Chapter 8: Hypothesis Testing

Section 8-1: Basics of Hypothesis Testing

1. Rejection of the claim about aspirin is more serious because it is a drug used for medical treatments. The wrong aspirin dosage could cause more serious adverse reactions than a wrong vitamin C dosage. It would be wise to use a smaller significance level for testing the claim about the aspirin.

2. Estimates and hypothesis tests are both methods of inferential statistics, but they have different objectives. We could use the sample body temperatures to construct a confidence interval estimate of the population mean, but hypothesis testing is used to test some claim made about the value of the mean body temperature.

3. a. H_0: $\mu = 174.1$ cm

   b. H_1: $\mu \neq 174.1$ cm

   c. Reject the null hypothesis or fail to reject the null hypothesis.

   d. No, in this case, the original claim becomes the null hypothesis. For the claim that the mean height of men is equal to 174.1 cm, we can either reject that claim or fail to reject it, but we cannot state that there is sufficient evidence to *support* that claim.

4. The P-value of 0.001 is preferred because it corresponds to the sample evidence that most strongly supports the alternative hypothesis that the method is effective.

5. a. $p > 0.5$ (more than a majority)

   b. H_0: $p = 0.5$; H_1: $p > 0.5$

6. a. $p < 0.95$

   b. H_0: $p = 0.95$; H_1: $p < 0.95$

7. a. $\mu = 69$ bpm

   b. H_0: $\mu = 69$ bpm; H_1: $\mu \neq 69$ bpm

8. a. $\sigma = 11$ bpm

   b. H_0: $\sigma = 11$ bpm; H_1: $\sigma > 11$ bpm

9. There is sufficient evidence to support the claim that most adults would erase all of their personal information online if they could.

10. There is sufficient evidence to support the claim that fewer than 95% of adults have a cell phone.

11. There is not sufficient evidence to warrant rejection of the claim that the mean pulse rate (in beats per minute) of adult males is 69 bpm.

12. There is not sufficient evidence to support the claim that the standard deviation of pulse rates of adult males is more than 11 bpm.

13. $z = \dfrac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \dfrac{0.709 - 0.50}{\sqrt{\frac{(0.50)(0.50)}{983}}} = 13.11$

14. $z = \dfrac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \dfrac{0.87 - 0.95}{\sqrt{\frac{(0.95)(0.05)}{1128}}} = -12.33$ (or $z = -12.38$ if using $x = 0.87 \cdot 1128 = 981$)

15. $t = \dfrac{\overline{x} - \mu}{s/\sqrt{n}} = \dfrac{69.6 - 69}{11.3/\sqrt{153}} = 0.657$

16. $\chi^2 = \dfrac{(153-1)s^2}{\sigma^2} = \dfrac{(40-1)11.3^2}{11^2} = 160.404$

17. a. right-tailed

    b. $P\text{-value} = P(z > 1.00) = 0.1587$

    c. $0.1587 > 0.05$; Fail to reject H_0.

18. a. left-tailed

    b. $P\text{-value} = P(z < -2.50) = 0.0062$

    c. $0.0062 < 0.05$; Reject H_0.

19. a. two-tailed

    b. $P\text{-value} = 2 \cdot P(z > 2.01) = 0.0444$

    c. $0.0444 < 0.05$; Reject H_0.

20. a. two-tailed

    b. $P\text{-value} = 2 \cdot P(z < -1.94) = 0.0524$

    c. $0.0524 > 0.05$; Fail to reject H_0.

21. a. Critical value: $z = 1.645$
    b. $1.00 < 1.645$; Fail to reject H_0.
22. a. Critical value: $z = -1.645$
    b. $-2.50 < -1.645$; Reject H_0.
23. a. Critical values: $z = \pm 1.645$
    b. $-2.01 < -1.645$; Reject H_0.
24. a. Critical values: $z = \pm 1.96$
    b. $-1.96 < -1.94 < 1.96$; Fail to reject H_0.
25. a. $0.2678 > 0.05$; Fail to reject H_0.
    b. There is not sufficient evidence to support the claim that more than 70% do not have hypertension.
26. a. $0.0003 < 0.05$; Reject H_0.
    b. There is sufficient evidence to support the claim that fewer than 90% of nurses have a cell phone.
27. a. $0.0095 < 0.05$; Reject H_0.
    b. There is sufficient evidence to warrant rejection of the claim that the mean pulse rate (in beats per minute) of adult males is 72 bpm.
28. a. $0.3045 > 0.05$; Fail to reject H_0.
    b. There is not sufficient evidence to support the claim that the standard deviation of pulse rates of adult males is more than 11 bpm.
29. Type I error: In reality $p = 0.1$, but we reject the claim that $p = 0.1$. Type II error: In reality $p \neq 0.1$, but we fail to reject the claim that $p = 0.1$.
30. Type I error: In reality $p = 0.35$, but we reject the claim that $p = 0.35$. Type II error: In reality $p \neq 0.35$, but we fail to reject the claim that $p \neq 0.35$.
31. Type I error: In reality $p = 0.87$, but we support the claim that $p > 0.87$. Type II error: In reality $p > 0.87$, but we fail to support that conclusion.
32. Type I error: In reality $p = 0.25$, but we support the claim that $p < 0.25$. Type II error: In reality $p < 0.25$, but we fail to support that conclusion.
33. The power of 0.96 shows that there is a 96% chance of rejecting the null hypothesis of $p = 0.08$ when the true proportion is actually 0.18. That is, if the proportion of Chantix users who experience abdominal pain is actually 0.18, then there is a 96% chance of supporting the claim that the proportion of Chantix users who experience abdominal pain is greater than 0.08.
34. a. From $p = 0.5$, $\hat{p} = 0.5 + 1.645\sqrt{\frac{(0.5)(0.5)}{64}} = 0.6028125$

    From $p = 0.65$, $z = \frac{0.6028125 - 0.65}{\sqrt{\frac{(0.65)(0.35)}{64}}} = -0.791$; Power $= P(z > -0.791) = 0.7852$ (Tech: 0.7857)

    b. Assuming that $p = 0.5$, as in the null hypothesis, the critical value of $z = 1.645$ corresponds to $\hat{p} = 0.6028125$, so any sample proportion greater than 0.6028125 causes us to reject the null hypothesis, as shown in the shaded critical region of the top graph. If p is actually 0.65, then the null hypothesis of $p = 0.5$ is false, and the actual probability of rejecting the null hypothesis is found by finding the area greater than $\hat{p} = 0.6028125$ in the bottom graph, which is the shaded area. That is, the shaded area in the bottom graph represents the probability of rejecting the false null hypothesis.
35. From $p = 0.5$, $\hat{p} = 0.5 + 1.645\sqrt{\frac{(0.5)(0.5)}{n}}$; from $p = 0.55$, since $P(z > -0.842) = 0.8000$,

    $\hat{p} = 0.55 - 0.842\sqrt{\frac{(0.55)(0.45)}{n}}$

35. (continued)

$$\text{So: } 0.5+1.645\sqrt{\frac{(0.5)(0.5)}{n}}=0.55-0.842\sqrt{\frac{(0.55)(0.45)}{n}}$$

$$0.5\sqrt{n}+1.645\sqrt{0.25}=0.55\sqrt{n}-0.842\sqrt{0.2475}$$

$$0.05\sqrt{n}=1.645\sqrt{0.25}+0.842\sqrt{0.2475}$$

$$n=\left(\frac{1.645\sqrt{0.25}+0.842\sqrt{0.2475}}{0.05}\right)^2=617$$

Section 8-2: Testing a Claim About a Proportion

1. a. $0.53\cdot 510=270$

   b. $\hat{p}=0.53$; The symbol $\hat{p}$ is used to represent a sample proportion.

2. $H_0\colon p=0.5$; $H_1\colon p>0.5$

3. The method based on a confidence interval is not equivalent to the P-value method and the critical value method.

4. a. The first requirement is violated because the sample is a voluntary response sample instead of being a simple random sample. Because a requirement is violated, the methods of this section should not be used to test the claim.

   b. If the P-value is very low (such as less than or equal to 0.05), "the null must go" means that we should reject the null hypothesis.

   c. The statement that "if the P is high, the null will fly" suggests that with a high P-value, the null hypothesis has been proved or is supported, but we should never make such a conclusion.

   d. Choosing a significance level with a number like 0.0483 would make it seem like you're scheming to reach a desired conclusion.

5. a. left-tailed.

   b. $z=-4.46$

   c. P-value $=0.000004$

   d. $H_0\colon p=0.1$; Reject the null hypothesis.

   e. There is sufficient evidence to support the claim that less than 10% of treated subjects experience headaches.

6. a. right-tailed.

   b. $z=2.92$

   c. P-value $=0.0017$

   d. $H_0\colon p=0.25$; Reject the null hypothesis.

   e. There is sufficient evidence to support the claim that more than 1/4 of adults feel comfortable in a self-driving vehicle.

7. a. two-tailed.

   b. $z=-1.36$

   c. P-value $=0.174$

   d. $H_0\colon p=0.67$; Fail to reject the null hypothesis.

   e. There is not sufficient evidence to warrant rejection of the claim that 92% of adults own cell phones.

8. a. two-tailed.

   b. $z=1.33$

   c. P-value $=0.1840$

   d. $H_0\colon p=0.5$; Fail to reject the null hypothesis.

   e. There is not sufficient evidence to warrant rejection of the claim that half of us say that we should replace passwords with biometric security.

9. $H_0: p = 0.5$; $H_1: p > 0.5$; Test statistic: $z = \dfrac{\frac{239}{291} - 0.5}{\sqrt{\frac{(0.5)(0.5)}{291}}} = 10.96$;

$P\text{-value} = P(z > 10.96) = 0.0001$ (Tech: 0.0000); Critical value: $z = 2.33$;

Reject H_0. There is sufficient evidence to support the claim that the YSORT method is effective in increasing the likelihood that a baby will be a boy.

10. $H_0: p = 0.03$; $H_1: p \neq 0.03$; Test statistic: $z = \dfrac{\frac{153}{5924} - 0.03}{\sqrt{\frac{(0.03)(0.97)}{5924}}} = -1.88$;

$P\text{-value} = 2 \cdot P(z < -1.88) = 0.0602$ (Tech: 0.0597); Critical values: $z = \pm 1.96$;

Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that 3% of Eliquis users develop nausea. The rate of nausea appears to be quite low, so it is not a problematic adverse reaction.

11. $H_0: p = 0.5$; $H_1: p \neq 0.5$; Test statistic: $z = \dfrac{\frac{481}{882} - 0.5}{\sqrt{\frac{(0.5)(0.5)}{882}}} = 2.69$;

$P\text{-value} = 2 \cdot P(z > 2.69) = 0.0072$ (Tech: 0.0071); Critical values: $z = \pm 2.575$ (Tech: $z = \pm 2.576$);

Reject H_0. There is sufficient evidence to reject the claim that the proportion of those in favor is equal to 0.5. The result suggests that the politician is wrong in claiming that the responses are random guesses equivalent to a coin toss.

12. $H_0: p = 0.06$; $H_1: p > 0.06$; Test statistic: $z = \dfrac{\frac{72}{724} - 0.06}{\sqrt{\frac{(0.06)(0.94)}{724}}} = 4.47$;

$P\text{-value} = P(z > 1.47) = 0.0001$ (Tech: 0.0000); Critical value: $z = 1.645$;

Reject H_0. There is sufficient evidence to warrant rejection of the claim that the rate of nausea is the same as the rate experienced by flu patients given a placebo. It appears the rate of nausea is greater for the Tamiflu patients, which would be of concern.

13. $H_0: p = 0.20$; $H_1: p > 0.20$; Test statistic: $z = \dfrac{\frac{52}{227} - 0.20}{\sqrt{\frac{(0.20)(0.80)}{227}}} = 1.10$;

$P\text{-value} = P(z > 1.10) = 0.1357$ (Tech: 0.1367); Critical value: $z = 1.645$;

Fail to reject H_0. There is not sufficient evidence to support the claim that more than 20% of OxyContin users develop nausea. However, with $\hat{p} = 0.229$, we see that a large percentage of OxyContin users experience nausea, so that rate does appear to be very high.

14. $H_0: p = 0.5$; $H_1: p > 0.5$; Test statistic: $z = \dfrac{\frac{856}{1228} - 0.5}{\sqrt{\frac{(0.5)(0.5)}{1228}}} = 13.81$;

$P\text{-value} = P(z > 13.81) = 0.0001$ (Tech: 0.0000); Critical value: $z = 2.33$;

Reject H_0. There is sufficient evidence to support the claim that most medical malpractice lawsuits are dropped or dismissed. This should be comforting to physicians.

15. $H_0: p = 0.15$; $H_1: p < 0.15$; Test statistic: $z = \dfrac{\frac{717}{5000} - 0.15}{\sqrt{\frac{(0.15)(0.85)}{5000}}} = -1.31$;

$P\text{-value} = P(z < -1.31) = 0.1357$ (Tech: 0.1367); Critical value: $z = -2.33$;

Fail to reject H_0. There is not sufficient evidence to support the claim that the return rate is less than 15%.

16. H_0: $p = 0.10$; H_1: $p < 0.10$; Test statistic: $z = \dfrac{\frac{27}{300} - 0.10}{\sqrt{\frac{(0.10)(0.90)}{300}}} = -0.58$;

P-value $= P(z < -0.58) = 0.2810$ (Tech: 0.2819); Critical value: $z = -1.645$;

Fail to reject H_0. There is not sufficient evidence to support the claim that less than 10% of the test results are wrong. The sample results suggest that the test is wrong too often to be considered very reliable.

17. H_0: $p = 0.512$; H_1: $p \neq 0.512$; Test statistic: $z = \dfrac{\frac{426}{860} - 0.512}{\sqrt{\frac{(0.512)(0.488)}{860}}} = -0.98$;

P-value $= 2 \cdot P(z < -0.98) = 0.3270$ (Tech: 0.3286); Critical values: $z = \pm 1.96$;

Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that 51.2% of newborn babies are boys. The results do not *support* the belief that 51.2% of newborn babies are boys; the results merely show that there is not strong evidence against the rate of 51.2%.

18. H_0: $p = 0.25$; H_1: $p \neq 0.25$; Test statistic: $z = \dfrac{\frac{152}{580} - 0.25}{\sqrt{\frac{(0.25)(0.25)}{580}}} = 0.67$;

P-value $= 2 \cdot P(z > 0.67) = 0.5028$ (Tech: 0.5021); Critical values: $z = \pm 2.575$ (Tech: $z = \pm 2.576$);

Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that 25% of offspring peas are yellow. There is not sufficient evidence to conclude that Mendel's rate of 25% is wrong.

19. H_0: $p = 0.5$; H_1: $p \neq 0.5$; Test statistic: $z = \dfrac{0.55 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{104}}} = 1.02$; $(z = 0.98 \text{ using } x = 57)$

P-value $= 2 \cdot P(z > 1.02) = 0.3078$; $(0.3268,$ Table: $0.3270 \text{ using } x = 57)$ Critical value: $z = \pm 1.96$;

Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that women who guess the gender of their babies have a success rate equal to 50%.

20. H_0: $p = 0.5$; H_1: $p \neq 0.5$; Test statistic: $z = \dfrac{\frac{32}{45} - 0.5}{\sqrt{\frac{(0.5)(0.5)}{104}}} = 2.83$;

P-value $= P(z > 2.83) = 0.0023$; Critical value: $z = 2.33$;

Reject H_0. There is sufficient evidence to support the claim that women with more than 12 years of education have a proportion of correct predictions that is greater than 0.5. It appears these women do have an ability to correctly predict the gender of their babies.

21. H_0: $p = 0.5$; H_1: $p \neq 0.5$; Test statistic: $z = \dfrac{\frac{123}{280} - 0.5}{\sqrt{\frac{(0.5)(0.5)}{280}}} = -2.03$;

P-value $= 2 \cdot P(z < -2.03) = 0.0424$ (Tech: 0.0422); Critical values: $z = \pm 1.645$;

Reject H_0. There is sufficient evidence to warrant rejection of the claim that touch therapists use a method equivalent to random guesses. However, their success rate of $123/280$, or 43.9%, indicates that they performed *worse* than random guesses, so they do not appear to be effective.

22. H_0: $p = 0.5$; H_1: $p \neq 0.5$; Test statistic: $z = \dfrac{\frac{123}{280} - 0.5}{\sqrt{\frac{(0.5)(0.5)}{280}}} = -2.03$;

$P\text{-value} = 2 \cdot P(z < -2.03) = 0.0424$ (Tech: 0.0422); Critical values: $z = \pm 2.575$ (Tech: $z = \pm 2.576$);

Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that touch therapists use a method equivalent to random guesses. After changing the significance level from 0.10 to 0.01, the conclusion does change.

23. H_0: $p = 0.00034$; H_0: $p \neq 0.00034$; Test statistic: $z = \dfrac{\frac{135}{420{,}095} - 0.000340}{\sqrt{\frac{(0.000340)(0.99966)}{420{,}095}}} = -0.66$;

$P\text{-value} = 2 \cdot P(z < -0.66) = 0.5092$ (Tech: 0.5122); Critical values: $z = \pm 2.81$;

Fail to reject H_0. There is not sufficient evidence to support the claim that the rate is different from 0.0340%. Cell phone users should not be concerned about cancer of the brain or nervous system.

24. H_0: $p = 0.80$; H_1: $p < 0.80$; Test statistic: $z = \dfrac{\frac{74}{98} - 0.80}{\sqrt{\frac{(0.80)(0.20)}{98}}} = -1.11$;

$P\text{-value} = P(z < -1.11) = 0.1335$ (Tech: 0.1332); Critical values: $z = \pm 1.96$;

Fail to reject H_0. There is not sufficient evidence support the claim that such polygraph results are correct less than 80% of the time. Opinions may vary, but it appears the polygraph tests are somewhat accurate.

25. H_0: $p = 0.5$; H_1: $p > 0.5$; Test statistic: $z = \dfrac{\frac{39}{71} - 0.5}{\sqrt{\frac{(0.5)(0.5)}{71}}} = 0.83$;

$P\text{-value} = P(z > 0.83) = 0.2033$ (Tech: 0.2031); Critical value: $z = 1.645$;

Fail to reject H_0. There is not sufficient evidence to support the claim that among smokers who try to quit with nicotine patch therapy, the majority are smoking one year after the treatment. There isn't sufficient evidence to conclude that the nicotine patch therapy is not effective.

26. H_0: $p = 0.5$; H_1: $p < 0.5$; Test statistic: $z = \dfrac{\frac{6062}{12{,}000} - 0.5}{\sqrt{\frac{(0.5)(0.5)}{12{,}000}}} = 1.13$;

$P\text{-value} = P(z > 1.13) = 0.8708$ (Tech: 0.8712); Critical value: $z = -1.645$;

Fail to reject H_0. There is not sufficient evidence to support the claim that less than 0.5 of the deaths occur the week before Thanksgiving. Based on these results, there is no indication that people can temporarily postpone their death to survive Thanksgiving. (With 50.5% of the deaths occurring before Thanksgiving, there is no way that the claim of a proportion *less than* 0.5 could be supported.)

27. H_0: $p = 0.80$; H_1: $p \neq 0.80$; Test statistic: $z = \dfrac{0.828 - 0.80}{\sqrt{\frac{(0.80)(0.20)}{198}}} = 0.98$ (using $x = 164, z = 0.99$);

$P\text{-value} = 2 \cdot P(z > 0.98) = 0.3270$ (Tech: 0.3246); Critical values: $z = \pm 2.575$ (Tech: $z = \pm 2.576$);

Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that 80% of patients stop smoking when given sustained care. With a success rate around 80%, it appears that sustained care is effective.

28. H_0: $p = 0.75$; H_1: $p > 0.75$; Test statistic: $z = \dfrac{0.817 - 0.75}{\sqrt{\frac{(0.75)(0.25)}{3005}}} = 8.48$;

P-value $= P(z > 8.48) = 0.0001$ (Tech: 0.0000); Critical value: $z = -2.33$;

Reject H_0. There is sufficient evidence to support the claim that more than $3/4$ of adults use at least one prescription medication. With more than $3/4$ of adults using at least one prescription medication, it appears that prescription use among adults is high.

29. H_0: $p = 0.5$; H_1: $p \neq 0.5$;

Normal approximation:

$$z = \frac{\frac{9}{10} - 0.5}{\sqrt{\frac{(0.5)(0.5)}{10}}} = 2.53; \quad P\text{-value} = 2 \cdot P(z > 2.53) = 0.0114$$

Exact:

$$P\text{-value} = 2 \cdot \left({}_{10}C_9\left(0.5^9\right)\left(0.5^1\right) + {}_{10}C_{10}\left(0.5^{10}\right)\left(0.5^0\right)\right) = 0.0215$$

Continuity Correction:

$$P\text{-value} = 2 \cdot \left({}_{10}C_9\left(0.5^9\right)\left(0.5^1\right) + {}_{10}C_{10}\left(0.5^{10}\right)\left(0.5^0\right)\right) - \frac{1}{2}\left({}_{10}C_9\left(0.5^9\right)\left(0.5^1\right)\right) = 0.0117$$

H_0: $p = 0.4$; H_1: $p \neq 0.4$;

Normal approximation:

$$z = \frac{\frac{9}{10} - 0.4}{\sqrt{\frac{(0.4)(0.6)}{10}}} = 3.23; \quad P\text{-value} = 2 \cdot P(z > 3.23) = 0.0012$$

Exact:

$$P\text{-value} = 2 \cdot \left({}_{10}C_9\left(0.4^9\right)\left(0.6^1\right) + {}_{10}C_{10}\left(0.4^{10}\right)\left(0.6^0\right)\right) = 0.0034$$

Continuity Correction:

$$P\text{-value} = 2 \cdot \left({}_{10}C_9\left(0.4^9\right)\left(0.6^1\right) + {}_{10}C_{10}\left(0.4^{10}\right)\left(0.6^0\right)\right) - \frac{1}{2}\left({}_{10}C_9\left(0.4^9\right)\left(0.6^1\right)\right) = 0.0018$$

H_0: $p = 0.5$; H_1: $p > 0.5$;

Normal approximation:

$$z = \frac{\frac{545}{1009} - 0.5}{\sqrt{\frac{(0.5)(0.5)}{1009}}} = 2.53; \quad P\text{-value} = P(z > 2.53) = 0.0057 \text{ (Tech: 0.0054)}$$

Exact:

$$P\text{-value} = {}_{1009}C_{545}\left(0.5^{545}\right)\left(0.5^{464}\right) + \cdots + {}_{1009}C_{1009}\left(0.5^{1009}\right)\left(0.5^0\right) = 0.0059$$

Continuity Correction:

$$P\text{-value} = {}_{1009}C_{545}\left(0.5^{545}\right)\left(0.5^{464}\right) + \cdots + {}_{1009}C_{1009}\left(0.5^{1009}\right)\left(0.5^0\right) - \frac{1}{2}\left({}_{1009}C_{545}\left(0.5^{545}\right)\left(0.5^{464}\right)\right)$$
$$= 0.0054$$

The P-values agree reasonably well with the large sample size of $n = 1009$. The normal approximation to the binomial distribution works better as the sample size increases.

30. a. H_0: $p = 0.10$; H_1: $p \neq 0.10$;. Test statistic: $z = \dfrac{0.119 - 0.1}{\sqrt{\frac{(0.1)(0.9)}{1000}}} = 2.00$; Critical values: $z = \pm 1.96$;

Reject H_0. There is sufficient evidence to warrant rejection of the claim that the proportion of zeros is 0.1.

30. (continued)

b. H_0: $p = 0.10$; H_1: $p \neq 0.10$; Test statistic: $z = \dfrac{0.119 - 0.1}{\sqrt{\frac{(0.1)(0.9)}{1000}}} = 2.00$;

P-value $= 2 \cdot P(z > 2.00) = 0.0456$ (Tech: 0.0452);

Reject H_0. There is sufficient evidence to warrant rejection of the claim that the proportion of zeros is 0.1.

c. 95% CI: $\hat{p} \pm z_{\alpha/2}\sqrt{\dfrac{\hat{p}\hat{q}}{n}} = 0.119 \pm 1.96\sqrt{\dfrac{(0.119)(0.881)}{1000}} \Rightarrow 0.0989 < p < 0.139$; Because 0.1 is contained within the confidence interval, fail to reject H_0: $p = 0.10$. There is not sufficient evidence to warrant rejection of the claim that the proportion of zeros is 0.1.

d. The traditional and P-value methods both lead to rejection of the claim, but the confidence interval method does not lead to rejection of the claim.

31. a. From $p = 0.40$, $\hat{p} = 0.4 - 1.645\sqrt{\dfrac{(0.4)(0.6)}{50}} = 0.286$

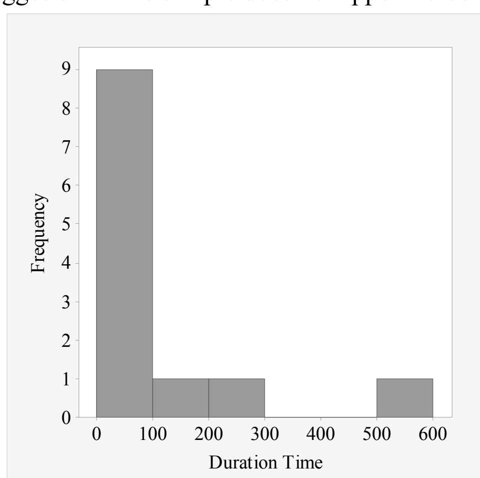
From $p = 0.25$, $z = \dfrac{0.286 - 0.25}{\sqrt{\frac{(0.25)(0.75)}{50}}} = 0.588$; Power $= P(z < 0.588) = 0.7224$ (Tech: 0.7219)

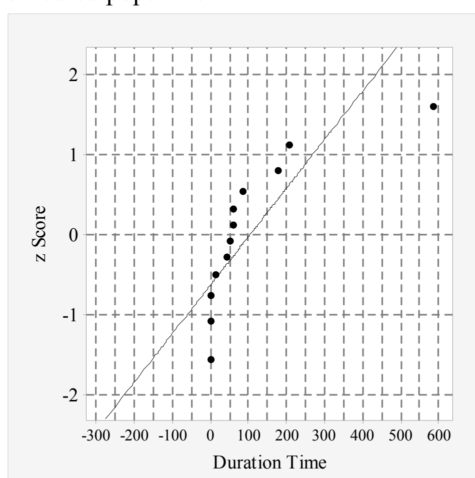
b. $1 - 0.7224 = 0.2776$ (Tech: 0.2781)

c. The power of 0.7224 shows that there is a reasonably good chance of making the correct decision of rejecting the false null hypothesis. It would be better if the power were even higher, such as greater than 0.8 or 0.9.

Section 8-3: Testing a Claim About a Mean

1. The requirements are (1) the sample must be a simple random sample, and (2) either or both of these conditions must be satisfied: The population is normally distributed or $n > 30$. There is not enough information given to determine whether the sample is a simple random sample. Because the sample size is not greater than 30, we must check for normality, but the value of 583 sec appears to be an outlier, and a normal quantile plot or histogram suggests that the sample does not appear to be from a normally distributed population.





2. df denotes the number of degrees of freedom. For the sample of 12 times, df $= 12 - 1 = 11$.

3. A t test is a hypothesis test that uses the Student t distribution, such as the method of testing a claim about a population mean as presented in this section. The t test methods are much more likely to be used than the z test methods because the t test does not require a known value of σ, and realistic hypothesis tests of claims about μ typically involve a population with an unknown value of σ.

4. For a 0.05 significance level used in a one-tailed test, use a 90% confidence level. The given confidence interval does contain the value of 90 sec, so it is possible that the value of μ is equal to 90 sec or some lower value, so there is not sufficient evidence to support the claim that the mean is greater than 90 sec.

5. P-value = 0.0437 (Table: $0.025 < P\text{-value} < 0.05$)

6. P-value = 0.0000 (Table: $P\text{-value} < 0.01$)

7. P-value = 0.2581 (Table: $P\text{-value} > 0.10$)

8. P-value = 0.1324 (Table: $0.05 < P\text{-value} < 0.10$)

9. H_0: $\mu = 98.6^\circ$F; H_1: $\mu \neq 98.6^\circ$F; Test statistic: $t = -3.8654$; P-value = 0.0004; Critical values (with $\alpha = 0.05$): $t = \pm 2.0262$ (Table: ± 2.028);
Reject H_0. There is sufficient evidence to warrant rejection of the claim that the mean body temperature is equal to 98.6°F.

10. H_0: $\mu = 98.6$°F; H_1: $\mu \neq 98.6$°F; Test statistic: $t = -13.272$; P-value < 0.0001;
Critical values (with $\alpha = 0.05$): $t = \pm 1.995$ (Table: ± 1.994);
Reject H_0. There is sufficient evidence to warrant rejection of the claim that the mean body temperature is equal to 98.6°F.

11. H_0: $\mu = 270$ (1000 cells/μL); H_1: $\mu < 270$ (1000 cells/μL); Test statistic: $t = -2.7637$; P-value = 0.0032;
Critical value (with $\alpha = 0.05$): $t = -1.665$; (Table: -1.660);
Reject H_0. There is sufficient evidence to support the claim that the population of adult females has a mean platelet count less than 270 (1000 cells/μL).

12. H_0: $\mu = 220$ (1000 cells/μL); H_1: $\mu > 220$ (1000 cells/μL); Test statistic: $t = 0.89$; P-value = 0.188; Critical value (with $\alpha = 0.05$): $t = 1.665$; (Table: 1.660);
Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that the mean platelet count for adult males is different than 220 (1000 cells/μL).

13. The data appear to follow the loose definition for a normal distribution and $n > 30$.
H_0: $\mu = 3000$ g; H_1: $\mu > 3000$ g;
Test statistic: $t = \dfrac{3037.1 - 3000}{706.3/\sqrt{205}} = 0.752$; Critical value: $t = 2.345$;
P-value = 0.2264 (Table: $P\text{-value} > 0.10$);
Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that the population mean of birth weights of females is greater than 3000 g.

14. The data appear to follow the loose definition for a normal distribution and $n > 30$.
H_0: $\mu = 3400$ g; H_1: $\mu < 3400$ g;
Test statistic: $t = \dfrac{3272.8 - 3400}{660.2/\sqrt{195}} = -2.690$; Critical value: $t = 2.345$;
P-value = 0.0039 (Table: $P\text{-value} < 0.01$);
Reject H_0. There is sufficient evidence to support the claim that the population mean of birth weights of males is greater than 3400 g.

15. The data cannot be verified to follow a normal distribution, but $n > 30$.
H_0: $\mu = 0$; H_1: $\mu > 0$;
Test statistic: $t = \dfrac{0.4 - 0}{21.0/\sqrt{49}} = -0.133$; Critical value: $t = 1.677$ (Table: $t \approx 1.676$);
P-value $= 0.4472$ (Table: P-value > 0.10);
Fail to reject H_0. There is not sufficient evidence to support the claim that with garlic treatment, the mean change in LDL cholesterol is greater than 0. The results suggest that the garlic treatment is not effective in reducing LDL cholesterol levels.

16. The data cannot be verified to follow a normal distribution and $n < 30$, so proceed with caution.
H_0: $\mu = 102.8$ min; H_1: $\mu < 102.8$ min;
Test statistic: $t = \dfrac{98.9 - 102.8}{42.3/\sqrt{16}} = -0.369$; Critical value ($\alpha = 0.05$): $t = -1.753$;
P-value $= 0.3587$ (Table: P-value > 0.10);
Fail to reject H_0. There is not sufficient evidence to support the claim that after treatment with zopiclone, subjects have a mean wake time less than 102.8 min. These results suggest that the zopiclone treatment is not effective.

17. The data cannot be verified to follow a normal distribution, but $n > 30$.
H_0: $\mu = 0$; H_1: $\mu > 0$;
Test statistic: $t = \dfrac{3.0 - 0.0}{4.9/\sqrt{40}} = 3.872$; Critical value: $t = 2.426$;
P-value $= 0.0002$ (Table: P-value < 0.005);
Reject H_0. There is sufficient evidence to support the claim that the mean weight loss is greater than 0. Although the diet appears to have statistical significance, it does not appear to have practical significance, because the mean weight loss of only 3.0 lb does not seem to be worth the effort and cost.

18. The data cannot be verified to follow a normal distribution, but $n > 30$.
H_0: $\mu = 69.5$ years; H_1: $\mu > 69.5$ years;
Test statistic: $t = \dfrac{73.4 - 69.5}{8.7/\sqrt{35}} = 2.652$; Critical value: $t = 1.691$;
P-value $= 0.0060$ (Table: $0.005 < P$-value < 0.01);
Reject H_0. There is sufficient evidence to support the claim that the mean life span for conductors is greater than 69.5 years.

19. The data cannot be verified to follow a normal distribution, but $n > 30$.
H_0: $\mu = 4.00$; H_1: $\mu \neq 4.00$;
Test statistic: $t = \dfrac{3.97 - 4.00}{0.55/\sqrt{436}} = -1.139$; Critical values: $t = \pm 1.965$; (Table: 1.966);
P-value $= 0.2554$ (Table: P-value > 0.20);
Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that the population of student course evaluations has a mean equal to 4.00. Because the data are from the University of Texas at Austin, they don't necessarily apply to a larger population that extends beyond that one institution.

20. The data cannot be verified to follow a normal distribution and $n < 30$, so proceed with caution.
H_0: $\mu = 0$; H_1: $\mu > 0$;
Test statistic: $t = \dfrac{4.00 - 0}{2.17/\sqrt{21}} = 8.447$; Critical value: $t = 2.584$;
P-value $= 0.0000$ (Table: P-value < 0.005);

20. (continued)

Reject H_0. There is sufficient evidence to support the claim that the mean fatigue level is positive. It appears the treatment is effective.

21. The sample data meet the loose requirement of having a normal distribution.
H_0: $\mu = 14\ \mu\text{g/g}$; H_1: $\mu < 14\ \mu\text{g/g}$;
Test statistic: $t = \dfrac{11.05 - 14.0}{6.46/\sqrt{10}} = -1.444$; Critical value: $t = -1.883$;
$P\text{-value} = 0.0913$ (Table: $P\text{-value} > 0.05$);
Fail to reject H_0. There is not sufficient evidence to support the claim that the mean lead concentration for all such medicines is less than $14\ \mu\text{g/g}$.

22. The sample data meet the loose requirement of having a normal distribution.
H_0: $\mu = 60$ sec; H_1: $\mu \neq 60$ sec;
Test statistic: $t = \dfrac{62.67 - 60.00}{9.48/\sqrt{15}} = 0.530$; Critical values: $t = \pm 2.145$;
$P\text{-value} = 0.6043$ (Table: $P\text{-value} > 0.20$);
Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that the times are from a population with a mean equal to 60 seconds. Although some of the individual estimates are off by a large amount, the group of 15 students yielded the mean of 62.7 sec, which is not significantly different from 60 sec, so as a group they appear to be reasonably good at estimating one minute.

23. The data do not appear to follow a normal distribution and $n < 30$, so proceed with caution.
H_0: $\mu = 1000$ hic; H_1: $\mu < 1000$ hic;
Test statistic: $t = \dfrac{704 - 1000}{273/\sqrt{6}} = -2.661$; Critical value: $t = -3.365$;
$P\text{-value} = 0.0224$ (Table: $0.1 < P\text{-value} < 0.025$);
Fail to reject H_0. There is not sufficient evidence to support the claim that the population mean is less than 1000 hic. There is not strong evidence that the mean is less than 1000 hic, and one of the booster seats has a measurement of 1210 hic, which does not satisfy the specified requirement of being less than 1000 hic.

24. The sample data meet the loose requirement of having a normal distribution.
H_0: $\mu = 162$ cm; H_1: $\mu > 162$ cm;
Test statistic: $t = \dfrac{117.25 - 162}{1.84/\sqrt{16}} = 33.082$; Critical value: $t = 2.602$;
$P\text{-value} = 0.0000$ (Table: $P\text{-value} < 0.005$);
Reject H_0. There is sufficient evidence to support the claim that supermodels have heights with a mean that is greater than the mean height of 162 cm for women in the general population. Supermodels appear to be taller than typical women.

25. The data appear to follow a normal distribution and $n > 30$.
H_0: $\mu = 75$ bpm; H_1: $\mu < 75$ bpm;
Test statistic: $t = \dfrac{74.04 - 75.00}{12.54/\sqrt{147}} = -0.927$; Critical value: $t = -1.655$ (Table: $t \approx -1.660$);
$P\text{-value} = 0.1777$ (Table: $P\text{-value} > 0.10$);
Fail to reject H_0. There is not sufficient evidence to support the claim that the mean pulse rate of adult females is less than 75 bpm.

26. The data appear to follow a normal distribution and $n > 30$.
H_0: $\mu = 75$ bpm; H_1: $\mu < 75$ bpm;
Test statistic: $t = \dfrac{69.58 - 75}{11.33/153} = -5.91$; Critical value: $t = -1.655$ (Table: $t \approx -1.660$);
P-value = 0.0000 (Table: P-value < 0.005);
Reject H_0. There is sufficient evidence to support the claim that the mean pulse rate of adult males is less than 75 bpm.

27. The data appear to follow a normal distribution and $n > 30$.
H_0: $\mu = 90$ mm Hg; H_1: $\mu < 90$ mm Hg;
Test statistic: $t = \dfrac{70.163 - 90}{11.22/\sqrt{147}} = -21.435$; Critical value: $t = -1.655$ (Table: $t \approx -1.660$);
P-value = 0.0000 (Table: P-value < 0.005);
Reject H_0. There is sufficient evidence to support the claim that the adult female population has a mean diastolic blood pressure level less than 90 mm Hg. The conclusion addresses the mean of a population, not individuals, so we cannot conclude that there are no female adults in the sample with hypertension.

28. The data appear to follow a normal distribution and $n > 30$.
H_0: $\mu = 90$ mm Hg; H_1: $\mu < 90$ mm Hg;
Test statistic: $t = \dfrac{71.32 - 90}{11.994/\sqrt{153}} = -19.265$; Critical value: $t = -1.655$ (Table: $t \approx -1.660$);
P-value = 0.0000 (Table: P-value < 0.005);
Reject H_0. There is sufficient evidence to support the claim that the adult male population has a mean diastolic blood pressure level less than 90 mm Hg. The conclusion addresses the mean of a population, not individuals, so we cannot conclude that there are no male adults in the sample with hypertension.

29. $A = \dfrac{1.645(8 \cdot 146 + 3)}{8 \cdot 146 + 1} = 1.647814371$ and $t = \sqrt{146\left(e^{1.647814371^2/146} - 1\right)} = -1.6554$; The approximation yields a critical t score of -1.6554, which is the same as the value of -1.6554 found from Statdisk. The approximation appears to work quite well.

30. The power of 0.4943 shows that the chance of recognizing that $\mu < 7$ hours is not very high when in reality $\mu = 6.0$ hours. It would be better if the power was higher, such as 0.8 or greater. $\beta = 1 - 0.4943 = 0.5057$, so there is better than a 50% chance of failing to recognize that $\mu < 7$ hours when in reality $\mu = 6$.

Section 8-4: Testing a Claim About a Standard Deviation of Variance

1. The sample must be a simple random sample and the sample must be from a normally distributed population. The normality requirement for a hypothesis test of a claim about a standard deviation is much more strict, meaning that the distribution of the population must be much closer to a normal distribution.

2. H_0: $\sigma = 0.470$ kg; H_1: $\sigma \neq 0.470$ kg; Test statistic: $\chi^2 = \dfrac{(n-1)s^2}{\sigma^2} = \dfrac{(16-1)0.6573^2}{0.470^2} = 29.337$ (29.339 if using the original data values.); The sampling distribution of the test statistic is the χ^2 (chi-square) distribution.

3. a. Reject H_0.
   b. Reject the claim that the sample is from a population with a standard deviation equal to 0.470 kg.
   c. It appears that the vitamin supplement does affect the variation among birth weights.

4. All of the values in the confidence interval are greater than the standard deviation of 0.470 kg, so it appears that birth weights of babies born to mothers taking the supplement have a larger standard deviation.

5. H_0: $\sigma = 10$ bpm; H_1: $\sigma \neq 10$ bpm; Test statistic: $\chi^2 = 194.0888$; P-value $= 0.0239$;
Reject H_0. There is sufficient evidence to warrant rejection of the claim that pulse rates of men have a standard deviation equal to 10 beats per minute. Using the normal range of 60 to 100 beats per minute is not very good for estimating σ in this case.

6. H_0: $\sigma = 10$ bpm; H_1: $\sigma \neq 10$ bpm; Test statistic: $\chi^2 = 229.718$ or $\chi^2 = \dfrac{(147-1)12.5^2}{10^2} = 228.125$;
P-value < 0.0001;
Reject H_0. There is sufficient evidence to warrant rejection of the claim that pulse rates of women have a standard deviation equal to 10 beats per minute. Using the normal range of 60 to 100 beats per minute is not very good for estimating σ in this case.

7. H_0: $\sigma = 2.08^\circ\text{F}$; H_1: $\sigma < 2.08^\circ\text{F}$; Test statistic: $\chi^2 = \dfrac{(n-1)s^2}{\sigma^2} = \dfrac{(106-1)0.62^2}{2.08^2} = 9.329$;
P-value $= 0.0000$ (Table: P-value < 0.005); Critical value: $\chi^2 = 74.252$ (Table: $\chi^2 \approx 70.065$);
Reject H_0. There is sufficient evidence to support the claim that body temperatures have a standard deviation less than 2.08°F. It is very highly unlikely that the conclusion in the hypothesis test in Example 5 from Section 8-3 would change because of a standard deviation from a different sample.

8. H_0: $\sigma = 660.2$ hg; H_1: $\sigma \neq 660.2$ hg; Test statistic: $\chi^2 = \dfrac{(n-1)s^2}{\sigma^2} = \dfrac{(30-1)829.5^2}{660.2^2} = 45.780$;
P-value $= 0.0493$ (Table: P-value > 0.02); Critical values: $\chi^2 = 13.121$ and $\chi^2 = 52.336$;
Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that birth weights of girls have the same standard deviation as the birth weights of boys.

9. H_0: $\sigma = 15$; H_1: $\sigma < 15$; Test statistic: $\chi^2 = \dfrac{(n-1)s^2}{\sigma^2} = \dfrac{(18-1)9.6^2}{15^2} = 6.963$;
P-value $= 0.0160$ (Table: P-value < 0.025); Critical value: $\chi^2 = 8.672$;
Reject H_0. There is sufficient evidence to support the claim that IQ scores of physicians have a standard deviation less than 15.

10. H_0: $\sigma = 14.1$; H_1: $\sigma < 14.1$; Test statistic: $\chi^2 = \dfrac{(n-1)s^2}{\sigma^2} = \dfrac{(27-1)9.3^2}{14.1^2} = 11.311$;
P-value $= 0.0056$ (Table: P-value < 0.01); Critical value: $\chi^2 = 12.198$;
Reject H_0. There is sufficient evidence to support the claim that the last class has less variation. The lower variation implies that the scores are closer together, but it does not imply that the scores are higher, so the lower standard deviation does not suggest that the class is doing better.

11. H_0: $\sigma = 4.25$ g; H_1: $\sigma < 4.25$ g; Test statistic: $\chi^2 = \dfrac{(n-1)s^2}{\sigma^2} = \dfrac{(71-1)3.87^2}{4.25^2} = 58.042$;
P-value $= 0.1545$ (Table: P-value > 0.10); Critical value: $\chi^2 = 51.739$;
Fail to reject H_0. There is not sufficient evidence to support the claim that the new process dispenses amounts with a standard deviation less than the standard deviation of 4.25 g for the old process. The new process does not appear to be better in the sense of dispensing amounts that are more consistent.

12. H_0: $\sigma = 7460$ words; H_1: $\sigma > 7460$ words; Test statistic: $\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(56-1)7871^2}{7460^2} = 61.227$;
P-value = 0.2625 (Table: P-value > 0.10); Critical value: $\chi^2 = 82.292$ (Table: $\chi^2 \approx 76.145$);
Fail to reject H_0. There is not sufficient evidence to support the claim that males have a standard deviation that is greater than the standard deviation of 7460 words for females.

13. H_0: $\sigma = 6.0$ lb; H_1: $\sigma \neq 6.0$ lb; Test statistic: $\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(40-1)4.9^2}{6.0^2} = 26.011$;
P-value = 0.1101 (Table: P-value > 0.10); Critical values: $\chi^2 = 19.996$ and $\chi^2 = 65.475$ (Table: 20.707, 66.766);
Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that the amounts of weight loss have a standard deviation equal to 6.0 lb.

14. H_0: $\sigma = 1.5$ mm Hg; H_1: $\sigma > 1.5$ mm Hg; Test statistic: $\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(11-1)9.65^2}{1.5^2} = 413.878$;
P-value = 0.0000 (Table: P-value < 0.005); Critical value: $\chi^2 = 18.307$;
Reject H_0. There is sufficient evidence to support the claim that the sample of devices has a standard deviation greater than 1.5 mm Hg. It appears the accuracy of the tested devices is suspect.

15. H_0: $\sigma = 109.3$ sec; H_1: $\sigma < 109.3$ sec; Test statistic: $\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(10-1)28.601^2}{109.3^2} = 0.616$;
P-value = 0.001 (Table: P-value < 0.005); Critical value: $\chi^2 = 3.325$;
Reject H_0. There is sufficient evidence to support the claim that with a single waiting line, the waiting times have a standard deviation less than 109.3 sec. Because the variation among waiting times appears to be reduced with the single waiting line, patients are happier because their waiting times are closer to being the same.

16. H_0: $\sigma = 0.03$ g; H_1: $\sigma \neq 0.03$ g; Test statistic: $\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(15-1)0.25875^2}{0.03^2} = 10.415$;
P-value = 0.5375 (Table: P-value > 0.20); Critical values: $\chi^2 = 5.629$ and $\chi^2 = 26.119$;
Reject H_0. There is sufficient evidence to warrant rejection of the claim that the weights of bumblebee bats have a standard deviation of 0.03 g. It appears the weights of these bats have a different variation than those from the region in Thailand.

17. Critical $\chi^2 = \frac{1}{2}\left(2.33 + \sqrt{2 \cdot 55 - 1}\right)^2 = 81.54$ (or 81.494 if using $z = 2.326348$ found from technology), which is reasonably close to the value of 22.465 obtained from STATDISK and Minitab.

18. Critical $\chi^2 = 55\left(1 - \frac{2}{9 \cdot 55} + \left(2.33\sqrt{\frac{2}{9 \cdot 55}}\right)\right)^3 = 82.360$ (or 82.309 if using $z = 2.326348$ found from technology), which is very close to the value of 82.292 obtained from STATDISK and Minitab.

Chapter Quick Quiz

1. a. t distribution
   b. normal distribution
   c. chi-square distribution

2. a. two-tailed
   b. left-tailed
   c. right-tailed

3. a. H_0: $p = 0.5$; H_1: $p > 0.5$

   b. Test statistic: $z = \dfrac{0.53 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{532}}} = 1.38$ (using $x = 282$, $z = 1.39$)

   c. Fail to reject H_0.

   d. There is not sufficient evidence to support the claim that the majority of Internet users aged 18–29 use Instagram.

4. P-value $= 0.0100$

5. true

6. false

7. false

8. No, all critical values of χ^2 are always positive.

9. The t test requires that the sample is from a normally distributed population, and the test is robust in the sense that the test works reasonably well if the departure from normality is not too extreme. The χ^2 (chi-square) test is not robust against a departure from normality, meaning that the test does not work well if the population has a distribution that is far from normal.

10. The only true statement is the one given in part (a).

Review Exercises

1. a. false
   b. true
   c. false
   d. false
   e. false

2. H_0: $p = 0.5$; H_1: $p > 0.5$; Test statistic: $z = \dfrac{0.64 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{21{,}346}}} = 40.91$ ($z = 40.90$ using $x = 13{,}661$);

   P-value $= P(z > 40.91) = 0.0001$ (Tech: 0.0000); Critical value: $z = 2.33$;

   Reject H_0. There is sufficient evidence to support the claim that most people believe that the Loch Ness monster exists. Because the sample is a voluntary response sample, the conclusion about the population might not be valid.

3. The data cannot be verified to have a normal distribution, but $n > 30$.
   H_0: $\mu = 5.4$ million cells per microliter; H_1: $\mu < 5.4$ million cells per microliter;

   Test statistic: $t = \dfrac{4.932 - 5.4}{0.504/\sqrt{40}} = -5.873$; Critical value: $t = -2.426$;

   P-value $= 0.0000$ (Table: P-value < 0.005);

   Reject H_0. There is sufficient evidence to support the claim that the sample is from a population with a mean less than 5.4 million cells per microliter. The test deals with the distribution of sample means, not individual values, so the result does not suggest that each of the 40 males has a red blood cell count below 5.4 million cells per microliter.

4. H_0: $p = 0.204$; H_1: $p \neq 0.204$; Test statistic: $z = \dfrac{\frac{124}{572} - 0.204}{\sqrt{\frac{(0.204)(0.796)}{572}}} = 0.76$;

   P-value $= 2 \cdot P(z > 0.76) = 0.4472$ (Tech: 0.4480); Critical values: $z = \pm 1.96$;

   Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that the rate of smoking by adult males is now the same as in 2008. The smoking rate appears to be about the same.

5. The data appear to follow a normal distribution.
H_0: $\mu = 25$ mg; H_1: $\mu \neq 25$ mg;
Test statistic: $t = \dfrac{24.89 - 25}{0.590/\sqrt{15}} = -0.744$; Critical values: $t = \pm 2.145$;
P-value $= 0.4694$ (Table: P-value > 0.10);
Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that the pills come from a population in which the mean amount of atorvastatin is equal to 25 mg.

6. The sample data meet the requirement of having a normal distribution.
H_0: $\mu = 20.16$; H_1: $\mu < 20.16$;
Test statistic: $t = \dfrac{18.76 - 20.16}{1.186/\sqrt{10}} = -3.732$; Critical value: $t = -2.821$;
P-value $= 0.0023$ (Table: P-value < 0.005);
Reject H_0. There is sufficient evidence to support the claim that the population of recent winners has a mean BMI less than 0.16. Recent winners appear to be significantly smaller than those from the 1920s and 1930s.

7. The sample data meet the requirement of having a normal distribution.
H_0: $\sigma = 1.34$; H_1: $\sigma \neq 1.34$; Test statistic: $\chi^2 = \dfrac{(n-1)s^2}{\sigma^2} = \dfrac{(10-1)1.186^2}{1.34^2} = 7.053$;
P-value $= 0.7368$ (Table: P-value > 0.20); Critical values: $\chi^2 = 1.735$ and $\chi^2 = 23.589$;
Fail to reject H_0. There is not sufficient evidence to warrant rejection of the claim that the recent winners have BMI values with variation different from that of the 1920s and 1930s.

8. a. A type I error is the mistake of rejecting a null hypothesis when it is actually true. A type II error is the mistake of failing to reject a null hypothesis when in reality it is false.
   b. Type I error: In reality, the mean BMI is equal to 20.16, but we support the claim that the mean BMI is less than 20.16. Type II error: In reality, the mean BMI is less than 20.16, but we fail to support that claim.

Cumulative Review Exercises

1. a. $\bar{x} = \dfrac{23+26+27+28+29+32+34+38+43+44+45+48+51+51}{14} = 37.1$ deaths
   b. $Q_2 = \dfrac{34+38}{2} = 36.0$ deaths
   c. $s = \sqrt{\dfrac{(23-37.1)^2 + (26-37.1)^2 + \cdots + (51-37.1)^2 + (51-37.1)^2}{14-1}} = 9.8$ deaths
   d. $s^2 = 9.8^2 = 96.8$ deaths^2
   e. range $= 51 - 23 = 28.0$ deaths
   f. The pattern of the data over time is not revealed by the statistics. A time-series graph would be very helpful in understanding the pattern over time.

2. a. ratio
   b. discrete
   c. quantitative
   d. No, the data are from recent and consecutive years, so they are not randomly selected.

3. 99% CI: $\bar{x} \pm t_{\alpha/2}\dfrac{s}{\sqrt{n}} = 37.1 \pm 3.012 \cdot \dfrac{9.84}{\sqrt{14}} \Rightarrow 29.1$ deaths $< \mu < 45.0$ deaths; We have 99% confidence that the limits of 29.1 deaths and 45.0 deaths contain the value of the population mean.

4. The sample data meet the loose requirement of having a normal distribution.
H_0: $\mu = 72.6$ deaths; H_1: $\mu < 72.6$ deaths;

Test statistic: $t = \dfrac{37.07 - 72.6}{9.84/\sqrt{14}} = -13.509$; Critical value: $t = -2.650$;

P-value = 0.0000 (Table: P-value < 0.005);

Reject H_0. There is sufficient evidence to support the claim that the mean number of annual lightning deaths is now less than the mean of 72.6 deaths from the 1980s. Possible factors: Shift in population from rural to urban areas; better lightning protection and grounding in electric and cable and phone lines; better medical treatment of people struck by lightning; fewer people use phones attached to cords; better weather predictions.

5. Because the vertical scale starts at 50 and not at 0, the difference between the number of males and the number of females is exaggerated, so the graph is deceptive by creating the false impression that males account for nearly all lightning strike deaths. A comparison of the numbers of deaths shows that the number of male deaths is roughly 4 times the number of female deaths, but the graph makes it appear that the number of male deaths is around 25 times the number of female deaths.

6. H_0: $p = 0.5$; H_1: $p > 0.5$; Test statistic: $z = \dfrac{\frac{232}{287} - 0.5}{\sqrt{\frac{(0.5)(0.5)}{287}}} = 10.45$;

P-value $= P(z > 10.45) = 0.0001$ (Tech: 0.0000); Critical value: $z = 2.33$;

Reject H_0. There is sufficient evidence to support the claim that the proportion of male deaths is greater than $1/2$. More males are involved in certain outdoor activities such as construction, fishing, and golf.

7. 95% CI: $\hat{p} \pm z_{\alpha/2}\sqrt{\dfrac{\hat{p}\hat{q}}{n}} = \dfrac{232}{287} \pm 1.96\sqrt{\dfrac{\left(\frac{232}{287}\right)\left(\frac{55}{287}\right)}{287}} \Rightarrow 0.762 < p < 0.854$; Because the entire confidence interval is greater than 0.5, it does not seem feasible that males and females have equal chances of being killed by lightning.

8. a. $0.8 \cdot 0.8 \cdot 0.8 = 0.512$

b. $0.2 \cdot 0.2 \cdot 0.2 = 0.008$

c. $1 - 0.2 \cdot 0.2 \cdot 0.2 = 0.992$

d. ${}_5C_3\left(0.8^3\right)\left(0.2^2\right) = 0.205$

e. $\mu = np = 50 \cdot 0.8 = 40$ males; $\sigma = \sqrt{npq} = \sqrt{50 \cdot 0.8 \cdot 0.2} = 2.8$ males

f. Yes, using the range rule of thumb, values above $\mu + 2\sigma = 40.0 + 2(2.8) = 45.6$ are considered significantly high. Since 46 is greater than 45.6, 46 males would be a significantly high number in a group of 50.

9. $\text{OR} = \dfrac{331/4489}{358/5487} = 1.14$; The odds in favor of experiencing migraine headaches is about 1.14 times higher for overweight people than for people with normal weight.