $\alpha = .05$ and $\alpha / 2 = .05 / 2 = .025$. From Table III, Appendix A, $z_{.025}$ = 1.96. The Wilson adjusted 95% confidence interval is:

$$\tilde{p} \pm z_{.025}\sqrt{\frac{\tilde{p}(1-\tilde{p})}{n+4}} \Rightarrow .269 \pm 1.96\sqrt{\frac{.269(.731)}{22+4}} \Rightarrow .269 \pm .170 \Rightarrow (.099,\ \ .439)$$

d. We are 95% confident that the true proportion of fish fillets purchased from vendors across the U.S. that are really red snapper is between .099 and .439.

5.60 a. Of the 38 times, only one was negative. Thus, $\hat{p} = \frac{37}{38} = .974$.

b. The sample size is large enough if both $n\hat{p} \geq 15$ and $n\hat{q} \geq 15$. For this problem, $\hat{p} = .974$.

$n\hat{p} = 38(.974) = 37.012$ and $n\hat{q} = 38(.026) = 0.988$. Since one of these numbers is less than 15, the sample size may not be sufficiently large to conclude the normal approximation is reasonable.

The Wilson adjusted sample proportion is

$$\tilde{p} = \frac{x+2}{n+4} = \frac{37+2}{38+4} = \frac{39}{42} = .929$$

For confidence coefficient .95, α = .05 and $\alpha/2$ = .05/2 = .025. From Table III, Appendix A, $z_{.025}$ = 1.96. The Wilson adjusted 95% confidence interval is:

$$\tilde{p} \pm z_{.025}\sqrt{\frac{\tilde{p}(1-\tilde{p})}{n+4}} \Rightarrow .929 \pm 1.96\sqrt{\frac{.929(.071)}{38+4}} \Rightarrow .929 \pm .078 \Rightarrow (.851,\ \ 1.007)$$

We are 95% confident that the true proportion of "improved" sprint times is between .851 and 1.

5.62 First, we compute $\hat{p}$: $\hat{p} = \frac{x}{n} = \frac{88}{504} = .175$

The sample size is large enough if both $n\hat{p} \ge 15$ and $n\hat{q} \ge 15$.

$n\hat{p} = 504(.175) = 88.2$ and $n\hat{q} = 504(.825) = 415.8$. Since both of the numbers are greater than or equal to 15, the sample size is sufficiently large to conclude the normal approximation is reasonable.

For confidence coefficient .90, $\alpha = .10$ and $\alpha / 2 = .10 / 2 = .05$. From Table III, Appendix A, $z_{.05}$ = 1.645. The 90% confidence interval is:

$$\hat{p} \pm z_{.05}\sqrt{\frac{pq}{n}} \Rightarrow \hat{p} \pm 1.645\sqrt{\frac{\hat{p}\hat{q}}{n}} \Rightarrow .175 \pm 1.645\sqrt{\frac{.175(.825)}{504}} \Rightarrow .175 \pm .028 \Rightarrow (.147,\ .203)$$

We are 90% confident that the true proportion of all ice melt ponds in the Canadian Arctic that have first-year ice is between .147 and .203.

5.64 a. The estimate of the true proportion of all U. S. teenagers who have used at least one informal element in a school writing assignment is $\hat{p} = \frac{52}{2481} = .021$.

The sample size is large enough if both $n\hat{p} \ge 15$ and $n\hat{q} \ge 15$.

$n\hat{p} = 2481(.021) = 52.10$ and $n\hat{q} = 2481(.979) = 2428.90$. Since both of these numbers are greater than or equal to 15, the sample size is sufficiently large to conclude the normal approximation is reasonable.

For confidence coefficient .95, $\alpha = .05$ and $\alpha / 2 = .05 / 2 = .025$. From Table III, Appendix A, $z_{.025}$ = 1.96. The 95% confidence interval is:

$$\hat{p} \pm z_{.025}\sqrt{\frac{pq}{n}} \approx \hat{p} \pm 1.96\sqrt{\frac{\hat{p}\hat{q}}{n}} \Rightarrow .021 \pm 1.96\sqrt{\frac{.021(.979)}{2481}} \Rightarrow .021 \pm .0056 \Rightarrow (.0154,\ .0266)$$

We are 95% confident that the true proportion of all people in the world who suffer from ORS is between .0154 and .0266.

b. The population of interest is all people in the world. The sample was selected from 2,481 university students in Japan. This sample is probably not representative of the population. There are several problems with this sample. First, it is made up of only Japanese people, which may not be representative of all the adults in the world. Next, the age group is very limited, probably in the range of 18 to 24. Finally, those people who attend universities may not be representative of all people. Thus, the inference made from this sample may not be valid for all of the people of the world.

5.66 The Wilson adjusted sample proportion is

$$\tilde{p} = \frac{x+2}{n+4} = \frac{3+2}{13+4} = \frac{5}{17} = .294$$

For confidence coefficient .99, $\alpha = .01$ and $\alpha/2 = .01/2 = .005$. From Table III, Appendix A, $z_{.005}$ = 2.58. The Wilson adjusted 99% confidence interval is:

$$\tilde{p} \pm z_{.005}\sqrt{\frac{\tilde{p}(1-\tilde{p})}{n+4}} \Rightarrow .294 \pm 2.58\sqrt{\frac{.294(.706)}{13+4}} \Rightarrow .294 \pm .285 \Rightarrow (.009,\ \ .579)$$

We are 99% confident that the true proportion of all studies on the treatment of Alzheimer's disease with a Wong score below 18 is between .009 and .579.

5.68 The sampling error, SE, is half the width of the confidence interval.

5.70 The statement "For a fixed confidence level $(1-\alpha)$, increasing the sampling error SE will lead to a smaller n when determining sample size" is True.

5.72 The sample size will be larger than necessary for any p other than .5.

5.74 a. For confidence coefficient .95, $\alpha = .05$ and $\alpha/2 = .05/2 = .025$. From Table III, Appendix A, $z_{.025}$ = 1.96.

The sample size is $n = \frac{\left(z_{\alpha/2}\right)^2 pq}{(SE)^2} = \frac{(1.96)^2(.3)(.7)}{.06^2} = 224.1 \approx 225$

You would need to take n = 225 samples.

b. To compute the needed sample size, use:

$$n = \frac{\left(z_{\alpha/2}\right)^2 pq}{(SE)^2} = \frac{(1.96)^2(.5)(.5)}{.06^2} = 266.8 \approx 267$$

You would need to take n = 267 samples.

5.76 a. To compute the needed sample size, use

$$n = \frac{\left(z_{\alpha/2}\right)^2 \sigma^2}{(SE)^2} \text{ where } \alpha = 1 - .95 = .05 \text{ and } \alpha/2 = .05/2 = .025$$

From Table III, Appendix A, $z_{.025}$ = 1.96. For a width of 4 units, SE = 4/2 = 2.

$$n = \frac{(1.96)^2(12)^2}{2^2} = 138.298 \approx 139$$

You would need to take 139 samples at a cost of 139($10) = $1390. No, you do not have sufficient funds.

b. For confidence coefficient .90, $\alpha = .10$ and $\alpha/2 = .10/2 = .05$. From Table III, Appendix A, $z_{.05}$ = 1.645.

$$n = \frac{(1.645)^2(12)^2}{2^2} = 97.417 \approx 98$$

You would need to take 98 samples at a cost of 98($10) = $980. You now have sufficient funds but have an increased risk of error.

5.78 For confidence coefficient .95, $\alpha = .05$ and $\alpha/2 = .05/2 = .025$. From Table III, Appendix A, $z_{.025}$ = 1.96. Since we have no previous knowledge about the proportion of cul-de-sac homes in home town that were burglarized in the past year, we will use .5 to estimate p.

$$n = \frac{z_{\alpha/2}^2 pq}{(SE)^2} = \frac{1.96^2(.5)(.5)}{.02^2} = 2401$$

We need to find out whether each home in the sample of 2,401 cul-de-sac homes was burglarized in the last year or not.

5.80 a. The confidence level desired by the researchers is .95.

b. The sampling error desired by the researchers is SE = .001.

c. For confidence coefficient .95, $\alpha = .05$ and $\alpha/2 = .05/2 = .025$. From Table III, Appendix A, $z_{.025}$ = 1.96.

The sample size is $n = \frac{(z_{.025})^2\sigma^2}{(SE)^2} = \frac{1.96^2(.005)^2}{.001^2} = 96.04 \approx 97$.

5.82 For confidence coefficient .90, $\alpha = .10$ and $\alpha/2 = .10/2 = .05$. From Table III, Appendix A, $z_{.05}$ = 1.645. Since we have no estimate given for the value of p, we will use .5. The confidence interval is:

$$n = \frac{z_{\alpha/2}^2 pq}{(SE)^2} = \frac{1.645^2 .5(.5)}{.02^2} = 1,691.3 \approx 1,692$$

5.84 a. From Exercise 5.37, s = 129.565. We will use this to estimate the population standard deviation. The necessary sample size is:

$$n = \frac{z_{\alpha/2}^2\sigma^2}{(SE)^2} = \frac{1.96^2(129.656)^2}{45^2} = 31.89 \approx 32$$

b. Answers will vary. Since we want to have the sample spread fairly evenly through out the year, we might want to randomly sample 2 days from each month. That would give us a sample size of 24. Then, we can randomly select 8 more months and randomly select a 3rd day in each of those months.

A simpler method might be to use a sample size of 36 (to be conservative) and randomly select 3 days from each month.

c. Answers will vary. We will use the first plan above. First we randomly selected 8 of the 12 months to sample a third time. Those months were March, August, November, January, October, April, September, and February. Then, we randomly selected 2 or 3 days from each of the months. The data selected were:

Jan	566	573	591
Feb	611	630	650
Mar	688	704	754
Apr	762	771	810
May	886	891	
Jun	907	907	
Jul	904	873	
Aug	829	827	809
Sep	767	748	715
Oct	678	667	651
Nov	624	621	574
Dec	553	552	

Using MINITAB, the descriptive statistics are:

Descriptive Statistics: Daylight

```
Variable   N   Mean  StDev  Minimum     Q1  Median     Q3  Maximum
Daylight  32  721.7  117.2    552.0  621.8   709.5  822.8    907.0
```

d. For confidence coefficient .95, $\alpha = .05$ and $\alpha / 2 = .05 / 2 = .025$. From Table III, Appendix A, $z_{.025}$ = 1.96. The confidence interval is:

$$\bar{x} \pm z_{.025} \frac{s}{\sqrt{n}} \Rightarrow 721.7 \pm 1.96 \frac{117.2}{\sqrt{32}} \Rightarrow 721.7 \pm 40.61 \Rightarrow (681.09,\ 762.31)$$

We have estimated the true mean to within 40.61 minutes, which is less than the 45 desired. Thus, we have met the desired width.

5.86 From Exercise 5.60, our estimate of p is $\hat{p} = .974$. We will use this to estimate the sample size.

For confidence coefficient .95, $\alpha = .05$ and $\alpha / 2 = .05 / 2 = .025$. From Table III, Appendix A, $z_{.025}$ = 1.96.

$$n = \frac{z_{\alpha/2}^2 pq}{(SE)^2} = \frac{1.96^2 (.974)(.026)}{.03^2} = 108.09 \approx 109$$

We would need to sample 109 high school athletes.

5.88 a. The confidence interval might lead to an erroneous inference because the sample size used is probably too small. Recall that the sample size is large enough if both $n\hat{p} \geq 15$ and $n\hat{q} \geq 15$. For this problem, $\hat{p} = \frac{10}{18} = .556$. $n\hat{p} = 18(.556) = 10.008$ and $n\hat{q} = 18(.444) = 7.992$. Neither of these two values is greater than 15. Thus, the sample size is not sufficiently large to conclude the normal approximation is reasonable.

b. For confidence coefficient .95, $\alpha = .05$ and $\alpha / 2 = .05 / 2 = .025$. From Table III, Appendix A, $z_{.025}$ = 1.96. From the previous study, we will use $\hat{p} = \frac{10}{18} = .556$ to estimate p.

$$n = \frac{z_{\alpha/2}^2 pq}{(SE)^2} = \frac{1.96^2(.556)(.444)}{.04^2} = 592.72 \approx 593$$

5.90 $\sigma \approx \frac{\text{Range}}{4} = \frac{180-60}{4} = 30$

For confidence coefficient .95, $\alpha = .05$ and $\alpha / 2 = .05 / 2 = .025$. From Table III, Appendix A, $z_{.025}$ = 1.96.

The sample size is $n = \frac{z_{\alpha/2}^2 \sigma^2}{(SE)^2} = \frac{1.96^2(30)^2}{5^2} = 138.3 \approx 139$

5.92 a. For confidence coefficient .90, $\alpha = .10$ and $\alpha / 2 = .10 / 2 = .05$. From Table III, Appendix A, $z_{.05}$ = 1.645.

The sample size is $n = \frac{z_{\alpha/2}^2 \sigma^2}{(SE)^2} = \frac{1.645^2(2^2)}{.1^2} = 1082.4 \approx 1083$

b. In part **a**, we found n = 1083. If we used an n of only 100, the width of the confidence interval for μ would be wider since we would be dividing by a smaller number.

c. We know $SE = \frac{z_{\alpha/2}\sigma}{\sqrt{n}} \Rightarrow z_{\alpha/2} = \frac{SE\sqrt{n}}{\sigma} = \frac{.1\sqrt{100}}{2} = .5$

$P(-.5 \leq z \leq .5)$ = .1915 + .1915 = .3830. Thus, the level of confidence is approximately 38.3%.

5.94 We must take a random sample from the target population and the distribution of the population must be approximately normal.

5.96 a. $\alpha/2 = .05/2 = .025$; $\chi^2_{.025,7} = 16.0128$ and $\chi^2_{.975,7} = 1.68987$

b. $\alpha/2 = .10/2 = .05$; $\chi^2_{.05,16} = 26.2962$ and $\chi^2_{.95,16} = 7.96164$

c. $\alpha/2 = .01/2 = .005$; $\chi^2_{.005,20} = 39.9968$ and $\chi^2_{.995,20} = 7.43386$

d. $\alpha/2 = .05/2 = .025$; $\chi^2_{.025,20} = 34.1696$ and $\chi^2_{.975,20} = 9.59083$

5.98 To find the 90% confidence interval for σ, we need to take the square root of the end points of the 90% confidence interval for σ^2 from Exercise 5.97.

a. The 90% confidence interval for σ is:

$$\sqrt{4.537} \le \sigma \le \sqrt{8.809} \Rightarrow 2.130 \le \sigma \le 2.968$$

b. The 90% confidence interval for σ is:

$$\sqrt{.00024} \le \sigma \le \sqrt{.00085} \Rightarrow .0155 \le \sigma \le .0292$$

c. The 90% confidence interval for σ is:

$$\sqrt{641.86} \le \sigma \le \sqrt{1{,}809.09} \Rightarrow 25.335 \le \sigma \le 42.533$$

d. The 90% confidence interval for σ is:

$$\sqrt{.94859} \le \sigma \le \sqrt{12.6632} \Rightarrow .97396 \le \sigma \le 3.55854$$

5.100 a. The target parameter is the variance of the WR scores among all drug dealers.

b. For confidence level .99, $\alpha = .01$ and $\alpha/2 = .01/2 = .005$. From Table V, Appendix A, with df = n – 1 = 100 - 1 = 99, $\chi^2_{.005,99} \approx 140.169$ and $\chi^2_{.995,99} \approx 67.3276$. The 99% confidence interval is:

$$\frac{(n-1)s^2}{\chi^2_{.005}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{.995}} \Rightarrow \frac{(100-1)6^2}{140.169} \le \sigma^2 \le \frac{(100-1)6^2}{67.3276} \Rightarrow 25.426 \le \sigma^2 \le 52.935$$

c. This means that in repeated sampling, 99% of all confidence intervals constructed in the same manner will contain the target parameter.

d. We must assume that a random sample was selected from the target population and that the population is approximately normally distributed.

e. We can use the standard deviation rather than the variance to find a reasonable range for the value of the population mean.

f. To find the 99% confidence interval for σ, we need to take the square root of the end points of the 99% confidence interval for σ^2.

$$\sqrt{25.426} \le \sigma \le \sqrt{52.935} \Rightarrow 5.042 \le \sigma \le 7.276$$

We are 99% confident that the true standard deviation of the WR scores of drug dealers is between 5.042 and 7.276.

5.102 For confidence level .90, $\alpha = .10$ and $\alpha / 2 = .10 / 2 = .05$. From Table V, Appendix A, with df = n – 1 = 4-1 = 3, $\chi^2_{.05,3} = 7.81473$ and $\chi^2_{.95,3} = .351846$. The 90% confidence interval is:

$$\frac{(n-1)s^2}{\chi^2_{.05}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{.95}} \Rightarrow \frac{(4-1).13^2}{7.81473} \le \sigma^2 \le \frac{(4-1).13^2}{.351846} \Rightarrow .0065 \le \sigma^2 \le .1441$$

We are 90% confident that the true variance of the peptide scores for alleles of the antigen-produced protein is between .0065 and .1441.

5.104 a. For confidence level .95, $\alpha = .05$ and $\alpha / 2 = .05 / 2 = .025$. From Table V, Appendix A, with df = n – 1 = 18 - 1 = 17, $\chi^2_{.025,17} = 30.1910$ and $\chi^2_{.975,17} = 7.56418$. The 95% confidence interval for the variance is:

$$\frac{(n-1)s^2}{\chi^2_{.025}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{.975}} \Rightarrow \frac{(18-1)6.3^2}{30.1910} \le \sigma^2 \le \frac{(18-1)6.3^2}{7.56418} \Rightarrow 22.349 \le \sigma^2 \le 89.201$$

The 95% confidence interval for the standard deviation is:

$$\sqrt{22.349} \le \sigma \le \sqrt{89.201} \Rightarrow 4.727 \le \sigma \le 9.445$$

b. We are 95% confident that the true standard deviation of the conduction times of the prototype system is between 4.727 and 9.445.

c. No. Since 7 falls in the 95% confidence interval, it is a likely value for the population standard deviation. Thus, we cannot conclude that the true standard deviation is less than 7.

5.106 From Exercise 5.20, s = 11.34.

For confidence level .90, $\alpha = .10$ and $\alpha / 2 = .10 / 2 = .05$. From Table V, Appendix A, with df = n – 1 = 76 - 1 = 75, $\chi^2_{.05,75} \approx 90.5312$ and $\chi^2_{.95,75} \approx 51.7393$. The 90% confidence interval for the variance is:

$$\frac{(n-1)s^2}{\chi^2_{.05}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{.95}} \Rightarrow \frac{(76-1)11.34^2}{90.5312} \le \sigma^2 \le \frac{(76-1)11.34^2}{51.7393} \Rightarrow 106.534 \le \sigma^2 \le 186.409$$

We are 90% confident that the true variance of shell lengths of all green sea turtles in the lagoon is between 106.534 and 186.409.

5.108 Using MINITAB, the descriptive statistics for the two groups are:

Descriptive Statistics: Honey, DM

Variable	N	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Honey	35	10.714	2.855	4.000	9.000	11.000	12.000	16.000
DM	33	8.333	3.256	3.000	6.000	9.000	11.500	15.000

a. For confidence level .90, $\alpha = .10$ and $\alpha / 2 = .10 / 2 = .05$. From Table V, Appendix A, with df = n – 1 = 35 - 1 = 34, $\chi^2_{.05,34} \approx 43.7729$ and $\chi^2_{.95,34} \approx 18.4926$. The 90% confidence interval for the variance is:

$$\frac{(n-1)s^2}{\chi^2_{.05}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{.95}} \Rightarrow \frac{(35-1)2.855^2}{43.7729} \le \sigma^2 \le \frac{(35-1)2.855^2}{18.4926} \Rightarrow 6.331 \le \sigma^2 \le 14.986$$

The 90% confidence interval for the standard deviation is:

$$\sqrt{6.331} \le \sigma \le \sqrt{14.986} \Rightarrow 2.516 \le \sigma \le 3.871$$

We are 90% confident that the true standard deviation for the improvement scores for the honey dosage group is between 2.516 and 3.871.

b. For confidence level .90, $\alpha = .10$ and $\alpha / 2 = .10 / 2 = .05$. From Table V, Appendix A, with df = n – 1 = 33 - 1 = 32, $\chi^2_{.05,32} \approx 43.7729$ and $\chi^2_{.95,32} \approx 18.4926$. The 90% confidence interval for the variance is:

$$\frac{(n-1)s^2}{\chi^2_{.05}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{.95}} \Rightarrow \frac{(33-1)3.256^2}{43.7729} \le \sigma^2 \le \frac{(33-1)3.256^2}{18.4926} \Rightarrow 7.750 \le \sigma^2 \le 18.345$$

The 90% confidence interval for the standard deviation is:

$$\sqrt{7.750} \le \sigma \le \sqrt{18.345} \Rightarrow 2.784 \le \sigma \le 4.283$$

We are 90% confident that the true standard deviation for the improvement scores for the DM dosage group is between 2.784 and 4.283.

c. Since the two confidence intervals constructed in parts **a** and **b** overlap, there is no evidence to indicate either of the two groups has a smaller variation in improvement scores.

5.110 95% confident means that in repeated sampling, 95% of all confidence intervals constructed for the proportion of all PCs with a computer virus will contain the true proportion.

5.112 a. $P(t \le t_0) = .05$ where df = 17
$t_0 = -1.740$

b. $P(t \ge t_0) = .005$ where df = 14
$t_0 = 2.977$

c. $P(t \le -t_0 \text{ or } t \ge t_0) = .10$ where df = 6 is equivalent to
$P(t \ge t_0) = .10/2 = .05$ where df = 6
$t_0 = 1.943$

d. $P(t \le -t_0 \text{ or } t \ge t_0) = .01$ where df = 17 is equivalent to
$P(t \ge t_0) = .01/2 = .005$ where df = 22
$t_0 = 2.819$

5.114 a. For confidence coefficient .99, $\alpha = .01$ and $\alpha/2 = .01/2 = .005$. From Table III, Appendix A, $z_{.005} = 2.58$. The confidence interval is:

$$\bar{x} \pm z_{\alpha/2}\frac{s}{\sqrt{n}} \Rightarrow 32.5 \pm 2.58\frac{30}{\sqrt{225}} \Rightarrow 32.5 \pm 5.16 \Rightarrow (27.34,\ 37.66)$$

b. The sample size is $n = \dfrac{(z_{\alpha/2})^2 \sigma^2}{(SE)^2} = \dfrac{2.575^2(30)^2}{.5^2} = 23{,}870.25 \approx 23{,}871$

c. "99% confidence" means that if repeated samples of size 225 were selected from the population and 99% confidence intervals were constructed for the population mean, then 99% of all the intervals constructed will contain the population mean.

d. For confidence level .99, $\alpha = .01$ and $\alpha/2 = .10/2 = .005$. Using MINITAB with df = $n - 1 = 225 - 1 = 224$, $\chi^2_{.005,224} = 282.268$ and $\chi^2_{.995,224} = 173.238$. The 99% confidence interval for the variance is:

$$\frac{(n-1)s^2}{\chi^2_{.005}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{.995}} \Rightarrow \frac{(225-1)30^2}{282.268} \le \sigma^2 \le \frac{(225-1)30^2}{173.238} \Rightarrow 714.215 \le \sigma^2 \le 1{,}163.717$$

5.116 a. The point estimate for the mean personal network size of all older adults is $\bar{x} = 14.6$.

b. For confidence coefficient .95, $\alpha = .05$ and $\alpha/2 = .05/2 = .025$. From Table III, Appendix A, $z_{.025} = 1.96$. The 95% confidence interval is:

$$\bar{x} \pm z_{.025}\sigma_{\bar{x}} \Rightarrow \bar{x} \pm 1.96\frac{\sigma}{\sqrt{n}} \Rightarrow 14.6 \pm 1.96\frac{9.8}{\sqrt{2{,}819}} \Rightarrow 14.6 \pm .36 \Rightarrow (14.24,\ \ 14.96)$$

c. We are 95% confident that the mean personal network size of all older adults is between 14.24 and 14.96.

d. We must assume that we have a random sample from the target population and that the sample size is sufficiently large.

5.118 Using MINITAB, the descriptive statistics are:

Descriptive Statistics: Commitment

Variable	N	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Commitment	30	79.67	10.25	44.00	76.25	81.50	86.25	97.00

a. The point estimate for the mean charitable commitment of tax-exempt organizations is $\bar{x} = 79.67$.

b. For confidence coefficient .95, $\alpha = .05$ and $\alpha / 2 = .05 / 2 = .025$. From Table III, Appendix A, $z_{.025}$ = 1.96. The confidence interval is:

$$\bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}} \Rightarrow 79.67 \pm 1.96 \frac{10.25}{\sqrt{30}} \Rightarrow 79.67 \pm 3.67 \Rightarrow (76.00, 83.34)$$

c. Since the sample size is at least 30, the Central Limit Theorem applies. We must assume that we have a random sample from the population.

d. The probability of estimating the true mean charitable commitment exactly with a point estimate is zero. By using a range of values to estimate the mean (i.e. confidence interval), we can have a level of confidence that the range of values will contain the true mean.

e. For confidence level .95, $\alpha = .05$ and $\alpha / 2 = .05 / 2 = .025$. From Table V, Appendix A, with df = n – 1 = 30 - 1 = 29, $\chi^2_{.025,29} = 45.7222$ and $\chi^2_{.975,29} = 16.0471$. The 95% confidence interval for the variance is:

$$\frac{(n-1)s^2}{\chi^2_{.025}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{.975}} \Rightarrow \frac{(30-1)10.25^2}{45.7222} \le \sigma^2 \le \frac{(30-1)10.25^2}{16.0471} \Rightarrow 66.637 \le \sigma^2 \le 189.867$$

We are 95% confident that the true variance of all charitable commitments for all tax-exempt organizations is between 66.637 and 189.867.

5.120 a. The population of interest is all shoppers in Muncie, Indiana.

b. The characteristic of interest is the proportion of shoppers who think “Made in the USA” means 100% of US labor and materials.

c. The point estimate for the proportion of shoppers who think “Made in the USA” means 100% of US labor and materials is $\hat{p} = \frac{x}{n} = \frac{64}{106} = .604$.

The sample size is large enough if both $n\hat{p} \geq 15$ and $n\hat{q} \geq 15$.

$n\hat{p} = 106(.604) = 64$ and $n\hat{q} = 106(.396) = 42$. Since both of the numbers are greater than or equal to 15, the sample size is sufficiently large to conclude the normal approximation is reasonable.
For confidence coefficient .90, $\alpha = .10$ and $\alpha/2 = .10/2 = .05$. From Table III, Appendix A, $z_{.05}$ = 1.645. The 90% confidence interval is:

$$\hat{p} \pm z_{.05}\sigma_{\hat{p}} \Rightarrow \hat{p} \pm 1.645\sqrt{\frac{\hat{p}\hat{q}}{n}} \Rightarrow .604 \pm 1.645\sqrt{\frac{.604(.396)}{106}} \Rightarrow .604 \pm .078 \Rightarrow (.526, \ \ .682)$$

d. We are 90% confident that the proportion of shoppers who think "Made in the USA" means 100% of US labor and materials is between .526 and .682.

e. "90% confidence" means that in repeated samples of size 106, 90% of all confidence intervals formed for the proportion of shoppers who think "Made in the USA" means 100% of US labor and materials will contain the true population proportion.

5.122 For confidence coefficient .95, $\alpha = .05$ and $\alpha/2 = .05/2 = .025$. From Table III, Appendix A, $z_{.025}$ = 1.96.

The sample size is $n = \frac{(z_{.025})^2\sigma^2}{(SE)^2} = \frac{1.96^2(5)^2}{1^2} = 96.04 \approx 97$.

5.124 From Exercise 2.182, $\bar{x} = 1.471$ and s = .064.

For confidence coefficient .99, $\alpha = .01$ and $\alpha/2 = .01/2 = .005$. From Table IV, Appendix A with df $= n - 1 = 8 - 1 = 7$, $t_{.005}$ = 3.499. The 99% confidence interval is:

$$\bar{x} \pm t_{\alpha/2}\frac{s}{\sqrt{n}} \Rightarrow 1.471 \pm 3.499\frac{.064}{\sqrt{8}} \Rightarrow 1.471 \pm .079 \Rightarrow (1.392, 1.550)$$

We are 99% confident that the mean daily ammonia level in air in the tunnel is between 1.392 and 1.550. We must assume that the population of ammonia levels is normally distributed.

5.126 a. The point estimate of p is $\hat{p} = \frac{x}{n} = \frac{39}{150} = .26$.

b. The sample size is large enough if both $n\hat{p} \geq 15$ and $n\hat{q} \geq 15$.

$n\hat{p} = 150(.26) = 39$ and $n\hat{q} = 150(.74) = 111$. Since both of the numbers are greater than or equal to 15, the sample size is sufficiently large to conclude the normal approximation is reasonable.

For confidence coefficient .95, $\alpha = .05$ and $\alpha/2 = .05/2 = .025$. From Table III, Appendix A, $z_{.025}$ = 1.96. The confidence interval is:

$$\hat{p} \pm z_{.025}\sqrt{\frac{\hat{p}\hat{q}}{n}} \Rightarrow .26 \pm 1.96\sqrt{\frac{.26(.74)}{150}} \Rightarrow .26 \pm .070 \Rightarrow (.190, .330)$$

c. We are 95% confident that the true proportion of college students who experience "residual anxiety" from a scary TV show or movie is between .190 and .330.

5.128 Some preliminary calculations are:

$$\bar{x} = \frac{\sum x}{n} = \frac{160.9}{22} = 7.314 \qquad s^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1} = \frac{1{,}389.1 - \frac{160.9^2}{22}}{22-1} = 10.1112$$

$$s = \sqrt{10.1112} = 3.180$$

For confidence coefficient .95, $\alpha = .05$ and $\alpha/2 = .05/2 = .025$. From Table IV, Appendix A with df = n – 1 = 22 – 1 = 21, $t_{.025}$ = 2.080. The 95% confidence interval is:

$$\bar{x} \pm t_{\alpha/2}\frac{s}{\sqrt{n}} \Rightarrow 7.314 \pm 2.080\frac{3.180}{\sqrt{22}} \Rightarrow 7.314 \pm 1.410 \Rightarrow (5.904,\ 8.724)$$

We are 95% confident that the mean PMI for all human brain specimens obtained at autopsy is between 5.904 and 8.724.

Since 10 is not in the 95% confidence interval, it is not a likely value for the true mean. We would infer that the true mean PMI for all human brain specimens obtained at autopsy is less than 10 days.

5.130 a. First, we compute $\hat{p}$: $\hat{p} = \frac{x}{n} = \frac{15}{40} = .375$

The sample size is large enough if both $n\hat{p} \geq 15$ and $n\hat{q} \geq 15$.

$n\hat{p} = 40(.375) = 15$ and $n\hat{q} = 40(.625) = 25$. Since both of the numbers are greater than or equal to 15, the sample size is sufficiently large to conclude the normal approximation is reasonable.
For confidence coefficient .90, $\alpha = .10$ and $\alpha/2 = .10/2 = .05$. From Table III, Appendix A, $z_{.05}$ = 1.645. The 90% confidence interval is:

$$\hat{p} \pm z_{.05}\sqrt{\frac{pq}{n}} \Rightarrow \hat{p} \pm 1.645\sqrt{\frac{\hat{p}\hat{q}}{n}} \Rightarrow .375 \pm 1.645\sqrt{\frac{.375(.625)}{40}}$$
$$\Rightarrow .375 \pm .126 \Rightarrow (.249,\ \ .501)$$

We are 90% confident that the true dropout rate for exercisers who vary their routine in workouts is between .249 and .501.

b. First, we compute $\hat{p}$: $\hat{p} = \frac{x}{n} = \frac{23}{40} = .575$

The sample size is large enough if both $n\hat{p} \geq 15$ and $n\hat{q} \geq 15$.
$n\hat{p} = 40(.575) = 23$ and $n\hat{q} = 40(.425) = 17$. Since both of the numbers are greater than or equal to 15, the sample size is sufficiently large to conclude the normal approximation is reasonable.

For confidence coefficient .90, $\alpha = .10$ and $\alpha / 2 = .10 / 2 = .05$. From Table III, Appendix A, $z_{.05}$ = 1.645. The 90% confidence interval is:

$$\hat{p} \pm z_{.05}\sqrt{\frac{pq}{n}} \Rightarrow \hat{p} \pm 1.645\sqrt{\frac{\hat{p}\hat{q}}{n}} \Rightarrow .575 \pm 1.645\sqrt{\frac{.575(.425)}{40}}$$
$$\Rightarrow .575 \pm .129 \Rightarrow (.446,\ \ .704)$$

We are 90% confident that the true dropout rate for exercisers who have no set schedule for their workouts is between .446 and .704.

5.132 a. For confidence coefficient .90, $\alpha = .10$ and $\alpha / 2 = .10 / 2 = .05$. From Table III, Appendix A, $z_{.05}$ = 1.645. The confidence interval is:

$$\bar{x} \pm z_{.05}\frac{s}{\sqrt{n}} \Rightarrow 7.62 \pm 1.645\frac{8.91}{\sqrt{65}} \Rightarrow 7.62 \pm 1.82 \Rightarrow (5.80,\ 9.44)$$

b. We are 90% confident that the mean sentence complexity score of all low-income children is between 5.80 and 9.44.

c. Yes. We are 90% confident that the mean sentence complexity score of all low-income children is between 5.80 and 9.44. Since the mean score for middle-income children, 15.55, is outside this interval, there is evidence that the true mean for low-income children is different from 15.55.

5.134 a. For confidence coefficient .95, $\alpha = .05$ and $\alpha / 2 = .05 / 2 = .025$. From Table IV, Appendix A, with df = n – 1 = 15 – 1 = 14, $t_{.025}$ = 2.145. The confidence interval is:

$$\bar{x} \pm t_{.025}\frac{s}{\sqrt{n}} \Rightarrow 5.87 \pm 2.145\frac{1.51}{\sqrt{15}} \Rightarrow 5.87 \pm .836 \Rightarrow (5.034,\ 6.706)$$

We are 95% confident that the true mean response of the students is between 5.034 and 6.706.

b. In part **a**, the width of the interval is 6.706 − 5.034 = 1.672. The value of *SE* is 1.672/2 = .836. If we want the interval to be half as wide, the value of *SE* would be half that in part **a** or .836/2 = .418. The necessary sample size is:

$$n = \frac{z_{\alpha/2}^2 \sigma^2}{(SE)^2} = \frac{1.96^2 1.51^2}{.418^2} = 50.13 \approx 51$$

5.136 For confidence coefficient .95, $\alpha = .05$ and $\alpha/2 = .05/2 = .025$. From Table III, Appendix A, $z_{.025}$ = 1.96. From Exercise 5.107, a good approximation for p is .094.

The sample size is $n = \dfrac{(z_{\alpha/2})^2 pq}{(SE)^2} = \dfrac{(1.96)^2(.094)(.906)}{.02^2} = 817.9 \approx 818$

You would need to take n = 818 samples.

5.138 The sample size is large enough if both $n\hat{p} \ge 15$ and $n\hat{q} \ge 15$.

$n\hat{p} = 150(.11) = 16.5$ and $n\hat{q} = 150(.89) = 133.5$. Since both of the numbers are greater than or equal to 15, the sample size is sufficiently large to conclude the normal approximation is reasonable.

For confidence coefficient .95, $\alpha = .05$ and $\alpha/2 = .05/2 = .025$. From Table III, Appendix A, $z_{.025}$ = 1.96. The 95% confidence interval is:

$$\hat{p} \pm z_{.025}\sqrt{\frac{pq}{n}} \Rightarrow \hat{p} \pm 1.96\sqrt{\frac{\hat{p}\hat{q}}{n}} \Rightarrow .11 \pm 1.96\sqrt{\frac{.11(.89)}{150}} \Rightarrow .11 \pm .050 \Rightarrow (.06,\ \ .16)$$

We are 95% confident that the true proportion of all MSDS that are satisfactorily completed is between .06 and .16.

Yes. Since 20% or .20 is not contained in the confidence interval, it is not a likely value.

5.140 a. The point estimate of p is $\hat{p} = \dfrac{x}{n} = \dfrac{35}{55} = .636$.

The sample size is large enough if both $n\hat{p} \ge 15$ and $n\hat{q} \ge 15$.

$n\hat{p} = 55(.636) = 35$ and $n\hat{q} = 55(.364) = 20$. Since both of the numbers are greater than or equal to 15, the sample size is sufficiently large to conclude the normal approximation is reasonable.

Since no level of confidence is given, we will use 95% confidence. For confidence coefficient .95, $\alpha = .05$ and $\alpha/2 = .05/2 = .025$. From Table III, Appendix A, $z_{.025}$ = 1.96.

The 95% confidence interval is:

$$\hat{p} \pm z_{.025}\sqrt{\frac{pq}{n}} \Rightarrow \hat{p} \pm 1.96\sqrt{\frac{\hat{p}\hat{q}}{n}} \Rightarrow .636 \pm 1.96\sqrt{\frac{.636(.364)}{55}} \Rightarrow .636 \pm .127 \Rightarrow (.509,\ \ .763)$$

We are 95% confident that the true proportion of all fatal air bag accidents involving children is between .509 and .763.

b. The sample proportion of children killed by air bags who were not wearing seat belts or were improperly restrained is 24/35 = .686. This is a rather large proportion. Whether a child is killed by an air bag could be related to whether or not he/she was properly restrained. Thus, the number of children killed by air bags could possibly be reduced if the child were properly restrained.