Chapter 9

Estimating the Value of a Parameter Using Confidence Intervals

9.1 The Logic in Constructing Confidence Intervals about a Population Mean Where the Population Standard Deviation is Known

1. The margin of error of a confidence interval of a parameter depends on the level of confidence, the sample size, and the standard deviation of the population.

2. The margin of error increases as the level of confidence increases because, if we want to be more confident that the interval contains the population mean, then we need to make the interval wider.

3. The margin of error decreases as the sample size increases because the Law of Large Numbers states that as the sample size increases the sample mean approaches the value of the population mean.

4. The level of confidence refers to the proportion of samples that will result in intervals that capture the unknown parameter.

5. The mean age of the population is a fixed value (i.e., constant), so it is not probabilistic. The 95% level of confidence refers to confidence in the method by which the interval is obtained, not the specific interval. A better interpretation would be: "We are confident that the interval 21.4 years to 28.8 years, obtained by using our method, is one of the 95% of confidence intervals that contains the population mean." The precision of the interval could be increased by increasing the sample size or decreasing the level of confidence.

6. No, it would not make sense to construct a confidence interval for the population mean. Since the population size is so small, the professor could easily obtain the exact value of the population mean.

7. No, a Z-interval should not be constructed because the data are not normal since a point is outside the bounds of the normal probability plot. Also, the data contain an outlier which can be seen in the boxplot.

8. No, a Z-interval should not be constructed because the data are not normal since points are outside the bounds of the normal probability plot. Also, the data contain outliers which can be seen in the boxplot.

9. No, a Z-interval should not be constructed because the data are not normal since points are outside the bounds of the normal probability plot. From the boxplot, the data appear to be skewed right.

10. No, a Z-interval should not be constructed because the data are not normal since points are outside the bounds of the normal probability plot. From the box plot, the data appear to be skewed left.

11. Yes, a Z-interval can be constructed. The plotted points are all within the bounds of the normal probability plot, which also has a generally linear pattern. The boxplot shows that there are no outliers.

12. Yes, a Z-interval can be constructed. The plotted points are all within the bounds of the normal probability plot, which also has a generally linear pattern. The boxplot shows that there are no outliers.

13. For a 98% confidence interval, we use $\alpha = 1 - 0.98 = 0.02$, so $z_{\alpha/2} = z_{0.01}$ which is the z-score with area 0.99 below it. The closest area in Table V to 0.9900 is 0.9901 corresponding to $z_{0.01} = 2.33$

14. For a 94% confidence interval we use $\alpha = 1 - 0.94 = 0.06$, so $z_{\alpha/2} = z_{0.03}$ which is the z-score with area 0.97 below it. The closest area in Table V to 0.9700 is 0.9699 corresponding to $z_{0.03} = 1.88$.

15. For a 85% confidence interval we use $\alpha = 1 - 0.85 = 0.15$, so $z_{\alpha/2} = z_{0.075}$ which is the z-score with area 0.9250 below it. The closest area in Table V to 0.9250 is 0.9251 corresponding to $z_{0.075} = 1.44$.

16. For a 80% confidence interval we use $\alpha = 1 - 0.80 = 0.20$, so $z_{\alpha/2} = z_{0.10}$ which is the z-score with area 0.90 below it. The closest area in Table V to 0.9000 is 0.8997 corresponding to $z_{0.10} = 1.28$.

17. The point estimate of the population mean is the midpoint of the confidence interval. We obtain this value by averaging the lower and upper bounds. The margin of error is half the width of the confidence interval. We find this value by subtracting the lower bound from the upper bound and dividing the result by 2.

$$\bar{x} = \frac{18+24}{2} = \frac{42}{2} = 21$$

$$E = \frac{24-18}{2} = \frac{6}{2} = 3$$

18. The point estimate of the population mean is the midpoint of the confidence interval. We obtain this value by averaging the lower and upper bounds. The margin of error is half the width of the confidence interval. We find this value by subtracting the lower bound from the upper bound and dividing the result by 2.

$$\bar{x} = \frac{20+30}{2} = \frac{50}{2} = 25$$

$$E = \frac{30-20}{2} = \frac{10}{2} = 5$$

19. The point estimate of the population mean is the midpoint of the confidence interval. We obtain this value by averaging the lower and upper bounds. The margin of error is half the width of the confidence interval. We find this value by subtracting the lower bound from the upper bound and dividing the result by 2.

$$\bar{x} = \frac{5+23}{2} = \frac{28}{2} = 14$$

$$E = \frac{23-5}{2} = \frac{18}{2} = 9$$

20. The point estimate of the population mean is the midpoint of the confidence interval. We obtain this value by averaging the lower and upper bounds. The margin of error is half the width of the confidence interval. We find this value by subtracting the lower bound from the upper bound and dividing the result by 2.

$$\bar{x} = \frac{15+35}{2} = \frac{50}{2} = 25$$

$$E = \frac{35-15}{2} = \frac{20}{2} = 10$$

21. **(a)** For 95% confidence the critical value is $z_{0.025} = 1.96$. Then:

$$\text{Lower bound} = \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} = 34.2 - 1.96 \cdot \frac{5.3}{\sqrt{35}} \approx 34.2 - 1.76 = 32.44$$

$$\text{Upper bound} = \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} = 34.2 + 1.96 \cdot \frac{5.3}{\sqrt{35}} \approx 34.2 + 1.76 = 35.96$$

(b) $$\text{Lower bound} = \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} = 34.2 - 1.96 \cdot \frac{5.3}{\sqrt{50}} \approx 34.2 - 1.47 = 32.73$$

$$\text{Upper bound} = \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} = 34.2 + 1.96 \cdot \frac{5.3}{\sqrt{50}} \approx 34.2 + 1.47 = 35.67$$

Increasing the sample size decreases the margin of error.

(c) For 99% confidence the critical value is $z_{0.005} = 2.575$. Then:

$$\text{Lower bound} = \bar{x} - z_{0.005} \cdot \frac{\sigma}{\sqrt{n}} = 34.2 - 2.575 \cdot \frac{5.3}{\sqrt{35}} \approx 34.2 - 2.31 = 31.89$$

$$\text{Upper bound} = \bar{x} + z_{0.005} \cdot \frac{\sigma}{\sqrt{n}} = 34.2 + 2.575 \cdot \frac{5.3}{\sqrt{35}} \approx 34.2 + 2.31 = 36.51$$

Increasing the level of confidence increases the margin of error.

(d) Since a sample size of $n = 15$ is less than 30, we can only compute a confidence interval in this way if the population from which we are sampling is normal.

22. **(a)** For 90% confidence the critical value is $z_{0.05} = 1.645$. Then:

$$\text{Lower bound} = \overline{x} - z_{0.05} \cdot \frac{\sigma}{\sqrt{n}} = 59.2 - 1.645 \cdot \frac{3.8}{\sqrt{45}} \approx 59.2 - 0.93 = 58.27$$

$$\text{Upper bound} = \overline{x} + z_{0.05} \cdot \frac{\sigma}{\sqrt{n}} = 59.2 + 1.645 \cdot \frac{3.8}{\sqrt{45}} \approx 59.2 + 0.93 = 60.13$$

(b) $$\text{Lower bound} = \overline{x} - z_{0.05} \cdot \frac{\sigma}{\sqrt{n}} = 59.2 - 1.645 \cdot \frac{3.8}{\sqrt{55}} \approx 59.2 - 0.84 = 58.36$$

$$\text{Upper bound} = \overline{x} + z_{0.05} \cdot \frac{\sigma}{\sqrt{n}} = 59.2 + 1.645 \cdot \frac{3.8}{\sqrt{55}} \approx 59.2 + 0.84 = 60.04$$

Increasing the sample size decreases the margin of error.

(c) For 98% confidence the critical value is $z_{0.01} = 2.33$. Then:

$$\text{Lower bound} = \overline{x} - z_{0.01} \cdot \frac{\sigma}{\sqrt{n}} = 59.2 - 2.33 \cdot \frac{3.8}{\sqrt{45}} \approx 59.2 - 1.32 = 57.88$$

$$\text{Upper bound} = \overline{x} + z_{0.01} \cdot \frac{\sigma}{\sqrt{n}} = 59.2 + 2.33 \cdot \frac{3.8}{\sqrt{45}} \approx 59.2 + 1.32 = 60.52$$

Increasing the level of confidence increases the margin of error.

(d) Since a sample size of $n = 15$ is less than 30, we can only compute a confidence interval in this way if the population from which we are sampling is normal.

23. **(a)** For 96% confidence the critical value is $z_{0.02} = 2.05$. Then:

$$\text{Lower bound} = \overline{x} - z_{0.02} \cdot \frac{\sigma}{\sqrt{n}} = 108 - 2.05 \cdot \frac{13}{\sqrt{25}} \approx 108 - 5.3 = 102.7$$

$$\text{Upper bound} = \overline{x} + z_{0.02} \cdot \frac{\sigma}{\sqrt{n}} = 108 + 2.05 \cdot \frac{13}{\sqrt{25}} \approx 108 + 5.3 = 113.3$$

(b) $$\text{Lower bound} = \overline{x} - z_{0.02} \cdot \frac{\sigma}{\sqrt{n}} = 108 - 2.05 \cdot \frac{13}{\sqrt{10}} \approx 108 - 8.4 = 99.6$$

$$\text{Upper bound} = \overline{x} + z_{0.02} \cdot \frac{\sigma}{\sqrt{n}} = 108 + 2.05 \cdot \frac{13}{\sqrt{10}} \approx 108 + 8.4 = 116.4$$

Decreasing the sample size increases the margin of error.

(c) For 88% confidence the critical value is $z_{0.06} = 1.555$. Then:

$$\text{Lower bound} = \overline{x} - z_{0.06} \cdot \frac{\sigma}{\sqrt{n}} = 108 - 1.555 \cdot \frac{13}{\sqrt{25}} \approx 108 - 4.0 = 104.0$$

$$\text{Upper bound} = \overline{x} + z_{0.06} \cdot \frac{\sigma}{\sqrt{n}} = 108 + 1.555 \cdot \frac{13}{\sqrt{25}} \approx 108 + 4.0 = 112.0$$

Decreasing the level of confidence decreases the margin of error.

(d) No. Each sample size is too small to insure the $\overline{x}$ sampling distribution is normal.

(e) The outliers would have increased the sample mean, shifting the confidence interval to the right. If there are outliers then we should not use this approach to compute a confidence interval.

24. **(a)** For 94% confidence the critical value is $z_{0.03} = 1.88$. Then:

$$\text{Lower bound} = \bar{x} - z_{0.03} \cdot \frac{\sigma}{\sqrt{n}} = 123 - 1.88 \cdot \frac{17}{\sqrt{20}} \approx 123 - 7.1 = 115.9$$

$$\text{Upper bound} = \bar{x} + z_{0.03} \cdot \frac{\sigma}{\sqrt{n}} = 123 + 1.88 \cdot \frac{17}{\sqrt{20}} \approx 123 + 7.1 = 130.1$$

(b) $$\text{Lower bound} = \bar{x} - z_{0.03} \cdot \frac{\sigma}{\sqrt{n}} = 123 - 1.88 \cdot \frac{17}{\sqrt{12}} \approx 123 - 9.2 = 113.8$$

$$\text{Upper bound} = \bar{x} + z_{0.03} \cdot \frac{\sigma}{\sqrt{n}} = 123 + 1.88 \cdot \frac{17}{\sqrt{12}} \approx 123 + 9.2 = 132.2$$

Decreasing the sample size increases the margin of error.

(c) For 85% confidence the critical value is $z_{0.075} = 1.44$. Then:

$$\text{Lower bound} = \bar{x} - z_{0.075} \cdot \frac{\sigma}{\sqrt{n}} = 123 - 1.44 \cdot \frac{17}{\sqrt{20}} \approx 123 - 5.5 = 117.5$$

$$\text{Upper bound} = \bar{x} + z_{0.075} \cdot \frac{\sigma}{\sqrt{n}} = 123 + 1.44 \cdot \frac{17}{\sqrt{20}} \approx 123 + 5.5 = 128.5$$

Decreasing the level of confidence decreases the margin of error.

(d) No. Each sample size is too small to insure the $\bar{x}$ sampling distribution is normal.

(e) The outlier would have increased the mean, shifting the confidence interval to the right. If there are outliers then we should not use this approach to compute a confidence interval.

25. **(a)** Flawed; this interpretation implies that the population mean varies rather than the interval. For a given population, the mean number of hours worked is fixed, though typically unknown.

(b) This is the correct interpretation.

(c) Flawed; this interpretation makes an implication statement about individual values rather than the mean.

(d) Flawed; since the sample was of adult Americans, the interpretation should be about the mean number of hours worked by adult Americans, not just about adults in Idaho.

26. **(a)** Flawed; this interpretation makes an implication statement about individual values rather than the mean.

(b) Flawed; since the confidence interval was for the mean number of hours of sleep during a *weekday*, the interpretation should be about the mean number of hours of sleep during a weekday, not any day of the week.

(c) Flawed; this interpretation implies that the population mean varies rather than the interval. For a given population, the mean number of hours worked is fixed, though typically unknown.

(d) This is the correct interpretation.

27. Based on the sample data, we are 90% confident that the mean drive-through service time for Taco Bell restaurants is between 161.5 seconds and 164.7 seconds. We are confident that the sample drawn is one of the 90% of the same size that would produce an interval containing the true mean.

28. Based on the sample data, we are 95% confident that the population mean time spent per month per user on MySpace.com is between 151.4 minutes and 190.6 minutes. We are confident that the sample drawn is one of the 95% of the same size that would produce an interval containing the true mean.

29. The precision of the interval could be increased by either increasing the sample size, or lowering the confidence level.

30. The precision of the interval could be increased by either increasing the sample size, or lowering the confidence level.

31. **(a)** For 90% confidence the critical value is $z_{0.05} = 1.645$. Then:

$$\text{Lower bound } = \bar{x} - z_{0.05} \cdot \frac{\sigma}{\sqrt{n}} = 0.16 - 1.645 \cdot \frac{0.08}{\sqrt{1200}} \approx 0.16 - 0.004 = 0.156 \text{ g/dL}$$

$$\text{Upper bound } = \bar{x} + z_{0.05} \cdot \frac{\sigma}{\sqrt{n}} = 0.16 + 1.645 \cdot \frac{0.08}{\sqrt{1200}} \approx 0.16 + 0.004 = 0.164 \text{ g/dL}$$

The researcher is 90% confident that the population mean BAC is between 0.156 and 0.164 g/dL for drivers involved in fatal accidents who have a positive BAC value.

(b) Yes, it is possible that the true mean BAC is not captured in the interval from part (a). Using 90% confidence, only 90% of all possible samples of the same size from the population will produce intervals containing the true value of the mean. The interval from part (a) could be one of the 10% that do not capture the true mean.

32. **(a)** For 90% confidence the critical value is $z_{0.05} = 1.645$. Then:

$$\text{Lower bound } = \bar{x} - z_{0.05} \cdot \frac{\sigma}{\sqrt{n}} = 1.23 - 1.645 \cdot \frac{0.65}{\sqrt{1120}} \approx 1.23 - 0.032 = 1.198 \text{ hours}$$

$$\text{Upper bound } = \bar{x} + z_{0.05} \cdot \frac{\sigma}{\sqrt{n}} = 1.23 + 1.645 \cdot \frac{0.65}{\sqrt{1120}} \approx 1.23 + 0.032 = 1.262 \text{ hours}$$

The Bureau of Labor Statistics is 90% confident that the population mean amount of time that Americans aged 15 and older spend eating or drinking each day is between 1.198 and 1.262 hours.

(b) Yes, it is possible that the true mean time spent eating and drinking is not captured in the interval from part (a). Using 90% confidence, only 90% of all possible samples of the same size from the population will produce intervals containing the true value of the mean. The interval from part (a) could be one of the 10% that do not capture the true mean.

33. **(a)** For 95% confidence the critical value is $z_{0.025} = 1.96$. Then:

$$\text{Lower bound } = \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} = 4.113 - 1.96 \cdot \frac{0.110}{\sqrt{900}} \approx 4.113 - 0.0072 = 4.1058$$

$$\text{Upper bound } = \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} = 4.113 + 1.96 \cdot \frac{0.110}{\sqrt{900}} \approx 4.113 + 0.0072 = 4.1202$$

The EIA is 95% confident that the population mean price per gallon for regular-grade gasoline on July 14, 2008 was between \$4.1058 and \$4.1202.

(b) No; the interval from part (a) is about the mean price for gasoline in the entire nation. A different result in a specific town does not make the interval inaccurate.

34. **(a)** For 95% confidence the critical value is $z_{0.025} = 1.96$. Then:

$$\text{Lower bound } = \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} = 24.2 - 1.96 \cdot \frac{18.5}{\sqrt{50}} \approx 24.2 - 5.13 = 19.07 \text{ minutes}$$

$$\text{Upper bound } = \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} = 24.2 + 1.96 \cdot \frac{18.5}{\sqrt{50}} \approx 24.2 + 5.13 = 29.33 \text{ minutes}$$

The economist is 95% confident that the population mean travel time to work is between 19.07 and 29.33 minutes.

(b) No; the interval from part (a) is about the mean travel time to work nationally. The mean travel time to work in a specific town may differ.

35. (a) $\bar{x} = \frac{1446 + 743 + 1581 + \ldots + 995}{12}$

$= \frac{14{,}451}{12} \approx \$1204.3$

(b) Yes, the conditions for a Z-interval are satisfied. The plotted points are all within the bounds of the normal probability plot, which also has a generally linear pattern. The boxplot shows that there are no outliers.

(c) For 95% confidence the critical value is $z_{0.025} = 1.96$. Then:

$$\begin{aligned} \text{Lower bound } &= \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 1204.3 - 1.96 \cdot \frac{450}{\sqrt{12}} \\ &\approx 1204.3 - 254.6 \\ &= \$949.7 \quad \text{[Tech: \$949.6]} \end{aligned}$$

$$\begin{aligned} \text{Upper bound } &= \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 1204.3 + 1.96 \cdot \frac{450}{\sqrt{12}} \\ &\approx 1204.3 + 254.6 \\ &= \$1458.9 \end{aligned}$$

The IIHS is 95% confident that the population mean cost of repairs is between \$949.7 and \$1458.9.

(d) For 90% confidence the critical value is $z_{0.05} = 1.645$. Then:

$$\begin{aligned} \text{Lower bound } &= \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 1204.3 - 1.645 \cdot \frac{450}{\sqrt{12}} \\ &\approx 1204.3 - 213.7 \\ &= \$990.6 \end{aligned}$$

$$\begin{aligned} \text{Upper bound } &= \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 1204.3 + 1.645 \cdot \frac{450}{\sqrt{12}} \\ &\approx 1204.3 + 213.7 \\ &= \$1418.0 \quad \text{[Tech: \$1417.9]} \end{aligned}$$

The IIHS is 90% confident that the population mean cost of repairs is between \$990.6 and \$1418.0.

(e) As the confidence level decreased, the interval width decreased. This is reasonable because, for a fixed standard error, we are less confident that a smaller interval will capture the true mean.

36. (a) $\bar{x} = \frac{4.16 + 4.02 + \ldots + 2.96 + 4.06}{20}$

$= \frac{74.54}{20} = \$3.727$

(b) Yes, the conditions for a Z-interval are satisfied. The plotted points are all within the bounds of the normal probability plot, which also has a generally linear pattern. The boxplot shows that there are no outliers.

(c) For 95% confidence the critical value is $z_{0.025} = 1.96$. Then:

$$\begin{aligned} \text{Lower bound } &= \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 3.727 - 1.96 \cdot \frac{0.46}{\sqrt{20}} \\ &\approx 3.727 - 0.202 \\ &= \$3.525 \end{aligned}$$

$$\begin{aligned} \text{Upper bound } &= \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 3.727 + 1.96 \cdot \frac{0.46}{\sqrt{20}} \\ &\approx 3.727 + 0.202 \\ &= \$3.929 \end{aligned}$$

The economist is 95% confident that the population mean price of reduced fat milk is between \$3.525 and \$3.929.

(d) For 90% confidence the critical value is $z_{0.05} = 1.645$. Then:

$$\begin{aligned} \text{Lower bound } &= \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 3.727 - 1.645 \cdot \frac{0.46}{\sqrt{20}} \\ &\approx 3.727 - 0.169 \\ &= \$3.558 \end{aligned}$$

$$\begin{aligned} \text{Upper bound } &= \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 3.727 + 1.645 \cdot \frac{0.46}{\sqrt{20}} \\ &\approx 3.727 + 0.169 \\ &= \$3.896 \end{aligned}$$

The economist is 90% confident that the population mean price of reduced fat milk is between \$3.558 and \$3.896.

(e) As the confidence level decreased, the interval width decreased. This is reasonable because, for a fixed standard error, we are less confident that a smaller interval will capture the true mean.

37. (a) $\bar{x} = \frac{\sum x}{20} = \frac{829}{20} \approx 41.5$ years

(b) Yes. All the data values lie within the bounds on the normal probability plot, indicating that the data could come from a population that is normal. The boxplot does not show any outliers.

(c) For 95% confidence the critical value is $z_{0.025} = 1.96$. Then:

$$\begin{aligned}\text{Lower bound } &= \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 41.5 - 1.96 \cdot \frac{8.7}{\sqrt{20}} \\ &\approx 41.5 - 3.8 \\ &= 37.7 \text{ years [Tech: 37.6 yrs]}\end{aligned}$$

$$\begin{aligned}\text{Upper bound } &= \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 41.5 + 1.96 \cdot \frac{8.7}{\sqrt{20}} \\ &\approx 41.5 + 3.8 = 45.3 \text{ years}\end{aligned}$$

The agent is 95% confident that the population mean age of buyers in his area who purchase investment property is between 37.7 and 45.3 years.

(d) No, the real estate agent's clients do not appear to differ in age from the general population since the mean age, 39, reported by the National Association of Realtors is contained within the confidence interval.

38. (a) $\bar{x} = \frac{\sum x}{20} = \frac{984}{20} = 49.2$ years

(b) Yes. All the data values lie within the bounds on the normal probability plot, indicating that the data could come from a population that is normal. The boxplot does not show any outliers.

(c) For 95% confidence the critical value is $z_{0.025} = 1.96$. Then:

$$\begin{aligned}\text{Lower bound } &= \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 49.2 - 1.96 \cdot \frac{9.4}{\sqrt{20}} \\ &\approx 49.2 - 4.1 = 45.1 \text{ years}\end{aligned}$$

$$\begin{aligned}\text{Upper bound } &= \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 49.2 + 1.96 \cdot \frac{9.4}{\sqrt{20}} \\ &\approx 49.2 + 4.1 = 53.3 \text{ years}\end{aligned}$$

The agent is 95% confident that the population mean age of buyers in her area who purchase a vacation home is between 45.1 and 53.3 years.

(d) Yes, the buyers in the real estate agent's area who purchase a vacation home appear to differ in age from the general population since the mean age, 44, reported by the National Association of Realtors is not contained within the confidence interval.

39. (a) Since the length of dramas (i.e. the population) is not normally distributed, the sample must be large so that the $\bar{x}$ distribution will be approximately normal.

(b) For 99% confidence the critical value is $z_{0.005} = 2.575$. Then:

$$\begin{aligned}\text{Lower bound } &= \bar{x} - z_{0.05} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 138.3 - 2.575 \cdot \frac{27.3}{\sqrt{30}} \\ &\approx 138.3 - 12.83 \\ &= 125.47 \text{ minutes}\end{aligned}$$

$$\begin{aligned}\text{Upper bound } &= \bar{x} + z_{0.05} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 138.3 + 2.575 \cdot \frac{27.3}{\sqrt{30}} \\ &\approx 138.3 + 12.83 \\ &= 151.13 \text{ minutes}\end{aligned}$$

The student is 99% confident that the population mean length of a drama is between 125.47 and 151.13 minutes.

40. Since the population mean age of the presidents is known (census data), it makes no sense to find a confidence interval.

41. (a) $\bar{x} = \frac{\sum x}{34} = \frac{1{,}589{,}283}{34} \approx 46{,}743.6$ miles

(b) The distribution is skewed right.

(c) Because the population is not normally distributed, a large sample is needed to apply the Central Limit Theorem and say that the $\bar{x}$ distribution is approximately normal.

(d) For 99% confidence the critical value is $z_{0.025} = 2.575$. Then:

$$\begin{aligned}\text{Lower bound } &= \bar{x} - z_{0.005} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 46{,}743.6 - 2.575 \cdot \frac{19{,}000}{\sqrt{34}} \\ &\approx 46{,}743.6 - 8390.6 \\ &= 38{,}353.0 \text{ miles} \\ &[\text{Tech: } 38{,}350.3]\end{aligned}$$

$$\begin{aligned}\text{Upper bound } &= \bar{x} + z_{0.005} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 46{,}743.6 + 2.575 \cdot \frac{19{,}000}{\sqrt{34}} \\ &\approx 46{,}743.6 + 8390.6 \\ &= 55{,}134.2 \text{ miles} \\ &[\text{Tech: } 55{,}136.9]\end{aligned}$$

The used-car dealer is 99% confident that the mean number of miles on a 4-year-old Hummer H2 is between 38,353.0 and 55,134.2.

(e) For 95% confidence the critical value is $z_{0.025} = 1.96$. Then:

$$\begin{aligned}\text{Lower bound } &= \bar{x} - z_{0.005} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 46{,}743.6 - 1.96 \cdot \frac{19{,}000}{\sqrt{34}} \\ &\approx 46{,}743.6 - 6386.6 \\ &= 40{,}357.0 \text{ miles} \\ &[\text{Tech: } 40{,}357.1]\end{aligned}$$

$$\begin{aligned}\text{Upper bound } &= \bar{x} + z_{0.005} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 46{,}743.6 + 1.96 \cdot \frac{19{,}000}{\sqrt{34}} \\ &\approx 46{,}743.6 + 6386.6 \\ &= 53{,}130.2 \text{ miles} \\ &[\text{Tech: } 53{,}130.1]\end{aligned}$$

The used-car dealer is 95% confident that the mean number of miles on a 4-year-old Hummer H2 is between 40,357.0 and 53,130.2.

(f) Decreasing the level of confidence decreases the width of the interval.

(g) No; the used-car dealer only sampled Hummers in the Midwest. Therefore, her results cannot be generalized to the whole United States.

42. (a) $\bar{x} = \frac{\sum x}{40} = \frac{5336}{40} = 133.4$ minutes

(b) The distribution is skewed right.

(c) Because the population is not normally distributed, a large sample is needed to apply the Central Limit Theorem and say that the $\bar{x}$ distribution is approximately normal.

(d) For 99% confidence the critical value is $z_{0.025} = 2.575$. Then:

$$\begin{aligned}\text{Lower bound } &= \bar{x} - z_{0.005} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 133.4 - 2.575 \cdot \frac{45}{\sqrt{40}} \\ &\approx 133.4 - 18.3 \\ &= 115.1 \text{ minutes}\end{aligned}$$

$$\begin{aligned}\text{Upper bound } &= \bar{x} + z_{0.005} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 133.4 + 2.575 \cdot \frac{45}{\sqrt{40}} \\ &\approx 133.4 + 18.3 \\ &= 151.7 \text{ minutes}\end{aligned}$$

The tennis enthusiast is 99% confident that the population mean length of men's singles matches during Wimbledon is between 115.1 and 151.7 minutes.

(e) For 95% confidence the critical value is $z_{0.025} = 1.96$. Then:

$$\begin{aligned}\text{Lower bound } &= \bar{x} - z_{0.005} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 133.4 - 1.96 \cdot \frac{45}{\sqrt{40}} \\ &\approx 133.4 - 13.9 \\ &= 119.5 \text{ minutes}\end{aligned}$$

$$\begin{aligned}\text{Upper bound } &= \bar{x} + z_{0.005} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 133.4 + 1.96 \cdot \frac{45}{\sqrt{40}} \\ &\approx 133.4 + 13.9 \\ &= 147.3 \text{ minutes}\end{aligned}$$

The tennis enthusiast is 95% confident that the population mean length of men's singles matches during Wimbledon is between 119.5 and 147.3 minutes.

(f) Decreasing the level of confidence decreases the width of the interval.

(g) No; the tennis enthusiast only sampled Wimbledon matches. Therefore, his results cannot be generalized to all professional tournaments.

43. For 99% confidence, we use $z_{0.005} = 2.575$. So,

$$n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E}\right)^2 = \left(\frac{2.575 \cdot 13.4}{2}\right)^2 \approx 297.65,$$

which we must increase to 298 subjects.

For 95% confidence we use $z_{0.025} = 1.96$. So,

$n = \left(\frac{1.96 \cdot 13.4}{2}\right)^2 \approx 172.45$, which we must increase to 173 subjects.

Decreasing the level of confidence decreases the required sample size.

44. For 90% confidence, we use $z_{0.05} = 1.645$. So,

$$n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E}\right)^2 = \left(\frac{1.645 \cdot 12.5}{1.5}\right)^2 \approx 187.92,$$

which we must increase to 188 subjects.

For 98% confidence we use $z_{0.01} = 2.33$. So,

$n = \left(\frac{2.33 \cdot 12.5}{1.5}\right)^2 \approx 377.01$ which we must increase to 378 subjects.

Increasing the level of confidence increases the required sample size.

45. For 95% confidence, we use $z_{0.025} = 1.96$. So,

$$n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E}\right)^2 = \left(\frac{1.96 \cdot 16.6}{1}\right)^2 \approx 1058.59,$$

which we must increase to 1059 subjects.

46. For 95% confidence, we use $z_{0.025} = 1.96$. So,

$$n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E}\right)^2 = \left(\frac{1.96 \cdot 82.25}{5}\right)^2 \approx 1039.55,$$

which we must increase to 1040 subjects.

47. (a) For 90% confidence, we use $z_{0.05} = 1.645$. So,

$$n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E}\right)^2 = \left(\frac{1.645 \cdot 19,000}{2000}\right)^2, \approx 244.22$$

which we must increase to 245 vehicles.

(b) For 90% confidence, we use $z_{0.05} = 1.645$. So,

$$n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E}\right)^2 = \left(\frac{1.645 \cdot 19,000}{1000}\right)^2, \approx 976.88$$

which we must increase to 977 vehicles.

(c) Doubling the required accuracy (that is, cutting the margin of error in half) will approximately quadruple the require sample size. This increase is expected because the sample size is inversely proportional to the square of the error. To half the error we must increase the sample size by a factor of $\left(\frac{1}{1/2}\right)^2 = 2^2 = 4$.

48. (a) For 98% confidence, we use $z_{0.01} = 2.33$.

So, $n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E}\right)^2 = \left(\frac{2.33 \cdot 45}{10}\right)^2 \approx 109.94$

which we must increase to 110 matches.

(b) For 98% confidence, we use $z_{0.01} = 2.33$.

So, $n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E}\right)^2 = \left(\frac{2.33 \cdot 45}{5}\right)^2 \approx 439.74$

which we must increase to 440 matches.

(c) Doubling the required accuracy (that is, cutting the margin of error in half) will approximately quadruple the require sample size. This increase is expected because the sample size is inversely proportional to the square of the error. To half the error we must increase the sample size by a factor of $\left(\frac{1}{1/2}\right)^2 = 2^2 = 4$.

49. (a), (b) Answers will vary.

(c) 95% of 20 is $0.95(20) = 19$. We would expect about 19 of the 20 samples to generate confidence intervals that include the population mean. The actual results will vary.

50. **(a), (b)** Answers will vary.

(c) 90% of 30 is $0.90(30) = 27$. We would expect about 27 of the 30 samples to generate confidence intervals that include the population mean. The actual results will vary.

51. **(a), (b)** Answers will vary.

(c) If these were truly "95% confidence intervals," then we would expect approximately 95% of the 100 samples, or 95 samples, to generate confidence intervals that include the population mean. Actual results will vary.

(d) Since the sample size, $n = 6$, is small and since we are sampling from a non-normal population, the sampling distribution of the sample mean is not normal, so our method for computing a 95% confidence interval is not valid. In other words, it is not true that close to 95% of our intervals contain the population mean.

52. **(a)** $\bar{x} = \frac{\sum x}{n} = \frac{603}{12} \approx 50.3$. For a 95% confidence interval we use $z_{.025} = 1.96$.

$$\begin{aligned}\text{Lower bound } &= \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 50.3 - 1.96 \cdot \frac{10}{\sqrt{12}} \\ &\approx 50.3 - 5.7 = 44.6\end{aligned}$$

$$\begin{aligned}\text{Upper bound } &= \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 50.3 + 1.96 \cdot \frac{10}{\sqrt{12}} \\ &\approx 50.3 + 5.7 = 56.0\end{aligned}$$

[Tech: 55.9]

(b) We sort the data (including the incorrectly entered observation) in order and find that $Q_1 = 43.5$ and $Q_3 = 53$. Then $IQR = 53 - 43.5 = 9.5$, and the lower fence is given by $Q_1 - 1.5 \cdot IQR = 43.5 - 1.5(9.5) = 29.25$. Since 14 is below this value, it is an outlier.

(c) $\bar{x} = \frac{\sum x}{n} = \frac{576}{12} = 48$. We now get:

$$\begin{aligned}\text{Lower bound } &= \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 48 - 1.96 \cdot \frac{10}{\sqrt{12}} \\ &\approx 48 - 5.7 = 42.3\end{aligned}$$

$$\begin{aligned}\text{Upper bound } &= \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 48 + 1.96 \cdot \frac{10}{\sqrt{12}} \\ &\approx 48 + 5.7 = 53.7\end{aligned}$$

The outlier shifts the sample mean to the left and, thus, shifts the confidence interval to the left by the same amount. This decreases the likelihood that such a confidence interval will contain the true population mean.

(d) $\bar{x} = \frac{\sum x}{n} = \frac{1809}{36} \approx 50.3$, which is the same mean as for the smaller sample.

(e) For the confidence interval based on this larger sample we get:

$$\begin{aligned}\text{Lower bound } &= \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 50.3 - 1.96 \cdot \frac{10}{\sqrt{36}} \\ &\approx 50.3 - 3.3 = 47.0\end{aligned}$$

$$\begin{aligned}\text{Upper bound } &= \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 50.3 + 1.96 \cdot \frac{10}{\sqrt{36}} \\ &\approx 50.3 + 3.3 = 53.6\end{aligned}$$

[Tech: 53.5]

The width of this confidence interval is $53.6 - 47.0 = 6.6$ compared to the confidence interval in (a) which has a width of $56.0 - 44.6 = 11.4$. Increasing the sample size decreases the width of the confidence interval.

(f) We sort the data (including the incorrectly entered observation) in ascending order and find that $Q_1 = 43$ and $Q_3 = 56$. Then $IQR = 56 - 43 = 13$, and the lower fence is given by $Q_1 - 1.5 \cdot IQR = 43 - 1.5(13) = 23.5$. Since 14 is below this value, it is an outlier.

(g) The new sample mean is
$\bar{x} = \frac{\sum x}{n} = \frac{1782}{36} = 49.5$. We now get:

$$\text{Lower bound } = \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} = 49.5 - 1.96 \cdot \frac{10}{\sqrt{36}} = 49.5 - 3.3 = 46.2$$

$$\text{Upper bound } = \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} = 49.5 + 1.96 \cdot \frac{10}{\sqrt{36}} = 49.5 + 3.3 = 52.8$$

Just as with the smaller sample, the outlier shifts the sample mean to the left and, thus, shifts the confidence interval to the left by the same amount.. This effect is independent of the sample size. However, an outlier in a large sample does not affect the sample mean as much as does an outlier in a small sample, so it also has less effect on the confidence interval.

53. The sample size must be increased by a factor of 4. This is because the sample size, n, is inversely proportional to the square of the error, E. To decrease the error by a factor of $\frac{1}{2}$ we must increase the sample size by a factor of $\left(\frac{1}{1/2}\right)^2 = 4$.

54. The sample size must be increased by a factor of 4. This is because the sample size, n, is directly proportional to the square of the standard deviation, σ. Since the standard deviation of population B is 2 times as large as that of population A, the sample size will need to be $(2)^2 = 4$ times as large.

55. (a) Data Set I: $\bar{x} = \frac{\sum x}{n} = \frac{793}{8} \approx 99.1$;

Data Set II: $\bar{x} = \frac{\sum x}{n} = \frac{1982}{20} = 99.1$;

Data Set III: $\bar{x} = \frac{\sum x}{n} = \frac{2971}{30} \approx 99.0$

(b) For 95% confidence the critical value is $z_{0.025} = 1.96$.

Set I:
$$\text{Lower bound } = \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} = 99.1 - 1.96 \cdot \frac{15}{\sqrt{8}} \approx 99.1 - 10.4 = 88.7$$

$$\text{Upper bound } = \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} = 99.1 + 1.96 \cdot \frac{15}{\sqrt{8}} \approx 99.1 + 10.4 = 109.5$$

Set II:
$$\text{Lower bound } = \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} = 99.1 - 1.96 \cdot \frac{15}{\sqrt{20}} \approx 99.1 - 6.6 = 92.5$$

$$\text{Upper bound } = \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} = 99.1 + 1.96 \cdot \frac{15}{\sqrt{20}} \approx 99.1 + 6.6 = 105.7$$

Set III:
$$\text{Lower bound } = \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} = 99.0 - 1.96 \cdot \frac{15}{\sqrt{30}} \approx 99.0 - 5.4 = 93.6$$

[Tech: 93.7]

$$\text{Upper bound } = \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} = 99.0 + 1.96 \cdot \frac{15}{\sqrt{30}} \approx 99.0 + 5.4 = 104.4$$

(c) As the size of the sample increases, the width of the confidence interval decreases.

(d) Set I: $\bar{x} = \frac{\sum x}{n} = \frac{703}{8} = 87.9$;

Lower bound $= \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}}$
$= 87.9 - 1.96 \cdot \frac{15}{\sqrt{8}}$
$\approx 87.9 - 10.4 = 77.5$

Upper bound $= \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}}$
$= 87.9 + 1.96 \cdot \frac{15}{\sqrt{8}}$
$\approx 87.9 + 10.4 = 98.3$

Set II: $\bar{x} = \frac{\sum x}{n} = \frac{1892}{20} = 94.6$;

Lower bound $= \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}}$
$= 94.6 - 1.96 \cdot \frac{15}{\sqrt{20}}$
$\approx 94.6 - 6.6 = 88.0$

Upper bound $= \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}}$
$= 94.6 + 1.96 \cdot \frac{15}{\sqrt{20}}$
$\approx 94.6 + 6.6 = 101.2$

Set III: $\bar{x} = \frac{\sum x}{n} = \frac{2881}{30} \approx 96.0$

Lower bound $= \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}}$
$= 96.0 - 1.96 \cdot \frac{15}{\sqrt{30}}$
$\approx 96.0 - 5.4 = 90.6$
[Tech: 90.7]

Upper bound $= \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}}$
$= 96.0 + 1.96 \cdot \frac{15}{\sqrt{30}}$
$\approx 96.0 + 5.4 = 101.4$

(e) The confidence intervals for both Data Set II and Data Set III still capture the population mean, 100, with the incorrect entry. The interval from Data Set I does not capture the population mean when the incorrect entry is made. The concept of robustness is illustrated. As the sample size increases the distribution of $\bar{x}$ becomes more normal, making the confidence interval more robust against the effect of the incorrect entry.

56. **(a)** Answers will vary depending on the results from the applet. You should expect 95% of the intervals to contain the population mean.

(b) Answers will vary depending on the results from the applet.

(c) Answers will vary depending on the results from the applet. You should expect 99% of the intervals to contain the population mean.

57. **(a)** Answers will vary depending on the results from the applet. You should expect 95% of the intervals to contain the population mean.

(b) Answers will vary depending on the results from the applet.

(c) Answers will vary depending on the results from the applet. You should expect 95% of the intervals to contain the population mean.

(d) Confidence intervals for the samples of size $n = 10$ should be wider than the confidence intervals for the samples of size $n = 50$.

58. **(a)** Answers will vary depending on the results from the applet. You should expect 95% of the intervals to contain the population mean.

(b) Answers will vary depending on the results from the applet. You should expect 95% of the intervals to contain the population mean.

(c) Non-normal data when the sample size is small will impact the proportion of intervals that actually contain the population mean, making the true proportion less than what is expected.

59. **(a)** Because all subjects were randomly assigned to the treatments, this is a completely randomized design.

(b) The treatment is the smoking cessation program. There are two levels: internet and phone-based intervention, and self-help booklet.

(c) The response variable was abstinence after 12 months (whether or not the subject quit smoking).

(d) The statistics reported are 22.3% of participants in the experimental group reported abstinence and 13.1% of participants in the control group reported abstinence.

(e) $\frac{p(1-q)}{q(1-p)} = \frac{0.223(1-0.131)}{0.131(1-0.223)} \approx 1.90$; this indicates that abstinence is more likely to occur using the experimental cessation program.

(f) The authors are 95% confident that the population odds ratio is between 1.12 and 3.26.

(g) Answers will vary. One possibility: Smoking cessation is more likely when the Happy Ending Intervention program is used rather than the control method.

9.2 Confidence Intervals for a Population Mean When the Population Standard Deviation is Unknown

1. We can construct a Z-interval if the sample is random, the population from which the sample is drawn is normal or the sample size is large $(n \geq 30)$, and the population standard deviation, σ, is known. A t-interval should be constructed if the sample is random, the population from which the sample is drawn is normal, but the population standard deviation, σ, is unknown. Neither interval can be constructed if the sample is not random, the population is not normal and the sample size is small, or when there are outliers.

2. As the degrees of freedom increase, the t-distribution has less spread because as n increases the value of s becomes closer to the value of σ, by the Law of Large Numbers.

3. Robust means that the procedure is accurate when there are moderate departures from the requirements, such as normality in the distribution of the population.

4. t-values are more dispersed than z-values because the sample standard deviation, s, is a random variable while σ is a constant.

5. Similarities: Both the standard normal distribution and the t-distribution are probability density functions; both have mean $\mu = 0$, and both are symmetric about their means.
Differences: t-distributions vary for different sample sizes while there is only one standard normal distribution. t-distributions have longer and thicker tails than the standard normal distribution, which has most of its area between –3 and 3.

6. Answers will vary. One possibility follows: The degrees of freedom are the number of values that are free to vary after a sample statistic, such as $\bar{x}$, has been calculated.

7. (a) From the row with df = 25 and the column headed 0.10, we read $t_{0.10} = 1.316$.

   (b) From the row with df = 30 and the column headed 0.05, we read $t_{0.05} = 1.697$.

   (c) From the row with df = 18 and the column headed 0.01, we read $t = 2.552$. This is the value with area 0.01 to the *right*. By symmetry, the t-value with an area to the *left* of 0.01 is $t_{0.99} = -2.552$.

   (d) For a 90% confidence interval, we want the t-value with an area in the right tail of 0.05. With df = 20, we read from Table VI that $t_{0.05} = 1.725$.

8. (a) From the row with df = 19 and the column headed 0.02 we read $t_{0.02} = 2.205$.

   (b) From the row with df = 32 and the column headed 0.10 we read $t_{0.10} = 1.309$.

   (c) From the row with df = 6 and the column headed 0.05 we read $t = 1.943$. This is the value with area 0.05 to the *right*. By symmetry, the t-value with an area to the *left* of 0.05 is $t_{0.95} = -1.943$.

   (d) For a 95% confidence interval we want the t-value with an area in the right tail of 0.025. With df = 16 we read from Table VI that $t_{0.025} = 2.120$.

9. **(a)** For 96% confidence, $\alpha/2 = 0.02$. Since $n = 25$, then df = 24. The critical value is $t_{0.02} = 2.172$. Then:

$$\text{Lower bound} = \bar{x} - t_{0.02} \cdot \frac{s}{\sqrt{n}} = 108 - 2.172 \cdot \frac{10}{\sqrt{25}} \approx 108 - 4.3 = 103.7$$

$$\text{Upper bound} = \bar{x} + t_{0.02} \cdot \frac{s}{\sqrt{n}} = 108 + 2.172 \cdot \frac{10}{\sqrt{25}} \approx 108 + 4.3 = 112.3$$

(b) Since $n = 10$, then df = 9. The critical value is $t_{0.02} = 2.398$. Then:

$$\text{Lower bound} = \bar{x} - t_{0.02} \cdot \frac{s}{\sqrt{n}} = 108 - 2.398 \cdot \frac{10}{\sqrt{10}} \approx 108 - 7.6 = 100.4$$

$$\text{Upper bound} = \bar{x} + t_{0.02} \cdot \frac{s}{\sqrt{n}} = 108 + 2.398 \cdot \frac{10}{\sqrt{10}} \approx 108 + 7.6 = 115.6$$

Decreasing the sample size increases the margin of error.

(c) For 90% confidence, $\alpha/2 = 0.05$. With 24 degrees of freedom, $t_{0.05} = 1.711$. Then:

$$\text{Lower bound} = \bar{x} - t_{0.05} \cdot \frac{s}{\sqrt{n}} = 108 - 1.711 \cdot \frac{10}{\sqrt{25}} \approx 108 - 3.4 = 104.6$$

$$\text{Upper bound} = \bar{x} + t_{0.05} \cdot \frac{s}{\sqrt{n}} = 108 + 1.711 \cdot \frac{10}{\sqrt{25}} \approx 108 + 3.4 = 111.4$$

Decreasing the level of confidence decreases the margin of error.

(d) No, because in all cases the sample was small ($n < 30$), so the population must be normally distributed.

10. **(a)** For 98% confidence, $\alpha/2 = 0.01$. Since $n = 20$, then df = 19 and $t_{0.01} = 2.539$. Then:

$$\text{Lower bound} = \bar{x} - t_{0.01} \cdot \frac{s}{\sqrt{n}} = 50 - 2.539 \cdot \frac{8}{\sqrt{20}} \approx 50 - 4.5 = 45.5$$

$$\text{Upper bound} = \bar{x} + t_{0.01} \cdot \frac{s}{\sqrt{n}} = 50 + 2.539 \cdot \frac{8}{\sqrt{20}} \approx 50 + 4.5 = 54.5$$

(b) Since $n = 15$, then df = 14. The critical value is $t_{0.01} = 2.624$. Then:

$$\text{Lower bound} = \bar{x} - t_{0.01} \cdot \frac{s}{\sqrt{n}} = 50 - 2.624 \cdot \frac{8}{\sqrt{15}} \approx 50 - 5.4 = 44.6$$

$$\text{Upper bound} = \bar{x} + t_{0.01} \cdot \frac{s}{\sqrt{n}} = 50 + 2.624 \cdot \frac{8}{\sqrt{15}} \approx 50 + 5.4 = 55.4$$

Decreasing the sample size increases the margin of error.

(c) For 95% confidence, $\alpha/2 = 0.025$. With 19 degrees of freedom, $t_{0.025} = 2.093$. Then:

$$\text{Lower bound} = \bar{x} - t_{0.025} \cdot \frac{s}{\sqrt{n}} = 50 - 2.093 \cdot \frac{8}{\sqrt{20}} \approx 50 - 3.7 = 46.3$$

$$\text{Upper bound} = \bar{x} + t_{0.025} \cdot \frac{s}{\sqrt{n}} = 50 + 2.093 \cdot \frac{8}{\sqrt{20}} \approx 50 + 3.7 = 53.7$$

Decreasing the level of confidence decreases the margin of error.

(d) No, because in all cases the sample was small ($n < 30$) and so the population must be normally distributed.

11. **(a)** For 95% confidence, $\alpha/2 = 0.025$. Since $n = 35$, then df = 34. The critical value is $t_{0.025} = 2.032$. Then:

$$\text{Lower bound } = \bar{x} - t_{0.025} \cdot \frac{s}{\sqrt{n}}$$
$$= 18.4 - 2.032 \cdot \frac{4.5}{\sqrt{35}}$$
$$\approx 18.4 - 1.55 = 16.85$$

$$\text{Upper bound } = \bar{x} + t_{0.025} \cdot \frac{s}{\sqrt{n}}$$
$$= 18.4 + 2.032 \cdot \frac{4.5}{\sqrt{35}}$$
$$\approx 18.4 + 1.55 = 19.95$$

(b) Since $n = 50$, then df = 49, but since there is no row in the tables for 49 degrees of freedom we use the closest value, df = 50, instead. The critical value is $t_{0.025} = 2.009$. Then:

$$\text{Lower bound } = \bar{x} - t_{0.025} \cdot \frac{s}{\sqrt{n}}$$
$$= 18.4 - 2.009 \cdot \frac{4.5}{\sqrt{50}}$$
$$\approx 18.4 - 1.28 = 17.12$$

$$\text{Upper bound } = \bar{x} + t_{0.025} \cdot \frac{s}{\sqrt{n}}$$
$$= 18.4 + 2.009 \cdot \frac{4.5}{\sqrt{50}}$$
$$\approx 18.4 + 1.28 = 19.68$$

Increasing the sample size decreases the margin of error.

(c) For 99% confidence, $\alpha/2 = 0.005$. With 34 degrees of freedom, $t_{0.005} = 2.728$. Then:

$$\text{Lower bound } = \bar{x} - t_{0.005} \cdot \frac{s}{\sqrt{n}}$$
$$= 18.4 - 2.728 \cdot \frac{4.5}{\sqrt{35}}$$
$$\approx 18.4 - 2.08 = 16.32$$

$$\text{Upper bound } = \bar{x} + t_{0.005} \cdot \frac{s}{\sqrt{n}}$$
$$= 18.4 + 2.728 \cdot \frac{4.5}{\sqrt{35}}$$
$$\approx 18.4 + 2.08 = 20.48$$

$\left[\text{Tech: } (16.33, 20.48)\right]$

Increasing the level of confidence increases the margin of error.

(d) For a small sample ($n = 15 < 30$), the population must be normally distributed.

12. **(a)** For 90% confidence, $\alpha/2 = 0.05$. Since $n = 40$, then df = 39 and $t_{0.05} = 1.685$. Then:

$$\text{Lower bound } = \bar{x} - t_{0.05} \cdot \frac{s}{\sqrt{n}}$$
$$= 35.1 - 1.685 \cdot \frac{8.7}{\sqrt{40}}$$
$$\approx 35.1 - 2.32 = 32.78$$

$$\text{Upper bound } = \bar{x} + t_{0.05} \cdot \frac{s}{\sqrt{n}}$$
$$= 35.1 + 1.685 \cdot \frac{8.7}{\sqrt{40}}$$
$$\approx 35.1 + 2.32 = 37.42$$

(b) Since $n = 100$, then df = 99, but since there is no row in the tables for 99 degrees of freedom we use the closest value, df = 100, instead. The critical value is $t_{0.05} = 1.660$. Then:

$$\text{Lower bound } = \bar{x} - t_{0.05} \cdot \frac{s}{\sqrt{n}}$$
$$= 35.1 - 1.660 \cdot \frac{8.7}{\sqrt{100}}$$
$$\approx 35.1 - 1.44 = 33.66$$

$$\text{Upper bound } = \bar{x} + t_{0.05} \cdot \frac{s}{\sqrt{n}}$$
$$= 35.1 + 1.660 \cdot \frac{8.7}{\sqrt{100}}$$
$$\approx 35.1 + 1.44 = 36.54$$

Increasing the sample size decreases the margin of error.

(c) For 98% confidence, $\alpha/2 = 0.01$. With 39 degrees of freedom, $t_{0.01} = 2.426$. Then:

$$\text{Lower bound } = \bar{x} - t_{0.01} \cdot \frac{s}{\sqrt{n}}$$
$$= 35.1 - 2.426 \cdot \frac{8.7}{\sqrt{40}}$$
$$\approx 35.1 - 3.34 = 31.76$$

$$\text{Upper bound } = \bar{x} + t_{0.01} \cdot \frac{s}{\sqrt{n}}$$
$$= 35.1 + 2.426 \cdot \frac{8.7}{\sqrt{40}}$$
$$\approx 35.1 + 3.34 = 38.44$$

Increasing the level of confidence increases the margin of error.

(d) For a small sample ($n = 18 < 30$), the population must be normally distributed.

13. For 99% confidence, $\alpha/2=0.005$. Since $n=1006$, then df = 1005. There is no row in the table for 1005 degrees of freedom, so we use the closest value, df = 1000, instead. The critical value is $t_{0.005}=2.581$. Then:

$$\text{Lower bound } =\bar{x}-t_{0.005}\cdot\frac{s}{\sqrt{n}} =13.4-2.581\cdot\frac{16.6}{\sqrt{1006}} \approx 13.4-1.35=12.05 \text{ books}$$

$$\text{Upper bound } =\bar{x}-t_{0.005}\cdot\frac{s}{\sqrt{n}} =13.4+2.581\cdot\frac{16.6}{\sqrt{1006}} \approx 13.4+1.35=14.75 \text{ books}$$

We are 99% confident that the population mean number of books read by Americans during 2005 was between 12.05 and 14.75 books.

14. **(a)** For 99% confidence, $\alpha/2=0.005$. Since $n=1006$, then df = 1005. There is no row in the table for 1005 degrees of freedom, so we use df = 1000 instead. The critical value is $t_{0.005}=2.581$. Then:

$$\text{Lower bound } =\bar{x}-t_{0.005}\cdot\frac{s}{\sqrt{n}} =18.8-2.581\cdot\frac{19.8}{\sqrt{1006}} \approx 18.8-1.61=17.19 \text{ books}$$

$$\text{Upper bound } =\bar{x}+t_{0.005}\cdot\frac{s}{\sqrt{n}} =18.8+2.581\cdot\frac{19.8}{\sqrt{1006}} \approx 18.8+1.61=20.41 \text{ books}$$

We are 99% confident that the population mean number of books read by Americans during 1978 was between 17.19 and 20.41 books.

(b) Since the lower bound for 1978 is above the upper bound for 2005, we conclude that Americans were reading more in 1978 than in 2005.

15. For 95% confidence, $\alpha/2=0.025$. Since $n=81$, then df = 80 and $t_{0.025}=1.990$. Then:

$$\text{Lower bound } =\bar{x}-t_{0.025}\cdot\frac{s}{\sqrt{n}} =4.6-1.990\cdot\frac{15.9}{\sqrt{81}} \approx 4.6-3.52=1.08 \text{ days}$$

$$\text{Upper bound } =\bar{x}+t_{0.025}\cdot\frac{s}{\sqrt{n}} =4.6+1.990\cdot\frac{15.9}{\sqrt{81}} \approx 4.6+3.52=8.12 \text{ days}$$

We are 95% confident that the population mean incubation period of patients with SARS is between 1.08 and 8.12 days.

16. For 90% confidence, $\alpha/2=0.05$. Since $n=72$, then df = 71. There is no row in the table for 71 degrees of freedom, so we use df = 70 instead. Thus, $t_{0.05}=1.667$. Then:

$$\text{Lower bound } =\bar{x}-t_{0.05}\cdot\frac{s}{\sqrt{n}} =242.2-1.667\cdot\frac{70.6}{\sqrt{72}} \approx 242.2-13.87=228.33 \text{ N}$$

$$\text{Upper bound } =\bar{x}+t_{0.05}\cdot\frac{s}{\sqrt{n}} =242.2+1.667\cdot\frac{70.6}{\sqrt{72}} \approx 242.2+13.87=256.07 \text{ N}$$

Researchers are 90% confident that the population mean tensile strength of the resin cement is between 228.33 and 256.07 newtons.

17. **(a)** For 90% confidence, $\alpha/2=0.05$. Since $n=547$, then df = 546. There is no row in the table for 546 degrees of freedom, so we use the closest value, df = 1000, instead. The critical value is $t_{0.05}=1.646$. Then:

$$\text{Lower bound } =\bar{x}-t_{0.05}\cdot\frac{s}{\sqrt{n}} =45.6-1.646\cdot\frac{31.4}{\sqrt{547}} \approx 45.6-2.21=43.39 \text{ min}$$

$$\text{Upper bound } =\bar{x}+t_{0.05}\cdot\frac{s}{\sqrt{n}} =45.6+1.646\cdot\frac{31.4}{\sqrt{547}} \approx 45.6+2.21=47.81 \text{ min}$$

Gallup is 90% confident that the population mean commute time of adult Americans employed full-time or part-time is between 43.39 minutes and 47.81 minutes.

(b) Yes; it is possible that the mean commute time is less than 40 minutes since it is possible that the true mean is not captured in the confidence interval. However, it is not very likely since we are 90% confident the true mean commute time is between 43.39 minutes and 47.81 minutes.

18. (a) For 95% confidence, $\alpha/2 = 0.025$. Since $n = 40$, then df = 39 and $t_{0.025} = 2.023$. Then:

$$\begin{aligned}\text{Lower bound } &= \bar{x} - t_{0.025} \cdot \frac{s}{\sqrt{n}} \\ &= 20 - 2.023 \cdot \frac{1.5}{\sqrt{40}} \\ &\approx 20 - 0.48 = 19.52\end{aligned}$$

$$\begin{aligned}\text{Upper bound } &= \bar{x} + t_{0.025} \cdot \frac{s}{\sqrt{n}} \\ &= 20 + 2.023 \cdot \frac{1.5}{\sqrt{40}} \\ &\approx 20 + 0.48 = 20.48\end{aligned}$$

We are 95% confident that the population mean gas mileage of 2007 Ford Mustangs (6 cylinder, 4 liter, 5-speed manual) is between 19.52 and 20.48 miles per gallon.

(b) Yes; it is possible that the mean commute time is greater than 23 miles per gallon since it is possible that the true mean is not captured in the confidence interval. However, it is not very likely since we are 95% confident the true mean gas mileage is between 19.52 and 20.48 miles per gallon.

19. Using technology, we find $\bar{x} = 3$ days and $s \approx 2.24$ days. For 95% confidence, $\alpha/2 = 0.025$. Since $n = 32$, then df = 31 and $t_{0.025} = 2.040$. Then:

$$\begin{aligned}\text{Lower bound } &= \bar{x} - t_{0.025} \cdot \frac{s}{\sqrt{n}} \\ &= 3 - 2.040 \cdot \frac{2.24}{\sqrt{32}} \\ &\approx 3 - 0.81 = 2.19 \text{ days}\end{aligned}$$

$$\begin{aligned}\text{Upper bound } &= \bar{x} + t_{0.025} \cdot \frac{s}{\sqrt{n}} \\ &= 3 + 2.040 \cdot \frac{2.24}{\sqrt{32}} \\ &\approx 3 + 0.81 = 3.81 \text{ days}\end{aligned}$$

The researcher is 95% confident that the population mean number of days per week in which adults engage in exercise activities is between 2.19 and 3.81 days.

20. Using technology, we find $\bar{x} \approx 18.71\,\%$ and $s \approx 2.76\,\%$. For 95% confidence, $\alpha/2 = 0.025$. Since $n = 14$, then df = 13 and $t_{0.025} = 2.160$. Then:

$$\begin{aligned}\text{Lower bound } &= \bar{x} - t_{0.025} \cdot \frac{s}{\sqrt{n}} \\ &= 18.71 - 2.160 \cdot \frac{2.76}{\sqrt{14}} \\ &\approx 18.71 - 1.59 = 17.12\end{aligned}$$

[Tech: 17.11]

$$\begin{aligned}\text{Upper bound } &= \bar{x} + t_{0.025} \cdot \frac{s}{\sqrt{n}} \\ &= 18.71 + 2.160 \cdot \frac{2.76}{\sqrt{14}} \\ &\approx 18.71 + 1.59 = 20.30\end{aligned}$$

The server is 95% confident that the population mean tip percentage per dinner is between 17.12% and 20.30%.

21. (a) $\bar{x} = \dfrac{\sum x}{n} = \dfrac{93.48}{40} = 2.337$ million shares

(b) Using technology, we find $s \approx 1.217$ million shares. For 90% confidence, $\alpha/2 = 0.05$. Since $n = 40$, then df = 39 and $t_{0.05} = 1.685$. Thus:

$$\begin{aligned}\text{Lower bound } &= \bar{x} - t_{0.05} \cdot \frac{s}{\sqrt{n}} \\ &= 2.337 - 1.685 \cdot \frac{1.217}{\sqrt{40}} \\ &\approx 2.337 - 0.324 = 2.013\end{aligned}$$

$$\begin{aligned}\text{Upper bound } &= \bar{x} + t_{0.05} \cdot \frac{s}{\sqrt{n}} \\ &= 2.337 + 1.685 \cdot \frac{1.217}{\sqrt{40}} \\ &\approx 2.337 + 0.324 = 2.661\end{aligned}$$

We are 90% confident that the population mean number of Harley-Davidson shares traded per day in 2007 was between 2.013 and 2.661 million shares.

(c) Using technology, we find $\bar{x} \approx 2.195$ million shares and $s \approx 0.815$ million shares. For 90% confidence, $\alpha/2 = 0.05$. Since $n = 40$, then df = 39 and $t_{0.05} = 1.685$. Thus:

$$\text{Lower bound} = \bar{x} - t_{0.05} \cdot \frac{s}{\sqrt{n}} = 2.195 - 1.685 \cdot \frac{0.815}{\sqrt{40}} \approx 2.195 - 0.217 = 1.978$$

$$\text{Upper bound} = \bar{x} + t_{0.05} \cdot \frac{s}{\sqrt{n}} = 2.195 + 1.685 \cdot \frac{0.815}{\sqrt{40}} \approx 2.195 + 0.217 = 2.412$$

We are 90% confident that the population mean number of Harley-Davidson shares traded per day in 2007 was between 1.978 and 2.412 million shares.

(d) The confidence intervals are different because of variation in sampling. The samples have different means and standard deviations that lead to different confidence intervals.

22. (a) $\bar{x} = \frac{\sum x}{n} = \frac{217.19}{40} \approx 5.430$ million shares

(b) Using technology, we find $s \approx 1.480$ million shares. For 95% confidence, $\alpha/2 = 0.025$. Since $n = 40$, then $\text{df} = 39$ and $t_{0.025} = 2.023$. Thus:

$$\text{Lower bound} = \bar{x} - t_{0.025} \cdot \frac{s}{\sqrt{n}} = 5.430 - 2.023 \cdot \frac{1.480}{\sqrt{40}} \approx 5.430 - 0.473 = 4.957$$

[Tech: 4.956]

$$\text{Upper bound} = \bar{x} + t_{0.025} \cdot \frac{s}{\sqrt{n}} = 5.430 + 2.023 \cdot \frac{1.480}{\sqrt{40}} \approx 5.430 + 0.473 = 5.903$$

We are 95% confident that the population mean number of PepsiCo shares traded per day in 2007 was between 4.957 and 5.903 million shares.

(c) Using technology, we find $\bar{x} \approx 5.810$ million shares and $s \approx 1.600$ million shares. For 95% confidence, $\alpha/2 = 0.025$. Since $n = 40$, then $\text{df} = 39$ and $t_{0.025} = 2.023$. Thus:

$$\text{Lower bound} = \bar{x} - t_{0.025} \cdot \frac{s}{\sqrt{n}} = 5.810 - 2.023 \cdot \frac{1.600}{\sqrt{40}} \approx 5.810 - 0.512 = 5.298$$

$$\text{Upper bound} = \bar{x} + t_{0.025} \cdot \frac{s}{\sqrt{n}} = 5.810 + 2.023 \cdot \frac{1.600}{\sqrt{40}} \approx 5.810 + 0.512 = 6.322$$

[Tech: 6.321]

We are 95% confident that the population mean number of PepsiCo shares traded per day in 2007 was between 5.298 and 6.322 million shares.

(d) The confidence intervals are different because of variation in sampling. The samples have different means and standard deviations that lead to different confidence intervals.

23. (a) Yes. The normal probability plot indicates that the data come from a population that is approximately normal, and the box plot indicates that there are no outliers.

(b) Using technology, $\bar{x} \approx 38.3$ weeks and $s \approx 10.0$ weeks.
For 95% confidence, $\alpha/2 = 0.025$. Since $n = 12$, then $\text{df} = 11$ and $t_{0.025} = 2.201$.

$$\text{Lower bound} = \bar{x} - t_{0.025} \cdot \frac{s}{\sqrt{n}} = 38.3 - 2.201 \cdot \frac{10.0}{\sqrt{12}} \approx 38.3 - 6.4 = 31.9 \text{ weeks}$$

$$\text{Upper bound} = \bar{x} + t_{0.025} \cdot \frac{s}{\sqrt{n}} = 38.3 + 2.201 \cdot \frac{10.0}{\sqrt{12}} \approx 38.3 + 6.4 = 44.7 \text{ weeks}$$

We are 95% confident that the population mean age at which a baby first crawls is between 31.9 and 44.7 weeks.

(c) The sample size could be increased in order to increase the accuracy of the interval without changing the confidence level.

24. **(a)** Yes. The normal probability plot indicates that the data come from a population that is approximately normal, and the box plot indicates that there are no outliers.

(b) Using technology, $\bar{x} = 9.47$ hours and $s \approx 2.14$ hours.
For 90% confidence, $\alpha/2 = 0.05$. Since $n = 10$, then df = 9 and $t_{0.05} = 1.833$.

$$\begin{aligned}\text{Lower bound } &= \bar{x} - t_{0.05} \cdot \frac{s}{\sqrt{n}} \\ &= 9.47 - 1.833 \cdot \frac{2.14}{\sqrt{10}} \\ &\approx 9.47 - 1.24 = 8.23 \text{ hours}\end{aligned}$$

$$\begin{aligned}\text{Upper bound } &= \bar{x} + t_{0.05} \cdot \frac{s}{\sqrt{n}} \\ &= 9.47 + 1.833 \cdot \frac{2.14}{\sqrt{10}} \\ &\approx 9.47 + 1.24 = 10.71 \text{ hours}\end{aligned}$$

We are 90% confident that the population mean number of hours the battery will last on this player is between 8.23 and 10.71 hours.

(c) The sample size could be increased in order to increase the accuracy of the interval without changing the confidence level.

25. **(a)** Yes. The normal probability plot indicates that the data come from a population that is approximately normal, and the box plot indicates that there are no outliers.

(b) Using technology, $\bar{x} = 167.5$ days and $s \approx 21.9$ days.
For 95% confidence, $\alpha/2 = 0.025$. Since $n = 10$, then df = 9 and $t_{0.025} = 2.262$.

$$\begin{aligned}\text{Lower bound } &= \bar{x} - t_{0.025} \cdot \frac{s}{\sqrt{n}} \\ &= 167.5 - 2.262 \cdot \frac{21.9}{\sqrt{10}} \\ &\approx 167.5 - 15.7 = 151.8 \text{ days}\end{aligned}$$

$$\begin{aligned}\text{Upper bound } &= \bar{x} + t_{0.025} \cdot \frac{s}{\sqrt{n}} \\ &= 167.5 + 2.262 \cdot \frac{21.9}{\sqrt{10}} \\ &\approx 167.5 + 15.7 = 183.2 \text{ days}\end{aligned}$$

We are 95% confident that the population mean length of the growing season in the Chicago area is between 151.8 and 183.2 days.

(c) The sample size could be increased in order to increase the accuracy of the interval without change the confidence level.

26. **(a)** Yes. The normal probability plot indicates that the data come from a population that is approximately normal, and the box plot indicates that there are no outliers.

(b) Using technology, $\bar{x} \approx 171.7$ grams and $s \approx 2.0$ grams.
For 95% confidence, $\alpha/2 = 0.025$. Since $n = 12$, then df = 11 and $t_{0.025} = 2.201$.

$$\begin{aligned}\text{Lower bound } &= \bar{x} - t_{0.025} \cdot \frac{s}{\sqrt{n}} \\ &= 171.7 - 2.201 \cdot \frac{2.0}{\sqrt{12}} \\ &\approx 171.7 - 1.3 = 170.4 \text{ grams}\end{aligned}$$

$$\begin{aligned}\text{Upper bound } &= \bar{x} + t_{0.025} \cdot \frac{s}{\sqrt{n}} \\ &= 171.7 + 2.201 \cdot \frac{2.0}{\sqrt{12}} \\ &\approx 171.7 + 1.3 = 173.0 \text{ grams}\end{aligned}$$

We are 95% confident that the population mean disc weight is between 170.4 and 173.0 grams.

27. **(a)** From the output:

$$\bar{x} = 22.150;\ E = \frac{22.209 - 22.091}{2} = 0.059$$

(b) We are 95% confident that the population mean age when first married is between 22.091 and 22.209 years.

(c) For 95% confidence, $\alpha/2 = 0.025$. Since $n = 26{,}540$, then df = 26,539. There is no row in the table for 26,539 degrees of freedom, so we use df = 1000 instead. and $t_{0.025} = 1.962$.

$$\begin{aligned}\text{Lower bound } &= \bar{x} - t_{0.025} \cdot \frac{s}{\sqrt{n}} \\ &= 22.150 - 1.962 \cdot \frac{4.885}{\sqrt{26{,}540}} \\ &\approx 22.150 - 0.059 = 22.091\end{aligned}$$

$$\begin{aligned}\text{Upper bound } &= \bar{x} + t_{0.025} \cdot \frac{s}{\sqrt{n}} \\ &= 22.150 + 1.962 \cdot \frac{4.885}{\sqrt{26{,}540}} \\ &\approx 22.150 + 0.059 = 22.209\end{aligned}$$

28. **(a)** From the output:

$$\bar{x} = 7.678\,;\ E = \frac{8.390 - 6.966}{2} = 0.712$$

(b) We are 95% confident that, of people who engaged in sexual intercourse during the past month, the mean number of times they engaged in intercourse during that month was between 6.966 and 8.390.

(c) For 95% confidence, $\alpha/2 = 0.025$. Since $n = 357$, then df = 356. There is no row in the table for 356 degrees of freedom, so we use df = 100 instead. and $t_{0.025} = 1.984$.

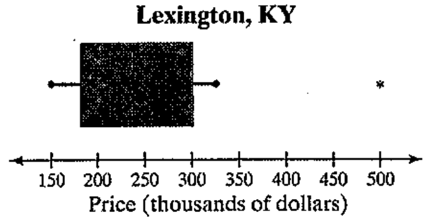
$$\text{Lower bound} = \bar{x} - t_{0.025} \cdot \frac{s}{\sqrt{n}} = 7.678 - 1.984 \cdot \frac{6.843}{\sqrt{357}} \approx 7.678 - 0.719 = 6.959$$

$$\text{Upper bound} = \bar{x} + t_{0.025} \cdot \frac{s}{\sqrt{n}} = 7.678 + 1.984 \cdot \frac{6.843}{\sqrt{357}} \approx 7.678 + 0.719 = 8.397$$

Note: differences in the bounds is due to the use of df = 100 instead of df = 356.

29. **(a)**

Home Asking Price, Lexington, KY



(b) Using technology with the outlier \$496,600 included, $\bar{x} \approx \$255{,}383.3$ and $s \approx \$93{,}444.0$.

For 99% confidence, $\alpha/2 = 0.005$. Since $n = 12$, then df = 11 and $t_{0.005} = 3.106$.

$$\text{Lower bound} = \bar{x} - t_{0.005} \cdot \frac{s}{\sqrt{n}} = 255{,}383.3 - 3.106 \cdot \frac{93{,}444.0}{\sqrt{12}} \approx 255{,}383.3 - 83{,}784.2 = \$171{,}599.1$$

$$\text{Upper bound} = \bar{x} + t_{0.005} \cdot \frac{s}{\sqrt{n}} = 255{,}383.3 + 3.106 \cdot \frac{93{,}444.0}{\sqrt{12}} \approx 255{,}383.3 + 83{,}784.2 = \$339{,}167.5$$

[Tech: (\$171,604.3, \$339,162.3)]

(c) Using technology with the outlier \$496,600 removed, $\bar{x} \approx \$233{,}454.5$ and $s \approx \$57{,}074.1$.

Since the outlier was removed, $n = 11$, df = 10, and $t_{0.005} = 3.169$.

$$\text{Lower bound} = \bar{x} - t_{0.005} \cdot \frac{s}{\sqrt{n}} = 233{,}454.5 - 3.169 \cdot \frac{57{,}074.1}{\sqrt{11}} \approx 233{,}454.5 - 54{,}533.7 = \$178{,}920.8$$

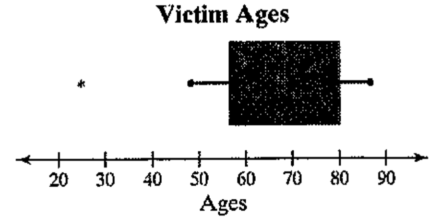
$$\text{Upper bound} = \bar{x} + t_{0.005} \cdot \frac{s}{\sqrt{n}} = 233{,}454.5 + 3.169 \cdot \frac{57{,}074.1}{\sqrt{11}} \approx 233{,}454.5 + 54{,}533.7 = \$287{,}988.2$$

[Tech: (\$178,916.2, \$287,992.9)]

(d) The inclusion of the outlier makes the confidence interval wider.

30. **(a)**

West Nile Virus Victim Ages



(b) Using technology with the outlier 25 included, $\bar{x} = 65.5$ victims and $s \approx 17.0$ victims.

For 99% confidence, $\alpha/2 = 0.005$. Since $n = 12$, then df = 11 and $t_{0.005} = 3.106$.

$$\text{Lower bound} = \bar{x} - t_{0.005} \cdot \frac{s}{\sqrt{n}} = 65.5 - 3.106 \cdot \frac{17.0}{\sqrt{12}} \approx 65.5 - 15.2 = 50.3$$

Upper bound $= \bar{x} + t_{0.005} \cdot \frac{s}{\sqrt{n}}$

$= 65.5 + 3.106 \cdot \frac{17.0}{\sqrt{12}}$

$\approx 65.5 + 15.2 = 80.7$

With the outlier included, the 99% confidence interval is from 50.3 to 80.7 victims.

(c) Using technology with the outlier 25 removed, $\bar{x} \approx 69.2$ victims and $s \approx 11.8$ victims.

Since the outlier was removed, $n = 11$, df = 10, and $t_{0.005} = 3.169$.

Lower bound $= \bar{x} - t_{0.005} \cdot \frac{s}{\sqrt{n}}$

$= 69.2 - 3.169 \cdot \frac{11.8}{\sqrt{11}}$

$\approx 69.2 - 11.3 = 57.9$

Upper bound $= \bar{x} + t_{0.005} \cdot \frac{s}{\sqrt{n}}$

$= 69.2 + 3.169 \cdot \frac{11.8}{\sqrt{11}}$

$\approx 69.2 + 11.3 = 80.5$

With the outlier removed, the 99% confidence interval is from 57.9 to 80.5 victims.

(d) The inclusion of the outlier makes the confidence interval wider.

31. **(a), (b), (c)** Answers will vary.

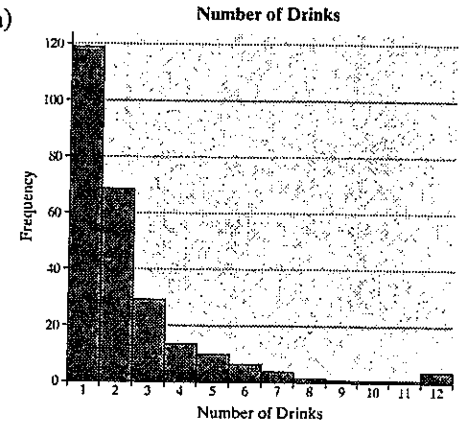
(d) 95% of 20 is $0.95(20) = 19$. We would expect about 19 of the 20 samples to generate confidence intervals that include the population mean. The actual results will vary.

32. **(a), (b)** Answers will vary.

(c) 90% of 30 is $0.90(30) = 27$. We would expect about 27 of the 30 samples to generate confidence intervals that include the population mean. The actual results will vary.

33. Answers will vary.

34. **(a)**



The distribution is skewed right.

(b) $\bar{x} = \frac{\sum x \cdot f}{243} = \frac{498}{243} \approx 2.0$ drinks

$s = \sqrt{\frac{\sum (x-\bar{x})^2 \cdot f}{\sum f - 1}} = \sqrt{\frac{\sum \left(x - \frac{498}{243}\right)^2 \cdot f}{242}}$

≈ 1.6 drinks

(c) The mode number of drinks is 1 since this happened with the most frequency.

(d) $P(2 \text{ drinks}) \approx \frac{66}{243} \approx 0.2716$

(e) $P(\text{at least 8 drinks}) \approx \frac{2}{243} \approx 0.0082$

This would be unusual since the probability is less than 0.05.

(f) Since the sample size is large $(243 > 30)$, we can apply the Central Limit Theorem. The distribution of the sample mean will be approximately normal.

(g) For 95% confidence, $\alpha/2 = 0.025$. Since $n = 243$, then df = 241. There is no row in the table for 241 degrees of freedom, so we use df = 100 instead. and $t_{0.025} = 1.984$.

Lower bound $= \bar{x} - t_{0.025} \cdot \frac{s}{\sqrt{n}}$

$= 2.0 - 1.984 \cdot \frac{1.6}{\sqrt{243}}$

$\approx 2.0 - 0.2 = 1.8$ [Tech: 1.9]

Upper bound $= \bar{x} + t_{0.025} \cdot \frac{s}{\sqrt{n}}$

$= 2.0 + 1.984 \cdot \frac{1.6}{\sqrt{243}}$

$\approx 2.0 + 0.2 = 2.2$

We are 95% confident that the mean number of drinks consumed, when people drink, is between 1.8 and 2.2 drinks.

Consumer Reports®: Consumer Reports Tests Tires

(a) All the data values lie within the bounds on the normal probability plot, indicating that the data is roughly normally distributed.

(b) The five-number summary for the data is 123.2, 128.8, 131.8, 136.9, 140.3.

The lower fence $= Q_1 - 1.5 \cdot \text{IQR}$

$$= 128.8 - 1.5(136.9 - 128.8)$$
$$= 116.65 \text{ feet}$$

The upper fence $= Q_3 + 1.5 \cdot \text{IQR}$

$$= 136.9 + 1.5(136.9 - 128.8)$$
$$= 149.05 \text{ feet}$$

No data values are below the lower fence or above the upper fence, so the data set does not contain any outliers.

(c) Using technology, $\bar{x} \approx 132.12$ feet and $s \approx 5.47$ feet.

For 95% confidence, $\alpha/2 = 0.025$. Since $n = 9$, then df = 8 and $t_{0.025} = 2.306$.

$$\text{Lower bound} = \bar{x} - t_{0.025} \cdot \frac{s}{\sqrt{n}} = 132.12 - 2.306 \cdot \frac{5.47}{\sqrt{9}} \approx 132.12 - 4.20 = 127.92 \text{ feet}$$

$$\text{Upper bound} = \bar{x} + t_{0.025} \cdot \frac{s}{\sqrt{n}} = 132.12 + 2.306 \cdot \frac{5.47}{\sqrt{9}} \approx 132.12 + 4.20 = 136.32 \text{ feet}$$

Based on this test, *Consumer Reports* is 95% confident that the mean dry braking distance for this brand of tires is between 127.92 and 136.32 feet.

9.3 Confidence Intervals for a Population Proportion

1. The best point estimate of the population proportion is the sample proportion, $\hat{p} = \frac{x}{n}$.

2. To construct a confidence interval about a population proportion, the sample size n must be no more than 5% of the population size N, and $n\hat{p}(1-\hat{p}) \geq 10$.

3. Answers will vary. One possibility follows: By using a prior estimate of p the researcher will get a better estimate of the required sample size, which will be smaller than the "worst-case" sample size given by using no prior estimate of p.

4. Answers will vary. One possibility follows: If a margin of error is not provided, then we have no estimate about how far off the results of the survey might be.

5. $\hat{p} = \frac{x}{n} = \frac{30}{150} = 0.20$. For 90% confidence, $z_{\alpha/2} = z_{0.05} = 1.645$.

$$\text{Lower bound} = \hat{p} - z_{0.05} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.20 - 1.645 \cdot \sqrt{\frac{0.2(1-0.2)}{150}} \approx 0.20 - 0.054 = 0.146$$

$$\text{Upper bound} = \hat{p} + z_{0.05} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.20 + 1.645 \cdot \sqrt{\frac{0.2(1-0.2)}{150}} \approx 0.20 + 0.054 = 0.254$$

6. $\hat{p} = \frac{x}{n} = \frac{80}{200} = 0.40$. For 98% confidence, $z_{\alpha/2} = z_{.01} = 2.33$.

$$\text{Lower bound} = \hat{p} - z_{0.01} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.40 - 2.33 \cdot \sqrt{\frac{0.4(1-0.4)}{200}} \approx 0.40 - 0.081 = 0.319$$

$$\text{Upper bound} = \hat{p} + z_{0.01} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.40 + 2.33 \cdot \sqrt{\frac{0.4(1-0.4)}{200}} \approx 0.40 + 0.081 = 0.481$$

7. $\hat{p} = \frac{x}{n} = \frac{120}{500} = 0.24$. For 99% confidence, $z_{\alpha/2} = z_{0.005} = 2.575$.

$$\begin{aligned}\text{Lower bound } &= \hat{p} - z_{0.005} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \\ &= 0.24 - 2.575 \cdot \sqrt{\frac{0.24(1-0.24)}{500}} \\ &\approx 0.24 - 0.049 = 0.191\end{aligned}$$

$$\begin{aligned}\text{Upper bound } &= \hat{p} + z_{0.005} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \\ &= 0.24 + 2.575 \cdot \sqrt{\frac{0.24(1-0.24)}{500}} \\ &\approx 0.24 + 0.049 = 0.289\end{aligned}$$

8. $\hat{p} = \frac{x}{n} = \frac{400}{1200} = 0.333$. For 95% confidence, $z_{\alpha/2} = z_{0.025} = 1.96$.

Lower bound

$$\begin{aligned}&= \hat{p} - z_{0.025} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \\ &= 0.333 - 1.96 \cdot \sqrt{\frac{0.333(1-0.333)}{1200}} \\ &\approx 0.333 - 0.027 = 0.306 \text{ [Tech: 0.307]}\end{aligned}$$

Upper bound

$$\begin{aligned}&= \hat{p} + z_{0.025} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \\ &= 0.333 + 1.96 \cdot \sqrt{\frac{0.333(1-0.333)}{1200}} \\ &\approx 0.333 + 0.027 = 0.360\end{aligned}$$

9. $\hat{p} = \frac{x}{n} = \frac{860}{1100} \approx 0.782$. For 94% confidence, $z_{\alpha/2} = z_{.03} = 1.88$.

$$\begin{aligned}\text{Lower bound } &= \hat{p} - z_{0.03} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \\ &= 0.782 - 1.88 \cdot \sqrt{\frac{0.782(1-0.782)}{1100}} \\ &\approx 0.782 - 0.023 \approx 0.759\end{aligned}$$

$$\begin{aligned}\text{Upper bound } &= \hat{p} + z_{0.03} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \\ &= 0.782 + 1.88 \cdot \sqrt{\frac{0.782(1-0.782)}{1100}} \\ &\approx 0.782 + 0.023 \approx 0.805\end{aligned}$$

10. $\hat{p} = \frac{x}{n} = \frac{540}{900} = 0.60$. For 96% confidence, $z_{\alpha/2} = z_{0.02} = 2.05$

$$\begin{aligned}\text{Lower bound } &= \hat{p} - z_{0.02} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \\ &= 0.60 - 2.05 \cdot \sqrt{\frac{0.6(1-0.6)}{900}} \\ &\approx 0.60 - 0.033 \\ &= 0.567 \text{ [Tech: 0.566]}\end{aligned}$$

$$\begin{aligned}\text{Upper bound } &= \hat{p} + z_{0.02} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \\ &= 0.60 + 2.05 \cdot \sqrt{\frac{0.6(1-0.6)}{900}} \\ &\approx 0.60 + 0.033 \\ &= 0.633 \text{ [Tech: 0.634]}\end{aligned}$$

11. $\hat{p} = \frac{0.249 + 0.201}{2} = \frac{0.45}{2} = 0.225$

$E = \frac{0.249 - 0.201}{2} = \frac{0.048}{2} = 0.024$

$x = n \cdot \hat{p} = 1200(0.225) = 270$

12. $\hat{p} = \frac{0.074 + 0.051}{2} = \frac{0.125}{2} = 0.0625$

$E = \frac{0.074 - 0.051}{2} = \frac{0.023}{2} = 0.0115$

$x = n \cdot \hat{p} = 1120(0.0625) = 70$

13. $\hat{p} = \frac{0.509 + 0.462}{2} = \frac{0.971}{2} = 0.4855$

$E = \frac{0.509 - 0.462}{2} = \frac{0.047}{2} = 0.0235$

$x = n \cdot \hat{p} = 1680(0.4855) = 815.64 \to 816$

14. $\hat{p} = \frac{0.871 + 0.853}{2} = \frac{1.724}{2} = 0.862$

$E = \frac{0.871 - 0.853}{2} = \frac{0.018}{2} = 0.009$

$x = n \cdot \hat{p} = 10,732(0.862) = 9250.984 \to 9251$

15. **(a)** Flawed; no interval has been provided about the population proportion.

(b) Flawed; this interpretation indicates that the level of confidence is varying.

(c) This is the correct interpretation.

(d) Flawed; this interpretation suggests that this interval sets the standard for all the other intervals, which is not true.

16. **(a)** This is the correct interpretation.

(b) Flawed; this interpretation indicates that the level of confidence is varying.

(c) Flawed; this interpretation suggest that this interval sets the standard for all the other intervals, which is not true.

(d) Flawed; no interval has been provided about the population proportion.

17. Rasmussen Reports is 95% confident that the population proportion of adult Americans who dread Valentine's Day is between 0.135 and 0.225.

18. Gallup is 95% confident that the population proportion of adult Americans who consider their commute to work very or somewhat stressful is between 0.20 and 0.28.

19. **(a)** $\hat{p} = \frac{x}{n} = \frac{47}{863} \approx 0.054$

(b) The sample size $n = 863$ is less than 5% of the population, and $n\hat{p}(1-\hat{p}) \approx 44.44 \geq 10$.

(c) For 90% confidence, $z_{\alpha/2} = z_{0.05} = 1.645$.

Lower bound

$$= \hat{p} - z_{0.05} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$= 0.054 - 1.645 \cdot \sqrt{\frac{0.054(1-0.054)}{863}}$$

$$\approx 0.054 - 0.013$$

$$= 0.041 \text{ [Tech: 0.042]}$$

Upper bound

$$= \hat{p} + z_{0.05} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$= 0.054 + 1.645 \cdot \sqrt{\frac{0.054(1-0.054)}{863}}$$

$$\approx 0.054 + 0.013 = 0.067$$

(d) We are 90% confident that the population proportion of Lipitor users who will have a headache as a side effect is between 0.041 and 0.067 (i.e., between 4.1% and 6.7%).

20. **(a)** $\hat{p} = \frac{x}{n} = \frac{58}{74} \approx 0.784$

(b) The sample size $n = 74$ is less than 5% of the population, and $n\hat{p}(1-\hat{p}) \approx 12.54 \geq 10$.

(c) For 99% confidence, $z_{\alpha/2} = z_{.005} = 2.575$.

Lower bound

$$= \hat{p} - z_{0.005} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$= 0.784 - 2.575 \cdot \sqrt{\frac{0.784(1-0.784)}{74}}$$

$$\approx 0.784 - 0.123 = 0.661$$

Upper bound

$$= \hat{p} + z_{0.005} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$= 0.784 + 2.575 \cdot \sqrt{\frac{0.784(1-0.784)}{74}}$$

$$\approx 0.784 + 0.123 = 0.907$$

(d) We are 99% confident that the population proportion of Pepcid users with ulcers who will experience ulcer healing is between 0.661 and 0.907 (i.e., between 66.1% and 90.7%).

21. **(a)** $\hat{p} = \frac{x}{n} = \frac{1322}{1979} \approx 0.668$

(b) The sample size $n = 1979$ is less than 5% of the population, and $n\hat{p}(1-\hat{p}) \approx 438.89 \geq 10$.

(c) For 96% confidence, $z_{\alpha/2} = z_{.02} = 2.05$.

Lower bound

$$= \hat{p} - z_{0.02} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$= 0.668 - 2.05 \cdot \sqrt{\frac{0.668(1-0.668)}{1979}}$$

$$\approx 0.668 - 0.022 = 0.646$$

Upper bound

$$= \hat{p} + z_{0.02} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$= 0.668 + 2.05 \cdot \sqrt{\frac{0.668(1-0.668)}{1979}}$$

$$\approx 0.668 + 0.022 = 0.690$$

(d) Yes; it is possible that the population proportion is below 60%, because it is possible that the true proportion is not captured in the confidence interval. Since we are 96% confident that the true proportion is between 0.646 and 0.690, it is not likely that the proportion of adult Americans who believe that traditional journalism is out of touch is below 60%.

22. **(a)** $\hat{p} = \frac{x}{n} = \frac{768}{1024} = 0.75$

(b) The sample size $n = 1024$ is less than 5% of the population, and $n\hat{p}(1-\hat{p}) = 192 \geq 10$.

(c) For 98% confidence, $z_{\alpha/2} = z_{.01} = 2.33$.
Lower bound

$$= \hat{p} - z_{0.01} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$
$$= 0.75 - 2.33 \cdot \sqrt{\frac{0.75(1-0.25)}{1024}}$$
$$\approx 0.75 - 0.032 = 0.718 \text{ [Tech: 0.719]}$$

Upper bound

$$= \hat{p} + z_{0.01} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$
$$= 0.75 + 2.33 \cdot \sqrt{\frac{0.75(1-0.25)}{1024}}$$
$$\approx 0.75 + 0.032 = 0.782 \text{ [Tech: 0.781]}$$

(d) Yes; it is possible that the population proportion is below 65%, because it is possible that the true proportion is not captured in the confidence interval. Since we are 98% confident that the true proportion is between 0.718 and 0.782, it is not likely that the proportion of adult Americans aged 18 or older for which the issue of family values is extremely or very important is below 65%.

23. **(a)** $\hat{p} = \frac{x}{n} = \frac{322}{2302} \approx 0.140$

(b) The sample size $n = 2302$ is less than 5% of the population, and $n\hat{p}(1-\hat{p}) \approx 277 \geq 10$.

(c) For 90% confidence, $z_{\alpha/2} = z_{0.05} = 1.645$.
Lower bound

$$= \hat{p} - z_{0.05} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$
$$= 0.140 - 1.645 \cdot \sqrt{\frac{0.140(1-0.140)}{2302}}$$
$$\approx 0.140 - 0.012 = 0.128$$

Upper bound

$$= \hat{p} + z_{0.05} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$
$$= 0.140 + 1.645 \cdot \sqrt{\frac{0.140(1-0.140)}{2302}}$$
$$\approx 0.140 + 0.012 = 0.152$$

(d) For 98% confidence, $z_{\alpha/2} = z_{.01} = 2.33$.
Lower bound

$$= \hat{p} - z_{0.01} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$
$$= 0.140 - 2.33 \cdot \sqrt{\frac{0.140(1-0.140)}{2302}}$$
$$\approx 0.140 - 0.017 = 0.123$$

Upper bound

$$= \hat{p} + z_{0.01} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$
$$= 0.140 + 2.33 \cdot \sqrt{\frac{0.140(1-0.140)}{2302}}$$
$$\approx 0.140 + 0.017 = 0.157$$

(e) Increasing the level of confidence widens the interval.

24. **(a)** $\hat{p} = \frac{x}{n} = \frac{604}{1010} \approx 0.598$

(b) The sample size $n = 1010$ is less than 5% of the population, and $n\hat{p}(1-\hat{p}) \approx 243 \geq 10$.

(c) For 95% confidence, $z_{\alpha/2} = z_{0.025} = 1.96$.
Lower bound

$$= \hat{p} - z_{0.05} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$
$$= 0.598 - 1.96 \cdot \sqrt{\frac{0.598(1-0.598)}{1010}}$$
$$\approx 0.598 - 0.030 = 0.568$$

Upper bound

$$= \hat{p} + z_{0.05} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$
$$= 0.598 + 1.96 \cdot \sqrt{\frac{0.598(1-0.598)}{1010}}$$
$$\approx 0.598 + 0.030 = 0.628$$

(d) For 99% confidence, $z_{\alpha/2} = z_{.005} = 2.575$.
Lower bound

$$= \hat{p} - z_{0.005} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$
$$= 0.598 - 2.575 \cdot \sqrt{\frac{0.598(1-0.598)}{1010}}$$
$$\approx 0.598 - 0.040 = 0.558$$

Upper bound

$$= \hat{p} + z_{0.005} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$
$$= 0.598 + 2.575 \cdot \sqrt{\frac{0.598(1-0.598)}{1010}}$$
$$\approx 0.598 + 0.040 = 0.638$$

(e) Increasing the level of confidence widens the interval.

25. For 99% confidence, $z_{\alpha/2} = z_{.005} = 2.575$.

(a) Using $\hat{p} = 0.69$,

$$n = \hat{p}(1-\hat{p})\left(\frac{z_{\alpha/2}}{E}\right)^2$$
$$= 0.69(1-0.69)\left(\frac{2.575}{0.03}\right)^2$$
$$\approx 1575.9$$

which we must increase to 1576. The researcher needs a sample size of 1576.

(b) Without using a prior estimate,

$$n = 0.25\left(\frac{z_{\alpha/2}}{E}\right)^2 = 0.25\left(\frac{2.575}{0.03}\right)^2$$
$$\approx 1841.8$$

which we must increase to 1842. Without the prior estimate, the researcher would need a sample size of 1842.

26. For 90% confidence, $z_{\alpha/2} = z_{.05} = 1.645$.

(a) Using $\hat{p} = 0.681$,

$$n = \hat{p}(1-\hat{p})\left(\frac{z_{\alpha/2}}{E}\right)^2$$
$$= 0.681(1-0.681)\left(\frac{1.645}{0.02}\right)^2$$
$$\approx 1469.6$$

which we must increase to 1470. The economist needs a sample size of 1470.

(b) Without using a prior estimate,

$$n = 0.25\left(\frac{z_{\alpha/2}}{E}\right)^2 = 0.25\left(\frac{1.645}{0.02}\right)^2$$
$$\approx 1691.3$$

which we must increase to 1692. Without the prior estimate, the economist would need a sample size of 1692.

27. For 98% confidence, $z_{\alpha/2} = z_{.01} = 2.33$.

(a) Using $\hat{p} = 0.15$,

$$n = \hat{p}(1-\hat{p})\left(\frac{z_{\alpha/2}}{E}\right)^2$$
$$= 0.15(1-0.15)\left(\frac{2.33}{0.02}\right)^2$$
$$\approx 1730.5$$

which we must increase to 1731. The researcher needs a sample size of 1731.

(b) Without using a prior estimate,

$$n = 0.25\left(\frac{z_{\alpha/2}}{E}\right)^2 = 0.25\left(\frac{2.33}{0.02}\right)^2$$
$$\approx 3393.1$$

which we must increase to 3394. Without the prior estimate, the researcher would need a sample size of 3394.

28. For 94% confidence, $z_{\alpha/2} = z_{0.03} = 1.88$.

(a) Using $\hat{p} = 0.34$,

$$n = \hat{p}(1-\hat{p})\left(\frac{z_{\alpha/2}}{E}\right)^2$$
$$= 0.34(1-0.34)\left(\frac{1.88}{0.025}\right)^2$$
$$\approx 1268.99$$

which we must increase to 1269. The administrator needs a sample size of 1269.

(b) $n = 0.25\left(\frac{z_{\alpha/2}}{E}\right)^2 = 0.25\left(\frac{1.88}{0.025}\right)^2$

$$\approx 1413.8$$

which we must increase to 1414. Without the prior estimate, the administrator would need a sample size of 1414.

29. For 95% confidence, $z_{\alpha/2} = z_{0.025} = 1.96$.

(a) Using $\hat{p} = 0.48$,

$$n = \hat{p}(1-\hat{p})\left(\frac{z_{\alpha/2}}{E}\right)^2$$
$$= 0.48(1-0.48)\left(\frac{1.96}{0.03}\right)^2$$
$$\approx 1065.4$$

which we must increase to 1066. The commentator needs a sample size of 1066.

(b) $n = 0.25\left(\frac{z_{\alpha/2}}{E}\right)^2 = 0.25\left(\frac{1.96}{0.03}\right)^2$

$$\approx 1067.1$$

which we must increase to 1068. Without the prior estimate, the commentator would need a sample size of 1068.

(c) The results are close because $0.48(1-0.48) = 0.2496$ is very close to 0.25. That is, the prior estimate is very close the conservative estimate of 0.5 that we use when no estimate is available.

30. For 90% confidence, $z_{\alpha/2} = z_{0.05} = 1.645$.

(a) Using $\hat{p} = 0.55$,

$$n = \hat{p}(1-\hat{p})\left(\frac{z_{\alpha/2}}{E}\right)^2$$
$$= 0.55(1-0.55)\left(\frac{1.645}{0.04}\right)^2$$
$$\approx 418.6$$

which we must increase to 419. The sociologist needs a sample size of 419.

(b) $n = 0.25\left(\frac{z_{\alpha/2}}{E}\right)^2 = 0.25\left(\frac{1.645}{0.04}\right)^2$
≈ 422.8
which we must increase to 423. Without the prior estimate, the sociologist needs a sample size of 423.

(c) The results are close because $0.55(1-0.55) = 0.2475$ is very close to 0.25. That is, the prior estimate is very close the conservative estimate of 0.5 that we use when no estimate is available.

31. For 95% confidence, $z_{\alpha/2} = z_{0.025} = 1.96$. Using $E = 0.03$ and $\hat{p} = 0.69$, we get

$$n = \hat{p}(1-\hat{p})\left(\frac{z_{\alpha/2}}{E}\right)^2$$
$$= 0.69(1-0.69)\left(\frac{1.96}{0.03}\right)^2$$
$$\approx 913.02$$

which we must increase to 914. At least 914 people were included in the survey.

32. For 95% confidence, $z_{\alpha/2} = z_{0.025} = 1.96$. Using $E = 0.035$ and $\hat{p} = 0.61$, we get

$$n = \hat{p}(1-\hat{p})\left(\frac{z_{\alpha/2}}{E}\right)^2$$
$$= 0.61(1-0.61)\left(\frac{1.96}{0.035}\right)^2$$
$$\approx 746.05$$

which we must increase to 747. At least 747 people were surveyed.

33. Answers may vary. One possibility follows: The confidence interval for the percentage intending to vote for George Bush was $(49-3,\ 49+3) = (46\%,\ 52\%)$. Likewise, the confidence interval for the percentage intending to vote for John Kerry was $(47-3,\ 47+3) = (44\%,\ 50\%)$. Since these intervals overlap, it is possible that the true percentage intending to vote for John Kerry could have been greater than the true percentage intending to vote for George Bush, or vice versa. Hence, the result was too close to call.

34. **(a)** Answers will vary. We would expect about 0.95 (i.e., 95%) of the intervals to contain the population proportion.

(b) Answers will vary.

(c) Answers will vary. We would expect about 0.99 (i.e., 99%) of the intervals to contain the population proportion.

35. **(a), (b), (c)** Answers will vary.

(d) As the sample size n increases, the proportion of intervals that capture p gets closer and closer to the level of confidence.

36. **(a)** $\hat{p} = \frac{x}{n} = \frac{921}{1001} \approx 0.920$

(b) The sample size $n = 1001$ is less than 5% of the population, and $n\hat{p}(1-\hat{p}) \approx 73.7 \geq 10$.

(c) For 95% confidence, $z_{\alpha/2} = z_{0.025} = 1.96$.

Lower bound

$$= \hat{p} - z_{0.05} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$
$$= 0.920 - 1.96 \cdot \sqrt{\frac{0.920(1-0.920)}{1001}}$$
$$\approx 0.920 - 0.017 = 0.903$$

Upper bound

$$= \hat{p} + z_{0.05} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$
$$= 0.920 + 1.96 \cdot \sqrt{\frac{0.920(1-0.920)}{1001}}$$
$$\approx 0.920 + 0.017 = 0.937$$

(d) Yes; it is possible that the population proportion is below 85%, because it is possible that the true proportion is not captured in the confidence interval. Since we are 95% confident that the true proportion is between 0.903 and 0.937, it is not likely that the proportion of adults who say they always wash their hands in public rest rooms is less than 85%.

(e) $\hat{p} = \frac{x}{n} = \frac{4679}{6076} \approx 0.770$

(f) The sample size $n = 6076$ is less than 5% of the population, and $n\hat{p}(1-\hat{p}) \approx 1076 \geq 10$.

(g) For 95% confidence, $z_{\alpha/2} = z_{0.025} = 1.96$.
Lower bound
$$= \hat{p} - z_{0.05} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$
$$= 0.770 - 1.96 \cdot \sqrt{\frac{0.770(1-0.770)}{6076}}$$
$$\approx 0.770 - 0.011 = 0.759$$
Upper bound
$$= \hat{p} + z_{0.05} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$
$$= 0.770 + 1.96 \cdot \sqrt{\frac{0.770(1-0.770)}{6076}}$$
$$\approx 0.770 + 0.011 = 0.781$$

(h) Yes; it is possible that the population proportion is greater than 85%, because it is possible that the true proportion is not captured in the confidence interval. Since we are 95% confident that the true proportion is between 0.759 and 0.781, it is not likely that the proportion of adults who wash their hands in public rest rooms is greater than 85%.

(i) The proportion who say they wash their hands in public restrooms is higher than the proportion who actually do. Explanations may vary. One possibility is that people lie about their hand washing habits out of embarrassment.

9.4 Confidence Intervals for a Population Standard Deviation

1. The chi-squared distribution is not symmetric but is skewed to the right; the actual shape of the chi-squared distribution depends on the degrees of freedom. As the number of degrees of freedom increases, the chi-squared distribution become more nearly symmetric. The values of χ^2 are nonnegative.

2. To construct a confidence interval about a population standard deviation, the sample must come from a normally distributed population.

3. After the population is shown to be normal, a confidence interval for the standard deviation is obtained by as follows:
Step 1: Compute the sample variance.
Step 2: Determine the critical values using the desired confidence level, the correct degrees of freedom, and the χ^2 distribution table.
Step 3: Construct the confidence interval for the population variance by using the formulas Lower bound $= \frac{(n-1)s^2}{\chi^2_{\alpha/2}}$ and Upper bound $= \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}$
Step 4: Compute the square root of the lower bound and the upper bound to find the confidence interval for the population standard deviation.

4. No; the confidence interval for a population standard deviation is not symmetric.

5. df $= n-1 = 19$ and, for 90% confidence, $\alpha/2 = 0.05$. From the table, $\chi^2_{0.05} = 30.144$ and $\chi^2_{1-0.05} = \chi^2_{0.95} = 10.117$.

6. df $= n-1 = 24$ and, for 95% confidence, $\alpha/2 = 0.025$. From the table, $\chi^2_{0.025} = 39.364$ and $\chi^2_{1-0.025} = \chi^2_{0.975} = 12.401$.

7. df $= n-1 = 22$ and, for 98% confidence, $\alpha/2 = 0.01$. From the table, $\chi^2_{0.01} = 40.289$ and $\chi^2_{1-0.01} = \chi^2_{0.99} = 9.542$.

8. df $= n-1 = 13$ and, for 99% confidence, $\alpha/2 = 0.005$. From the table, $\chi^2_{0.005} = 29.819$ and $\chi^2_{1-0.005} = \chi^2_{0.995} = 3.565$.

9. **(a)** df $= n-1 = 19$ and, for 90% confidence, $\alpha/2 = 0.05$.
From the table, $\chi^2_{0.05} = 30.144$ and $\chi^2_{1-0.05} = \chi^2_{0.95} = 10.117$.
Lower bound $= \frac{(n-1)s^2}{\chi^2_{\alpha/2}} = \frac{19 \cdot 12.6}{30.144} \approx 7.94$
Upper bound $= \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}} = \frac{19 \cdot 12.6}{10.117} \approx 23.66$

(b) $df = n-1 = 29$ and, for 90% confidence, $\alpha/2 = 0.05$.

From the table, $\chi^2_{0.05} = 42.557$ and $\chi^2_{1-0.05} = \chi^2_{0.95} = 17.708$.

$$\text{Lower bound} = \frac{(n-1)s^2}{\chi^2_{\alpha/2}} = \frac{29 \cdot 12.6}{42.557} \approx 8.59$$

$$\text{Upper bound} = \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}} = \frac{29 \cdot 12.6}{17.708} \approx 20.63$$

Increasing the sample size decreases the width of the confidence interval.

(c) $df = n-1 = 19$ and, for 98% confidence, $\alpha/2 = 0.01$.

From the table, $\chi^2_{0.05} = 36.191$ and $\chi^2_{1-0.01} = \chi^2_{0.99} = 7.633$.

$$\text{Lower bound} = \frac{(n-1)s^2}{\chi^2_{\alpha/2}} = \frac{19 \cdot 12.6}{36.191} \approx 6.61$$

$$\text{Upper bound} = \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}} = \frac{19 \cdot 12.6}{7.633} \approx 31.36$$

Increasing the level of confidence increases the width of the confidence interval.

10. (a) $df = n-1 = 9$ and, for 95% confidence, $\alpha/2 = 0.025$.

From the table, $\chi^2_{0.025} = 19.023$ and $\chi^2_{1-0.025} = \chi^2_{0.975} = 2.700$.

$$\text{Lower bound} = \frac{(n-1)s^2}{\chi^2_{\alpha/2}} = \frac{9 \cdot 19.8}{19.023} \approx 9.37$$

$$\text{Upper bound} = \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}} = \frac{9 \cdot 19.8}{2.700} \approx 66.00$$

(b) $df = n-1 = 24$ and, for 95% confidence, $\alpha/2 = 0.025$.

From the table, $\chi^2_{0.025} = 39.364$ and $\chi^2_{1-0.025} = \chi^2_{0.975} = 12.401$.

$$\text{Lower bound} = \frac{(n-1)s^2}{\chi^2_{\alpha/2}} = \frac{24 \cdot 19.8}{39.364} \approx 12.07$$

$$\text{Upper bound} = \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}} = \frac{24 \cdot 19.8}{12.401} \approx 38.32$$

Increasing the sample size decreases the width of the confidence interval.

(c) $df = n-1 = 9$ and, for 99% confidence, $\alpha/2 = 0.005$.

From the table, $\chi^2_{0.05} = 23.589$ and $\chi^2_{1-0.005} = \chi^2_{0.995} = 1.735$.

$$\text{Lower bound} = \frac{(n-1)s^2}{\chi^2_{\alpha/2}} = \frac{9 \cdot 19.8}{23.589} \approx 7.55$$

$$\text{Upper bound} = \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}} = \frac{9 \cdot 19.8}{1.735} \approx 102.71$$

Increasing the level of confidence increases the width of the confidence interval.

11. $df = n-1 = 12-1 = 11$ and, for 95% confidence, $\alpha/2 = 0.025$.

From the table, $\chi^2_{0.025} = 21.920$ and $\chi^2_{1-0.025} = \chi^2_{0.975} = 3.816$.

$$\text{Lower bound} = \sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2}}} = \sqrt{\frac{11 \cdot (10.00)^2}{21.920}} \approx 7.08 \text{ weeks}$$

$$\text{Upper bound} = \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}} = \sqrt{\frac{11 \cdot (10.00)^2}{3.816}} \approx 16.98 \text{ weeks}$$

Essential Baby can be 95% confident that the population standard deviation of the time at which babies first crawl is between 7.08 and 16.98 weeks.

12. $df = n-1 = 10-1 = 9$ and, for 95% confidence, $\alpha/2 = 0.025$.

From the table, $\chi^2_{0.025} = 19.023$ and $\chi^2_{1-0.025} = \chi^2_{0.975} = 2.700$.

$$\text{Lower bound} = \sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2}}} = \sqrt{\frac{9 \cdot (2.14)^2}{19.023}} \approx 1.47 \text{ hours}$$

$$\text{Upper bound} = \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}} = \sqrt{\frac{9 \cdot (2.14)^2}{2.700}} \approx 3.91 \text{ hours}$$

We can be 95% confident that the population standard deviation of the battery life of a fifth-generation iPod music player is between 1.47 and 3.91 hours.

13. df $= n-1 = 10-1 = 9$ and, for 99% confidence, $\alpha/2 = 0.005$.

From the table, $\chi^2_{0.005} = 23.589$ and $\chi^2_{1-0.005} = \chi^2_{0.995} = 1.735$.

$$\text{Lower bound} = \sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2}}} = \sqrt{\frac{9\cdot(21.88)^2}{23.589}} \approx 13.51 \text{ days}$$

$$\text{Upper bound} = \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}} = \sqrt{\frac{9\cdot(21.88)^2}{1.735}} \approx 49.83 \text{ days}$$

The agricultural researcher can be 99% confident that the population standard deviation of the growing season in Chicago is between 13.51 and 49.83 days.

14. df $= n-1 = 12-1 = 11$ and, for 99% confidence, $\alpha/2 = 0.005$.

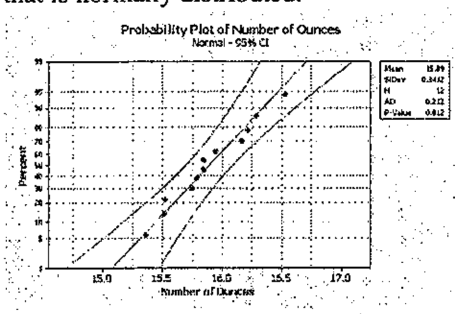
From the table, $\chi^2_{0.005} = 26.757$ and $\chi^2_{1-0.005} = \chi^2_{0.995} = 2.603$.

$$\text{Lower bound} = \sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2}}} = \sqrt{\frac{11\cdot(1.97)^2}{26.757}} \approx 1.26 \text{ grams}$$

$$\text{Upper bound} = \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}} = \sqrt{\frac{11\cdot(1.97)^2}{2.603}} \approx 4.05 \text{ grams}$$

We can be 99% confident that the population standard deviation of the disc weight of the distance drivers produced by this manufacturer is between 1.26 and 4.05 grams.

15. **(a)** From the probability plot shown below, the data appear to be from a population that is normally distributed.



(b) Using technology, we obtain $s \approx 0.349$ oz.

(c) df $= n-1 = 11$ and, for 90% confidence, $\alpha/2 = 0.05$. From the table, $\chi^2_{1-0.05} = \chi^2_{0.95} = 4.575$ and $\chi^2_{0.05} = 19.675$.

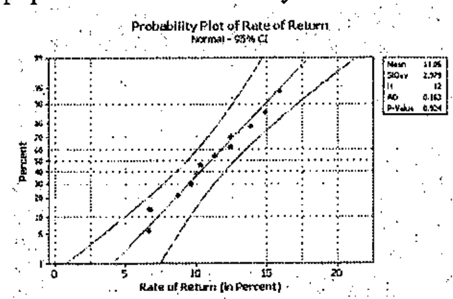
$$\text{Lower bound} = \sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2}}} = \sqrt{\frac{11(0.349)^2}{19.675}} \approx 0.261 \text{ oz}$$

$$\text{Upper bound} = \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}} = \sqrt{\frac{11(0.349)^2}{4.575}} \approx 0.541 \text{ oz}$$

The quality-control manager can be 90% confident that the population standard deviation of the number of ounces of peanuts is between 0.261 and 0.541 ounce.

(d) No; we are 90% confident that the population standard deviation is between 0.261 and 0.541 ounces, and so above 0.20 ounces.

16. **(a)** From the probability plot shown to the right, the data appear to be from a population that is normally distributed.



(b) Using technology, we obtain $s \approx 2.98$ %.

(c) df $= n-1 = 11$ and, for 95% confidence, $\alpha/2 = 0.025$. From the table, $\chi^2_{1-0.025} = \chi^2_{0.975} = 3.816$ and $\chi^2_{0.025} = 21.920$.

$$\text{Lower bound} = \sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2}}} = \sqrt{\frac{11(2.98)^2}{21.920}} \approx 2.11\%$$

$$\text{Upper bound} = \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}} = \sqrt{\frac{11(2.98)^2}{3.816}} \approx 5.06\%$$

The investment manager can be 95% confident that the population standard deviation of the rate of return is between 2.11% and 5.06%.

(d) Yes; we are 95% confident that the population standard deviation for the rate of return is between 2.11% and 5.06%, which is below 6%.

17. From the standard normal table, we get $z_{0.975} = -1.96$ and $z_{0.025} = 1.96$. With $\nu = 100$, we get

$$\chi^2_{0.975} \approx \frac{\left(z_{0.975} + \sqrt{2\nu - 1}\right)^2}{2} = \frac{\left(-1.96 + \sqrt{2 \cdot 100 - 1}\right)^2}{2} = 73.772$$

(compared to the table's value of 74.222)

$$\text{and } \chi^2_{0.025} \approx \frac{\left(z_{0.025} + \sqrt{2\nu - 1}\right)^2}{2} = \frac{\left(1.96 + \sqrt{2 \cdot 100 - 1}\right)^2}{2} = 129.070$$

(compared to the table's value of 129.561).

9.5 Putting It All Together: Which Procedure Do I Use?

1. We construct a t-interval when we are estimating the population mean, we do not know the population standard deviation, and the underlying population is normally distributed. If the underlying population is not normally distributed, we can construct a t-interval to estimate the population mean provide the sample size is large ($n \geq 30$).

   We construct a Z-interval when we are estimating the population mean, we know the population standard deviation, and the underlying population is normally distributed. If the underlying population is not normally distributed, we can construct a Z-interval to estimate the population mean provided the sample size is large ($n \geq 30$). We also construct a Z-interval when we are estimating the population proportion, provided the sample is smaller than 5% of the population and $n\hat{p}(1-\hat{p}) \geq 10$.

2. To construct a confidence interval about a population proportion, the sample must come from a normally distributed population with no outliers, the sample size n must be no more than 5% of the population size N, and $n\hat{p}(1-\hat{p}) \geq 10$.

3. We construct a Z-interval because we are estimating a population mean, we know the population standard deviation, and the underlying population is normally distributed. For 95% confidence the critical value is $z_{0.025} = 1.96$. Then:

$$\text{Lower bound} = \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} = 60 - 1.96 \cdot \frac{20}{\sqrt{14}} \approx 49.5$$

$$\text{Upper bound} = \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} = 60 + 1.96 \cdot \frac{20}{\sqrt{14}} \approx 70.5$$

4. We construct a Z-interval because we are estimating a population mean, we know the population standard deviation, and the underlying population is normally distributed. For 90% confidence the critical value is $z_{0.05} = 1.645$. Then:

$$\text{Lower bound} = \bar{x} - z_{0.05} \cdot \frac{\sigma}{\sqrt{n}} = 122.5 - 1.645 \cdot \frac{37}{\sqrt{22}} \approx 109.52$$

$$\text{Upper bound} = \bar{x} + z_{0.05} \cdot \frac{\sigma}{\sqrt{n}} = 122.5 + 1.645 \cdot \frac{37}{\sqrt{22}} \approx 135.48$$

5. $\hat{p} = \frac{35}{300} \approx 0.117$. The sample size $n = 300$ is less than 5% of the population, and $n\hat{p}(1-\hat{p}) \approx 31.0 \geq 10$, so we can construct a Z-interval. For 99% confidence the critical value is $z_{0.005} = 2.575$. Then:

   Lower bound

$$= \hat{p} - z_{0.005} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.117 - 2.575 \cdot \sqrt{\frac{0.117(1-0.117)}{300}} \approx 0.117 - 0.048 = 0.069$$

Upper bound

$= \hat{p} + z_{0.005} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$= 0.117 + 2.575 \cdot \sqrt{\frac{0.117(1-0.117)}{300}}$

$\approx 0.117 + 0.048 = 0.165$ [Tech: 0.164]

6. $\hat{p} = \frac{275}{785} \approx 0.350$. The sample size $n = 785$ is less than 5% of the population, and $n\hat{p}(1-\hat{p}) \approx 178.6 \geq 10$, so we can construct a Z-interval. For 95% confidence the critical value is $z_{0.025} = 1.96$. Then:

Lower bound

$= \hat{p} - z_{0.025} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$= 0.350 - 1.96 \cdot \sqrt{\frac{0.350(1-0.350)}{785}}$

$\approx 0.350 - 0.033 = 0.317$

Upper bound

$= \hat{p} + z_{0.025} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$= 0.350 + 1.96 \cdot \sqrt{\frac{0.350(1-0.350)}{785}}$

$\approx 0.350 + 0.033 = 0.383$ [Tech: 0.384]

7. We construct a t-interval because we are estimating the population mean, we do not know the population standard deviation, and the underlying population is normally distributed. For 90% confidence, $\alpha/2 = 0.05$. Since $n = 12$, then df = 11 and $t_{0.05} = 1.796$. Then:

Lower bound $= \bar{x} - t_{0.05} \cdot \frac{s}{\sqrt{n}}$

$= 45 - 1.796 \cdot \frac{14}{\sqrt{12}}$

$\approx 45 - 7.26 = 37.74$

Upper bound $= \bar{x} + t_{0.05} \cdot \frac{s}{\sqrt{n}}$

$= 45 + 1.796 \cdot \frac{14}{\sqrt{12}}$

$\approx 45 + 7.26 = 52.26$

8. We construct a t-interval because we are estimating the population mean, we do not know the population standard deviation, and the underlying population is normally distributed. For 95% confidence, $\alpha/2 = 0.025$. Since $n = 17$, then df = 16 and $t_{0.025} = 2.120$. Then:

Lower bound $= \bar{x} - t_{0.025} \cdot \frac{s}{\sqrt{n}}$

$= 3.25 - 2.120 \cdot \frac{1.17}{\sqrt{17}}$

$\approx 3.25 - 0.60 = 2.65$

Upper bound $= \bar{x} + t_{0.025} \cdot \frac{s}{\sqrt{n}}$

$= 3.25 + 2.120 \cdot \frac{1.17}{\sqrt{17}}$

$\approx 3.25 + 0.60 = 3.85$

9. We construct a t-interval because we are estimating the population mean, we do not know the population standard deviation, and the underlying population is normally distributed. For 99% confidence, $\alpha/2 = 0.005$. Since $n = 40$, then df = 39 and $t_{0.050} = 2.708$. Then:

Lower bound $= \bar{x} - t_{0.005} \cdot \frac{s}{\sqrt{n}}$

$= 120.5 - 2.708 \cdot \frac{12.9}{\sqrt{40}}$

$\approx 120.5 - 5.52 = 114.98$

Upper bound $= \bar{x} + t_{0.005} \cdot \frac{s}{\sqrt{n}}$

$= 120.5 + 2.708 \cdot \frac{12.9}{\sqrt{40}}$

$\approx 120.5 + 5.52 = 126.02$

10. We construct a t-interval because we are estimating the population mean, we do not know the population standard deviation, and the underlying population is normally distributed. For 90% confidence, $\alpha/2 = 0.05$. Since $n = 210$, then df = 209. There is no row in the table for 209 degrees of freedom, so we use df = 100 instead. The critical value is $t_{0.05} = 1.660$.

Lower bound $= \bar{x} - t_{0.05} \cdot \frac{s}{\sqrt{n}}$

$= 20.1 - 1.660 \cdot \frac{3.2}{\sqrt{210}}$

$\approx 20.1 - 0.37$

$= 19.73$ [Tech: 19.74]

Upper bound $= \bar{x} + t_{0.05} \cdot \frac{s}{\sqrt{n}}$

$= 20.1 + 1.660 \cdot \frac{3.2}{\sqrt{210}}$

$\approx 20.1 + 0.37$

$= 20.47$ [Tech: 20.46]

11. $\text{df} = n-1 = 12-1 = 11$ and, for 90% confidence, $\alpha/2 = 0.05$.
From the table, $\chi^2_{0.05} = 19.675$ and $\chi^2_{1-0.05} = \chi^2_{0.95} = 4.575$.

$$\text{Lower bound} = \frac{(n-1)s^2}{\chi^2_{\alpha/2}} = \frac{11 \cdot 23.7}{19.675} \approx 13.25$$

$$\text{Upper bound} = \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}} = \frac{11 \cdot 23.7}{4.575} \approx 56.98$$

12. $\text{df} = n-1 = 25-1 = 24$ and, for 95% confidence, $\alpha/2 = 0.025$.
From the table, $\chi^2_{0.025} = 39.364$ and $\chi^2_{1-0.025} = \chi^2_{0.975} = 12.401$.

$$\text{Lower bound} = \sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2}}} = \sqrt{\frac{24 \cdot 3.97}{39.364}} \approx 1.56$$

$$\text{Upper bound} = \sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2}}} = \sqrt{\frac{24 \cdot 3.97}{12.401}} \approx 2.77$$

13. We construct a t-interval because we are estimating the population mean, we do not know the population standard deviation, and the underlying population is normally distributed. For 95% confidence, $\alpha/2 = 0.025$. Since $n = 40$, then $\text{df} = 39$ and $t_{0.025} = 2.023$. Then:

$$\begin{aligned}\text{Lower bound} &= \bar{x} - t_{0.025} \cdot \frac{s}{\sqrt{n}} \\ &= 54 - 2.023 \cdot \frac{8}{\sqrt{40}} \\ &\approx 54 - 2.6 = 51.4 \text{ months}\end{aligned}$$

$$\begin{aligned}\text{Upper bound} &= \bar{x} + t_{0.025} \cdot \frac{s}{\sqrt{n}} \\ &= 54 + 2.023 \cdot \frac{8}{\sqrt{40}} \\ &\approx 54 + 2.6 = 56.6 \text{ months}\end{aligned}$$

We can be 95% confident that the population of felons convicted of aggravated assault serve a mean sentence between 51.4 and 56.6 months.

14. $\hat{p} = \frac{881}{1013} \approx 0.870$. The sample size $n = 1013$ is less than 5% of the population, and $n\hat{p}(1-\hat{p}) \approx 115 \geq 10$, so we can construct a Z-interval:
For 98% confidence the critical value is $z_{0.01} = 2.33$. Then:

$$\begin{aligned}\text{Lower bound} &= \hat{p} - z_{0.01} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \\ &= 0.870 - 2.33 \cdot \sqrt{\frac{0.870(1-0.870)}{1013}} \\ &\approx 0.870 - 0.025 = 0.845\end{aligned}$$

$$\begin{aligned}\text{Upper bound} &= \hat{p} + z_{0.01} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \\ &= 0.870 + 2.33 \cdot \sqrt{\frac{0.870(1-0.870)}{1013}} \\ &\approx 0.870 + 0.025 = 0.895 \quad \text{[Tech: 0.894]}\end{aligned}$$

The Harris Organization can be 98% confident that the population proportion of adults who always wear seatbelts is between 0.845 and 0.895 (i.e., between 84.5% and 89.5%).

15. We construct a Z-interval because we are estimating a population mean, we know the population standard deviation, and the underlying population is normally distributed. For 90% confidence the critical value is $z_{0.05} = 1.645$. Then:

$$\begin{aligned}\text{Lower bound} &= \bar{x} - z_{0.05} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 3421 - 1.645 \cdot \frac{2583}{\sqrt{100}} \\ &\approx 3421 - 424.9 = \$2996.1\end{aligned}$$

$$\begin{aligned}\text{Upper bound} &= \bar{x} + z_{0.05} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 3421 + 1.645 \cdot \frac{2583}{\sqrt{100}} \\ &\approx 3421 + 424.9 = \$3845.9\end{aligned}$$

The Internal Revenue Service can be 90% confident that the population mean additional tax owed is between \$2996.1 and \$3845.9.

16. We construct a t-interval because we are estimating the population mean, we do not know the population standard deviation, and the underlying population is normally distributed.
For 99% confidence, $\alpha/2 = 0.005$. Since $n = 50$, then $\text{df} = 49$. There is no row in the table for 49 degrees of freedom, so we use $\text{df} = 50$ instead. The critical value is $t_{0.005} = 2.678$.

$$\text{Lower bound} = \bar{x} - t_{0.005} \cdot \frac{s}{\sqrt{n}}$$
$$= 863 - 2.678 \cdot \frac{2.7}{\sqrt{50}}$$
$$\approx 863 - 1.02 = 861.98 \text{ m/s}$$
$$\text{Upper bound} = \bar{x} + t_{0.005} \cdot \frac{s}{\sqrt{n}}$$
$$= 863 + 2.678 \cdot \frac{2.7}{\sqrt{50}}$$
$$\approx 863 + 1.02 = 864.02 \text{ m/s}$$

We can be 99% confident that the population mean muzzle velocity is between 861.98 and 864.02 meters per second.

17. $\hat{p} = \frac{567}{1008} \approx 0.563$. The sample size $n = 1008$ is less than 5% of the population, and $n\hat{p}(1-\hat{p}) \approx 248 \geq 10$, so we can construct a Z-interval:
For 90% confidence the critical value is $z_{0.05} = 1.645$. Then:
Lower bound
$$= \hat{p} - z_{0.05} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$
$$= 0.563 - 1.645 \cdot \sqrt{\frac{0.563(1-0.563)}{1008}}$$
$$\approx 0.563 - 0.026 = 0.537$$
Upper bound
$$= \hat{p} + z_{0.05} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$
$$= 0.563 + 1.645 \cdot \sqrt{\frac{0.563(1-0.563)}{1008}}$$
$$\approx 0.563 + 0.026 = 0.589 \text{ [Tech: 0.588]}$$

The Gallup Organization can be 90% confident that the population proportion of adult Americans who are worried about having enough money for retirement is between 0.537 and 0.589 (i.e., between 53.7% and 58.9%).

18. We construct a Z-interval because we are estimating a population mean, we know the population standard deviation, and the underlying population is normally distributed. For 95% confidence the critical value is $z_{0.025} = 1.96$. Then:

$$\text{Lower bound} = \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}}$$
$$= 93.43 - 1.96 \cdot \frac{15}{\sqrt{40}}$$
$$\approx 93.43 - 4.65 = \$88.78$$
$$\text{Upper bound} = \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}}$$
$$= 93.43 + 1.96 \cdot \frac{15}{\sqrt{40}}$$
$$\approx 93.43 + 4.65 = \$98.08$$

We can be 95% confident that the population mean amount spent per person in one day at a certain theme park is between \$88.78 and \$98.08.

19. The normal probability plot and boxplot show that the data are normal with no outliers. We construct a Z-interval because we are estimating a population mean, and we know the population standard deviation. For 95% confidence the critical value is $z_{0.025} = 1.96$. The data give $n = 15$ and $\bar{x} = 69.85$ inches. Then:

$$\text{Lower bound} = \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}}$$
$$= 69.85 - 1.96 \cdot \frac{2.9}{\sqrt{15}}$$
$$\approx 69.85 - 1.47 = 68.38 \text{ inches}$$
[Tech: 68.39]
$$\text{Upper bound} = \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}}$$
$$= 69.85 + 1.96 \cdot \frac{2.9}{\sqrt{15}}$$
$$\approx 69.85 + 1.47 = 71.32 \text{ inches}$$

We are 95% confident that the population mean height of 20- to 29-year-old males is between 68.38 and 71.32 inches.

20. The normal probability plot and boxplot show that the data are normal with no outliers. We construct a Z-interval because we are estimating a population mean, and we know the population standard deviation. For 95% confidence the critical value is $z_{0.025} = 1.96$. The data give $n = 12$ and $\bar{x} = 267.75$ days. Then:

$$\text{Lower bound} = \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}}$$
$$= 267.75 - 1.96 \cdot \frac{16}{\sqrt{12}}$$
$$\approx 267.75 - 9.05 = 258.7 \text{ days}$$

$$\text{Upper bound } = \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}}$$
$$= 267.75 + 1.96 \cdot \frac{16}{\sqrt{15}}$$
$$\approx 267.75 + 9.05 = 276.8 \text{ days}$$

We can be 95% confident that the mean gestation period for humans is between 258.7 and 276.8 days.

21. The box plot indicates an outlier in the data, so we cannot compute a confidence interval using either method.

22. The normal probability plot and boxplot show that the data are normal with no outliers. We construct a t-interval because we are estimating the population mean, and we do not know the population standard deviation. The data give $n = 12$, $\bar{x} \approx 0.873$ grams and $s \approx 0.033$ grams. For 95% confidence, $\alpha/2 = 0.025$. With df = 11, we use $t_{0.025} = 2.201$. Then:

$$\text{Lower bound } = \bar{x} - t_{0.025} \cdot \frac{s}{\sqrt{n}}$$
$$= 0.873 - 2.201 \cdot \frac{0.033}{\sqrt{12}}$$
$$\approx 0.873 - 0.021 = 0.852 \text{ grams}$$

$$\text{Upper bound } = \bar{x} + t_{0.025} \cdot \frac{s}{\sqrt{n}}$$
$$= 0.873 + 2.201 \cdot \frac{0.033}{\sqrt{12}}$$
$$\approx 0.873 + 0.021 = 0.894 \text{ grams}$$

We can be 95% confident that the population mean weight of plain M&Ms is between 0.852 and 0.894 grams.

23. The normal probability plot and boxplot show that the data are normal with no outliers. We construct a t-interval because we are estimating the population mean and we do not know the population standard deviation. The data give $n = 15$, $\bar{x} \approx 109.3$ and $s \approx 14.4$. With df = 14, we use $t_{0.025} = 2.145$. Then:

$$\text{Lower bound } = \bar{x} - t_{0.025} \cdot \frac{s}{\sqrt{n}}$$
$$= 109.3 - 2.145 \cdot \frac{14.4}{\sqrt{15}}$$
$$\approx 109.3 - 8.0$$
$$= 101.3 \text{ beats per minute}$$

[Tech: 101.4]

$$\text{Upper bound } = \bar{x} + t_{0.025} \cdot \frac{s}{\sqrt{n}}$$
$$= 109.3 + 2.145 \cdot \frac{14.4}{\sqrt{15}}$$
$$\approx 109.3 + 8.0$$
$$= 117.3 \text{ beats per minute}$$

We can be 95% confident that the population mean pulse rate for women after 3 minutes of exercise is between 101.3 and 117.3 beats per minute.

24. We have a small sample and the normal probability plot shows that the data are not normal, so we cannot compute a confidence interval using either method.

Chapter 9 Review Exercises

1. **(a)** For a 99% confidence interval we want the t-value with an area in the right tail of 0.005. With df = 17, we read from the table that $t_{0.005} = 2.898$.

(b) For a 90% confidence interval we want the t-value with an area in the right tail of 0.05. With df = 26, we read from the table that $t_{0.05} = 1.706$.

2. **(a)** df $= n - 1 = 21$ and for 95% confidence, $\alpha/2 = 0.025$. From the table, $\chi^2_{1-0.025} = \chi^2_{0.975} = 10.283$ and $\chi^2_{0.025} = 35.479$.

(b) df $= n - 1 = 11$ and for 99% confidence, $\alpha/2 = 0.005$. From the table, $\chi^2_{1-0.005} = \chi^2_{0.995} = 2.603$ and $\chi^2_{0.005} = 26.757$.

3. In 100 samples, we would expect 95 of the 100 intervals to include the true population mean, 100. Random chance in sampling causes a particular interval to not include the true population mean.

4. In a 90% confidence interval, the 90% represents the proportion of intervals that would contain the parameter of interest (e.g. the population mean, population proportion, or population standard deviation) if a large number of different samples is obtained.