Chapter 10

Hypothesis Tests Regarding a Parameter

Section 10.1

1. Type I error: To reject the null hypothesis when it is true.
   Type II error: To not reject the null hypothesis when the alternative hypothesis is true.

2. Choose $\alpha = 0.01$. We choose a small α to make it difficult to reject H_0 and minimize the chance that we make a Type I error.

3. As we decrease α (the probability of rejecting a true H_0), we are effectively making it less likely that we will reject H_0 since we require stronger evidence against H_0 as α decreases. This means that it is also more likely that we will fail to reject H_0 when H_1 is really true, so β increases. Thus, β increases as α decreases.

4. If the level of significance is $= \alpha = 0.05$, then the probability of making a Type I error is 0.05.

5. Answers will vary. One possibility follows: In a hypothesis test we make a judgment about the validity of a hypothesis based on the available data. If the data contradicts H_0 then we reject H_0. However, if the available data do not contradict H_0, this does not guarantee that H_0 is true. Consider the court system in the U.S., where suspects are assumed to be innocent until proven guilty. An acquittal does not mean the suspect is innocent, merely that there was not enough evidence to reject the assumption of innocence.

6. Answers will vary. One possibility follows: Reasonable doubt is based on logic and reasoning. The phrase "beyond all reasonable doubt" means that there is sufficient and convincing evidence that a reasonable person would not hesitate to act upon it. It is not based on circumstantial evidence or unsubstantiated accusations. The primary difference is that "beyond all doubt" expects every possibility to be considered while "beyond all reasonable doubt" expects that all reasonably likely possibilities have been taken into account.

7. False; sample evidence will never prove a null hypothesis is true. We assume the null is true and try to gather enough evidence to say that the null is not true. Failing to reject a null hypothesis does not imply that the hypothesis is actually true, just that there was not enough evidence to reject the assumption that it is true.

8. False; decreasing the probability of making one type of error will increase the probability of making the other type (see #3 above).

9. Right-tailed since $H_1: \mu > 5$
   Parameter $= \mu$

10. Left-tailed since $H_1: p < 0.2$
    Parameter $= p$

11. Two-tailed since $H_1: \sigma \neq 4.2$
    Parameter $= \sigma$

12. Right-tailed since $H_1: p > 0.76$
    Parameter $= p$

13. Left-tailed since $H_1: \mu < 120$
    Parameter $= \mu$

14. Two-tailed since $H_1: \sigma \neq 7.8$
    Parameter $= \sigma$

15. **(a)** $H_0: p = 0.102$, $H_1: p > 0.102$
    The alternative hypothesis is > because the sociologist believes the percent has increased.

    (b) We make a Type I error if the sample evidence leads us to reject H_0 and conclude that the proportion of registered births to teenage mothers has increased above 0.102 when, in fact, it has not increased above 0.102.

    (c) We make a Type II error if the sample evidence does not lead us to conclude that the proportion of registered births to teenage mothers has increased above 0.102 when, in fact, the proportion of registered births to teenage mothers has increased above 0.102.

16. **(a)** $H_0: \mu = \$1623$; $H_1: \mu \neq \$1623$
The alternative hypothesis is $\neq$ because no direction of change is given. The researcher feels the level has changed but this could mean an increase or a decrease.

(b) We make a Type I error if the sample evidence leads us to reject H_0 and conclude that the mean charitable contribution per household in the U.S. is not $1623 when, in fact, it is $1623.

(c) We make a Type II error if we do not reject the null hypothesis that the mean charitable contribution is $1623 when, in fact, it is different than $1623.

17. **(a)** $H_0: \mu = \$299,800$; $H_1: \mu < \$299,800$
The alternative hypothesis is < because the real estate broker believes the mean price has decreased.

(b) We make a Type I error if the sample evidence leads us to reject H_0 and conclude that the mean price of a single-family home had decreased below $299,800 when, in fact, it has not decreased below $299,800.

(c) We make a Type II error if we do not conclude that the mean price of a single-family home has decreased below $299,800, when, in fact, it has decreased below $299,800.

18. **(a)** $H_0: \mu = 32$ ounces; $H_1: \mu < 32$ ounces
The alternative hypothesis is < because the consumer advocate believes the mean weight of the jars is less than stated on the label.

(b) We make a Type I error if the sample evidence leads us to reject H_0 and conclude that the mean weight of the jars is less than 32 ounces when, in fact, the true mean weight is 32 ounces.

(c) We make a Type II error if we do not reject the null hypothesis that the mean weight of the jars is 32 ounces when, in fact, the true mean weight is less than 32 ounces.

19. **(a)** $H_0: \sigma = 0.7$ p.s.i.; $H_1: \sigma < 0.7$ p.s.i.
The alternative hypothesis is < because the quality control manager believes the standard deviation of the required pressure has been reduced.

(b) We make a Type I error if the sample evidence leads us to reject H_0 and conclude that the standard deviation in the pressure required is less than 0.7 p.s.i when, in fact, the true standard deviation is 0.7 p.s.i.

(c) We make a Type II error if we do not reject the null hypothesis that the standard deviation in the pressure required is 0.7 p.s.i when, in fact, the true standard deviation is less than 0.7 p.s.i.

20. **(a)** $H_0: p = 0.16$; $H_1: p > 0.16$
The alternative hypothesis is > because the school nurse believes the percentage of 6-to 11-year-olds who are overweight is higher than 16%.

(b) We make a Type I error if the sample evidence leads us to reject H_0 and conclude that the proportion of 6- to 11-year-olds who are overweight is higher than 0.16 when, in fact, it is 0.16.

(c) We make a Type II error if we do not reject the null hypothesis that the proportion of 6- to 11-year-olds who are overweight is 0.16 when, in fact, it is more than 0.16.

21. **(a)** $H_0: \mu = \$49.94$; $H_1: \mu \neq \$49.94$
The alternative hypothesis is $\neq$ because no direction of change is given. The researcher feels the mean monthly bill has changed but this could mean an increase or a decrease.

(b) We make a Type I error if the sample evidence leads us to reject H_0 and conclude that the mean monthly cell phone bill is not $49.94 when, in fact, it is $49.94.

(c) We make a Type II error if we do not reject the null hypothesis that the mean monthly cell phone bill is $49.94 when, in fact, it is different than $49.94.

22. **(a)** $H_0: \sigma = 113$; $H_1: \sigma < 113$
The alternative hypothesis is < because the teacher believes the standard deviation of the SAT math scores has decreased.

(b) We make a Type I error if the sample evidence leads us to reject H_0 and conclude that the standard deviation of the SAT math scores is less than 113 points when, in fact, the true standard deviation is 113 points.

(c) We make a Type II error if we do not reject the null hypothesis that the standard deviation of the SAT math scores is 113 points when, in fact, the true standard deviation is less than 113 points.

23. There is sufficient evidence to conclude that the proportion of registered births in the U.S. to teenage mothers has increased above 0.102.

24. There is not sufficient evidence to conclude that the mean charitable contribution per household has changed from $1623.

25. There is not sufficient evidence to conclude that the mean price of a single-family home has decreased from $299,800.

26. There is sufficient evidence to conclude that the mean content of a jar of peanut butter from this manufacturer is less than 32 ounces.

27. There is not sufficient evidence to conclude that the standard deviation in the pressure required has been reduced from 0.7 p.s.i.

28. There is not sufficient evidence to conclude that the percentage of 6- to 11-year-olds who are overweight is higher than 16% in her school district.

29. There is sufficient evidence to conclude that the mean monthly cell phone bill is different from $49.94.

30. There is not sufficient evidence to conclude that the standard deviation in SAT Reasoning test scores has decreased from its 2007 level of 113.

31. There is not sufficient evidence to conclude that the proportion of registered births to teenage mothers has increased above 0.102.

32. There is sufficient evidence to conclude that the mean charitable contribution per household has changed from $1623.

33. There is sufficient evidence to conclude that the mean price of a single-family home has decreased from $299,800.

34. There is not sufficient evidence to conclude that the mean content in a jar of peanut butter from this manufacturer is less than 32 ounces.

35. (a) $H_0: \mu = 54$ quarts; $H_1: \mu > 54$ quarts

(b) Answers may vary. One possibility follows: Congratulations to the marketing department at popcorn.org. After a marketing campaign encouraging people to consume more popcorn, our researchers have determined that the mean annual consumption of popcorn is now greater than 54 quarts, the mean consumption prior to the campaign.

(c) A Type I error has been made, since the true mean consumption has not increased above 54 quarts. The probability of making a Type I error is 0.05.

36. (a) $H_0: \mu = 518$; $H_1: \mu > 518$

(b) There is not sufficient evidence to conclude that the mean SAT Math Reasoning exam of students who take the company's course is more than 518.

(c) A Type II error was committed because the researcher did not reject a false null hypothesis. If $\alpha = 0.01$, the probability of committing a Type I error is 0.01.

(d) To decrease the probability of making a Type II error, we need to increase the probability of making a Type I error. That is, we need to increase the significance level, α.

37. (a) $H_0: p = 0.152$; $H_1: p < 0.152$

(b) There is not sufficient evidence to conclude that changes in the DARE program have resulted in a decrease in the proportion of tenth graders who have tried marijuana.

(c) Since we failed to reject a false null hypothesis, a Type II error was committed.

38. (a) $H_0: p = 0.152$; $H_1: p < 0.152$

(b) Answers may vary. One possibility follows: We are proud to announce a new massage therapy program that is clinically proven to reduce the proportion of individuals who experience migraine headaches.

(c) Since we rejected a true null hypothesis, a Type I error was committed.

39. Let $\mu =$ the mean increase in gas mileage for cars using the Platinum Gasaver. Then the hypotheses would be $H_0 : \mu = 0$ versus $H_1 : \mu > 0$.

40. (a) Let $\mu =$ the mean number of hours that an engine, with Prolong added, can run without oil. The manufacturer's ad seems to imply that $\mu = 4$ hours. If we give the manufacturer the benefit of the doubt then we would test $H_0 : \mu = 4$ hours versus $H_1 : \mu < 4$ hours.

    (b) *Consumer Reports* might reject the null hypothesis and conclude that a car with Prolong will not run 4 hours without oil.

41. Answers will vary. One possibility follows: If you are going to accuse a company of wrongdoing, you should have fairly convincing evidence. In addition, you likely do not want to find out down the road that your accusations were unfounded. Therefore, it is likely more serious to make a Type I error. For this reason, we should probably make the level of significance $\alpha = 0.01$.

Section 10.2

1. The sample must have been obtained using simple random sampling, and either the population from which the sample is selected must be normally distributed or the sample size must be large ($n \geq 30$).

2. When $\alpha = 0.01$ and σ is known, the critical value for a right-tailed test is $z_{0.01} = 2.33$.

3. When $\alpha = 0.05$ and σ is known, the critical values for a two-tailed test are $-z_{0.025} = -1.96$ and $z_{0.025} = 1.96$.

4. To say that methods for testing hypotheses regarding a population mean if σ is known are robust means that minor departures from normality will not adversely affect the results of the test. (However, for small samples, if the data have outliers, then this procedure should not be used.)

5. The *P*-value is the probability of observing a sample statistic as extreme or more extreme than the one observed under the assumption that the null hypothesis is true. A small *P*-value indicates that the observed data are very unlikely to result from chance variation in samples, so this is evidence that the null hypothesis is not true. More specifically, if the *P*-value is less than the level of significance α, the null hypothesis is rejected.

6. If the *P*-value is 0.23, then the probability of obtaining a sample mean of $\bar{x}$ or smaller, assuming that 100 is the true mean, is 0.23. Since this is not an unusual outcome, we would conclude that our sample data do not provide strong enough evidence to reject H_0.

7. If the *P*-value is 0.02, then the probability of obtaining a sample mean more than $|z_0|$ standard deviations away from the hypothesized mean of 50, assuming that 50 is the true mean, is 0.02. Decisions regarding whether to reject H_0 will vary. One possibility follows: Since this is an unusual occurrence, we would conclude that our sample data provide strong enough evidence to reject H_0, although this depends on the level of significance that we are using.

8. Answers will vary. The classical method can be a little easier to perform by hand. However, the *P*-value approach is almost universal in practice, and is more versatile in that it allows a conclusion to be drawn at any level of significance.

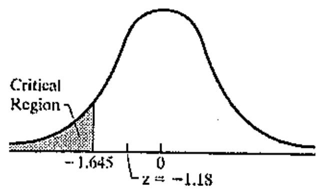
9. Answers will vary. Statistical significance typically refers to absolute differences; that is, whether the observed difference is due to chance. Practical significance typically refers to relative differences; that is, whether the observed difference is large enough to cause concern.

10. True. Performing a two-sided hypothesis test is equivalent to constructing a two-sided confidence interval and determining if it captures the assumed value of the parameter being tested. If the assumed value of the parameter (e.g. μ_0) is not contained within the confidence interval, we have evidence to support the alternative hypothesis so we reject H_0.

11. (a) $z_0 = \dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \dfrac{47.1-50}{12/\sqrt{24}} = -1.18$

(b) This is a left-tailed test, so the critical value is $-z_{0.05} = -1.645$.

(c)

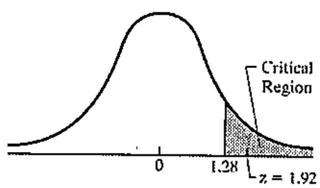


(d) Since $-1.18 > -1.645$, the test statistic is not in the critical region. The researcher will not reject the null hypothesis.

12. (a) $z_0 = \dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \dfrac{42.3-40}{6/\sqrt{25}} = 1.92$

(b) This is a right-tailed test, so the critical value is $z_{0.10} = 1.28$

(c)

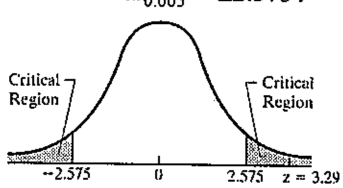


(d) Since $1.92 > 1.28$, the test statistic is in the critical region. The researcher will reject the null hypothesis.

13. (a) $z_0 = \dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \dfrac{104.8-100}{7/\sqrt{23}} = 3.29$

(b) This is a two-tailed test, so the critical values are $\pm z_{0.005} = \pm 2.575$.

(c)

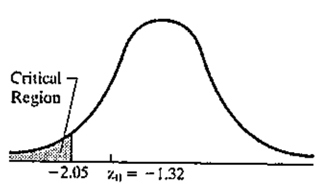


(d) Since $3.29 > 2.575$, the test statistic is in the critical region. The researcher will reject the null hypothesis.

14. (a) $z_0 = \dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \dfrac{76.9-80}{11/\sqrt{22}} = -1.32$

(b) This is a left-tailed test, so the critical value is $-z_{0.02} = -2.05$

(c)



(d) Since $-1.32 > -2.05$, the test statistic is not in the critical region. The researcher will not reject the null hypothesis.

15. (a) The test statistic is

$$z_0 = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{18.3-20}{3/\sqrt{18}} = -2.40.$$

This is a left-tailed test, so we get
P – value = $P(Z < -2.40) = 0.0082$
Fewer than 1 sample in 100 will result in a sample mean as extreme or more extreme as the one we obtained *if* the population mean is $\mu = 20$.

(b) Since P-value $= 0.0082 < 0.0500 = \alpha$, the researcher will reject the null hypothesis.

16. (a) The test statistic is

$$z_0 = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{4.9-4.5}{1.2/\sqrt{13}} = 1.20.$$

This is a right-tailed test so we have
P-value $= P(Z > 1.20)$
$= 1 - 0.8849 = 0.1151$
About 12 samples in 100 will result in a sample mean as extreme or more extreme as the one we obtained *if* the population mean is $\mu = 4.5$.

(b) Since P-value $= 0.1151 > 0.1000 = \alpha$, the researcher will not reject the null hypothesis.

17. (a) No, the population does not need to be normally distributed to compute the P-value since the sample size is large ($n \geq 30$).

(b) The test statistic is

$$z_0 = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{101.2-105}{12/\sqrt{35}} = -1.87.$$

This is a two-tailed test so we have
P-value $= 2 \cdot P(Z < -1.87)$
$= 2 \cdot 0.0307 = 0.0614$
About 6 samples in 100 will result in a sample mean as extreme or more extreme as the one we obtained *if* the population mean is $\mu = 105$.

(c) Since P-value $= 0.0614 > 0.05 = \alpha$, the researcher will not reject the null hypothesis.

(d) For a 95% confidence interval, we use $z_{\alpha/2} = z_{(1-0.95)/2} = z_{0.025}$.

$$\text{Lower bound} = \bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$
$$= 101.2 - 1.96 \cdot \frac{12}{\sqrt{35}} \approx 97.22$$
$$\text{Upper bound} = \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}}$$
$$= 101.2 + 1.96 \cdot \frac{12}{\sqrt{35}} \approx 105.18$$

Since 105 is contained in this interval, we do not reject the null hypothesis.

18. (a) No, the population does not need to be normally distributed to compute the P-value since the sample size is large ($n \ge 30$).

(b) The test statistic is
$$z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{48.3 - 45}{8/\sqrt{40}} = 2.61.$$
This is a two-tailed test so we get
$$\begin{aligned} P\text{-value} &= P(Z < -2.61) + P(Z > 2.61) \\ &= 2P(Z < -2.61) \\ &= 2(0.0045) \\ &= 0.0090 \end{aligned}$$
Fewer than 1 sample in 100 will result in a sample mean such as the one we obtained *if* the population mean is $\mu = 45$.

(c) Since $P\text{-value} = 0.0090 < 0.0500 = \alpha$, the researcher will reject the null hypothesis.

(d) For a 95% confidence interval, we use $z_{\alpha/2} = z_{(1-0.95)/2} = z_{0.025}$.

$$\text{Lower bound} = \bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$
$$= 48.3 - 1.96 \cdot \frac{8}{\sqrt{40}} \approx 45.82$$
$$\text{Upper bound} = \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$
$$= 48.3 + 1.96 \cdot \frac{8}{\sqrt{40}} \approx 50.78$$

Since 45 is not is contained in this interval, we reject the null hypothesis.

19. (a) $H_0: \mu = \$67$ versus $H_1: \mu > \$67$

(b) A P-value of 0.02 means that there is a 0.02 probability of obtaining a sample mean of \$73 or higher from a population whose mean is \$67. So, if we obtained 100 simple random samples of size $n = 40$ from a population whose mean is \$67, we would expect about 2 of these samples to result in sample means of \$73 or higher.

(c) Since $P\text{-value} = 0.02 < \alpha = 0.05$, we reject the statement in the null hypothesis. There is sufficient evidence to conclude that the mean dollar amount withdrawn from a PayEase ATM is more than the mean amount from a standard ATM. That is, the mean dollar amount withdrawn from a PayEase ATM is more than \$67.

20. (a) $H_0: \mu = 63.7$ in. versus $H_1: \mu > 63.7$ in.

(b) A P-value of 0.35 means that there is a 0.35 probability of obtaining a sample mean of 63.9 inches or higher from a population whose mean is 63.9 inches. So, if we obtained 100 simple random samples of size $n = 45$ from a population whose mean is 63.7 inches, we would expect about 35 of these samples to result in sample means of 63.9 inches or higher.

(c) Since $P\text{-value} = 0.35 > \alpha = 0.10$, we do not reject the statement in the null hypothesis. There is not sufficient evidence to conclude that the mean height of women 20 years of age or older is greater than it was in 1990.

21. (a) $H_0: \mu = 21.2$ versus $H_1: \mu > 21.2$

(b) $$z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{21.9 - 21.2}{4.9/\sqrt{200}} = 2.02$$

Classical approach:
This is a right-tailed test, so the critical value is $z_\alpha = z_{0.05} = 1.645$ and the critical region lies to the right of our critical value. Since $z_0 = 2.02 > z_{0.05} = 1.645$, the test statistic falls in the critical region. Therefore, we reject the null hypothesis.

P-value approach:
$$\begin{aligned} P\text{-value} &= P(Z > 2.02) = 1 - P(Z \le 2.02) \\ &= 1 - 0.9783 = 0.0217 \end{aligned}$$
Since $P\text{-value} = 0.0217 < \alpha = 0.05$, we reject the null hypothesis.

(c) There is sufficient evidence to conclude that students who take the core mathematics curriculum score better on the ACT.

22. (a) It is necessary for SAT math scores to be normally distributed because the sample size is small ($n < 30$).

(b) $H_0 : \mu = 516$ versus $H_1 : \mu \neq 516$

(c) Test Statistic:

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{522 - 516}{114 / \sqrt{20}} = 0.24$$

Classical approach:
This is a two-tailed test, so the critical values are $\pm z_{\alpha/2} = \pm z_{0.05} = \pm 1.645$. The critical region lies to the left of -1.645 and to the right of 1.645. Since $z_0 = 0.24$ is between $-z_{0.05} = -1.645$ and $z_{0.05} = 1.645$, the test statistic does not fall in the critical region. Therefore, we do not reject the null hypothesis.

P-value approach:

$$\begin{aligned} P\text{-value} &= P(Z < -0.24) + P(Z > 0.24) \\ &= 2 \cdot P(Z > 0.24) \\ &= 2\left[1 - P(Z \le 0.24)\right] \\ &= 2(1 - 0.5948) \\ &= 0.8104 \quad [\text{Tech: } 0.8139] \end{aligned}$$

Since $P\text{-value} = 0.8104 > \alpha = 0.10$, we do not reject the null hypothesis.

(d) There is not sufficient evidence to conclude that students whose first language is not English score differently on the math portion of the SAT exam.

23. (a) The data are all within the confidence bands of the normal probability plot, which also has a generally linear pattern. The boxplot shows that there are no outliers. Therefore, the conditions for a hypothesis test are satisfied.

(b) Hypotheses: $H_0 : \mu = 0.11$ mg/L versus $H_1 : \mu \neq 0.11$ mg/L

We compute the sample mean to be $\bar{x} = 0.1568$ mg/L.

Test Statistic:

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{0.1568 - 0.11}{0.08 / \sqrt{10}} \approx 1.85$$

Classical approach:
This is a two-tailed test, so the critical values are $\pm z_{\alpha/2} = \pm z_{0.025} = \pm 1.96$. Since $z_0 = 1.85$ is between $-z_{0.05} = -1.96$ and $z_{0.05} = 1.96$, the test statistic does not fall in the critical region. Therefore, we do not reject the null hypothesis.

P-value approach:

$$\begin{aligned} P\text{-value} &= P(Z < -1.85) + P(Z > 1.85) \\ &= 2 \cdot P(Z < -1.85) \\ &= 2(0.0322) \\ &= 0.0644 \quad [\text{Tech: } 0.0643] \end{aligned}$$

Since $P\text{-value} = 0.0644 > \alpha = 0.05$, we do not reject the null hypothesis.

Conclusion: There is not sufficient evidence to indicate that the calcium concentration in rainwater in Chautauqua, New York, has changed since 1990.

24. (a) The data are all within the confidence bands of the normal probability plot, which also has a generally linear pattern. The boxplot shows that there are no outliers. Therefore, the conditions for a hypothesis test are satisfied.

(b) Hypotheses: $H_0 : \mu = 5.27$ versus $H_1 : \mu \neq 5.27$

We compute the sample mean to be $\bar{x} \approx 5.336$

Test Statistic:

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{5.336 - 5.27}{0.27 / \sqrt{16}} \approx 0.98$$

Classical approach:
This is a two-tailed test, so the critical value is $\pm z_{\alpha/2} = \pm z_{0.005} = \pm 2.575$. The critical region lies to the left of -2.575 and to the right of 2.575. Since $z_0 = 0.98$ is between $-z_{0.005} = -2.575$ and $z_{0.005} = 2.575$, the test statistic does not fall in the critical region. Therefore, we do not reject the null hypothesis.

P-value approach:

$$\begin{aligned} P\text{-value} &= P(Z < -0.98) + P(Z > 0.98) \\ &= 2 \cdot P(Z < -0.98) \\ &= 2(0.1635) \\ &= 0.3270 \quad [\text{Tech: } 0.3264] \end{aligned}$$

Since $P\text{-value} = 0.3270 > \alpha = 0.05$, we do not reject the null hypothesis.

Conclusion: There is not sufficient evidence to indicate that the pH in the rain in Columbia River Gorge has changed since 2002.

25. (a) The data are all within the confidence bands of the normal probability plot, which also has a generally linear pattern. The boxplot shows that there are no outliers. Therefore, the conditions for a hypothesis test are satisfied.

(b) Hypotheses: $H_0: \mu = 64.05$ ounces versus $H_1: \mu \neq 64.05$ ounces

We compute the sample mean to be $\bar{x} \approx 64.007$ ounces.

Test Statistic:

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{64.007 - 64.05}{0.06 / \sqrt{22}} = -3.36$$

Classical approach:

This is a two-tailed test, so the critical values are $\pm z_{\alpha/2} = \pm z_{0.005} = \pm 2.575$.

Since $z_0 = -3.36 < -z_{0.005} = -2.575$, the test statistic falls in the critical region. Thus, we reject the null hypothesis.

P-value approach:

$$\begin{aligned} P\text{-value} &= P(Z < -3.36) + P(Z > 3.36) \\ &= 2 \cdot P(Z < -3.36) \\ &= 2(0.0004) \\ &= 0.0008 \end{aligned}$$

Since P-value $= 0.0008 < \alpha = 0.01$, we reject the null hypothesis.

Conclusion: There is sufficient evidence to indicate that the mean amount of juice in each bottle is not 64.05 ounces.

Since the null hypothesis has been rejected, the process should be stopped so the machine can be recalibrated.

(c) Answers will vary. Using $\alpha = 0.1$ means that we will reject a true null 10% of the time. Stopping the machine process to recalibrate, when unnecessary, delays production which can lead to increased costs, lost revenue, and lower profits.

26. (a) The data are all within the confidence bands of the normal probability plot, which also has a generally linear pattern. The boxplot shows that there are no outliers. Therefore, the conditions for a hypothesis test are satisfied.

(b) Hypotheses: $H_0: \mu = 5.6$ miles per gallon versus $H_1: \mu > 5.6$ miles per gallon

We compute the sample mean to be $\bar{x} \approx 6.308$ miles per gallon.

Test Statistic:

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{6.308 - 5.6}{0.5 / \sqrt{12}} = 4.91$$

Classical approach:

This is a right-tailed test, so the critical value is $z_{0.05} = 1.645$. Since $z_0 = 4.91 > z_{0.05} = 1.645$, the test statistic is in the critical region and we reject the null hypothesis.

P-value approach:

P-value $= P(Z > 4.91) < 0.0001$

Since P-value $< 0.05 = \alpha$, we reject the null hypothesis.

Conclusion: There is sufficient evidence to conclude that the catalyst has increased the gas mileage above 5.6 miles per gallon.

27. (a) Assessed home values are typically skewed right because there are a few homes with very high assessed values. Because the shape of the distribution of assessed home values is not approximately normal, we need a large sample size so that we can use the Central Limit Theorem, which means that the distribution of the sample mean will be approximately normal.

(b) Hypotheses: $H_0: \mu = \$11.38$ versus $H_1: \mu < \$11.38$

Test Statistic:

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{11.09 - 11.38}{8.02 / \sqrt{50}} \approx -0.26$$

Classical approach:

We use an $\alpha = 0.05$ level of significance. Since this is a left-tailed test, our critical value is $-z_\alpha = -z_{0.05} = -1.645$. Since $z_0 = -0.26 > -z_{0.05} = -1.645$, the test statistic does not fall in the critical region and we do not reject the null hypothesis.

P-value approach:

P-value $= P(Z < -0.26) = 0.3974$ [Tech: 0.3991]. Since P-value $> \alpha = 0.05$, we do not reject the null hypothesis.

Conclusion: There is not sufficient evidence to indicate that rural households pay less per \$1000 assessed valuation than all homes in the United States.

28. (a) Since the sample size is small $(n = 15 < 30)$, the population of birth weights must be approximately normal, and the data cannot have any outliers.

(b) Hypotheses: $H_0 : \mu = 3270$ grams versus $H_1 : \mu < 3270$ grams

Test Statistic:

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{3080 - 3270}{610 / \sqrt{15}} = -1.21$$

Classical approach:
This is a left-tailed test so our critical value is $-z_\alpha = -z_{0.1} = -1.28$.

Since $z_0 = -1.21 > -z_{0.1} = -1.28$, the test statistic does not fall in the critical region and we do not reject the null hypothesis.

P-value approach:
P-value $= P(Z < -1.21) = 0.1131$ [Tech: 0.1138]. Since P-value $= 0.1131 > \alpha = 0.1$, we do not reject the null hypothesis.

Conclusion: There is not sufficient evidence to indicate that mothers who are less than 15 years of age have babies with lower birth weights than the general population.

29. (a) The data are not normally distributed and have outliers, so a large sample is needed to conduct a hypothesis about the mean.

(b) Hypotheses: $H_0 : \mu = 24.1$ million shares versus $H_1 : \mu \neq 24.1$ million shares

Test Statistic:

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{21.97 - 24.1}{7.6 / \sqrt{40}} = -1.77$$

Classical approach:
This is a two-tailed test, so the critical values are $\pm z_{\alpha/2} = \pm z_{0.025} = \pm 1.96$. Since $z_0 = -1.77$ is between $-z_{0.025} = -1.96$ and $z_{0.025} = 1.96$, the test statistic does not fall in a critical region. Thus, we do not reject the null hypothesis.

P-value approach:

$$\begin{aligned} P\text{-value} &= P(Z < -1.77) + P(Z > 1.77) \\ &= 2 \cdot P(Z < -1.77) \\ &= 2(0.0384) \\ &= 0.0768 \text{ [Tech: 0.0763]} \end{aligned}$$

Since P-value $= 0.0768 > \alpha = 0.05$, we do not reject the null hypothesis.

Conclusion: There is not sufficient evidence to indicate that the volume of Dell stock is different from what it was in 2002.

30. (a) The data are not normally distributed and have outliers, so a large sample is needed to conduct a hypothesis about the mean.

(b) Hypotheses: $H_0 : \mu = 10.7$ million shares versus $H_1 : \mu \neq 10.7$ million shares

Test Statistic:

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{5.97 - 10.7}{5.7 / \sqrt{35}} = -4.91$$

Classical approach:
This is a two-tailed test, so the critical values are $\pm z_{\alpha/2} = \pm z_{0.025} = \pm 1.96$. Since $z_0 = -4.91 < -z_{0.025} = -1.96$, the test statistic falls in a critical region. Thus, we reject the null hypothesis.

P-value approach:
The critical value is so far to the left that the P-value < 0.0001. Since the P-value is less than $\alpha = 0.05$, we reject the null hypothesis.

Conclusion: There is sufficient evidence to indicate that the volume of Google stock is different from what it was in 2005.

31. Hypotheses: $H_0 : \mu = 0.11$ mg/L versus $H_1 : \mu \neq 0.11$ mg/L. From Problem 23, we have $\bar{x} = 0.1568$ mg/L, $\sigma = 0.08$ mg/L, and $n = 10$. Using $\alpha = 0.01$, we have $z_{\alpha/2} = z_{0.005} = 2.575$.

$$\begin{aligned} \text{Lower bound} &= \bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 0.1568 - 2.575 \cdot \frac{0.08}{\sqrt{10}} \\ &\approx 0.0917 \text{ [Tech: 0.0916] mg/L} \end{aligned}$$

$$\begin{aligned} \text{Upper bound} &= \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \\ &= 0.1568 + 2.575 \cdot \frac{0.08}{\sqrt{10}} \\ &\approx 0.2219 \text{ [Tech: 0.2220] mg/L} \end{aligned}$$

Since 0.11 lies in this interval, we do not reject the null hypothesis.

Conclusion: There is not sufficient evidence to indicate that the calcium concentration in rainwater in Chautauqua, New York, has changed since 1990.

32. Hypotheses: $H_0: \mu = 5.27$ versus $H_1: \mu \neq 5.27$. From Problem 24, we have $\bar{x} \approx 5.336$, $\sigma = 0.27$, and $n = 16$. Using $\alpha = 0.05$, we have $z_{\alpha/2} = z_{0.025} = 1.96$.

$$\text{Lower bound } = \bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 5.336 - 1.96 \cdot \frac{0.27}{\sqrt{16}} \approx 5.204$$

$$\text{Upper bound } = \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 5.336 + 1.96 \cdot \frac{0.27}{\sqrt{16}} \approx 5.468 \text{ [Tech: 5.469]}$$

Since 5.27 lies in this interval, we do not reject the null hypothesis.
Conclusion: There is not sufficient evidence to indicate that the pH in the rain in Columbia River Gorge has changed since 2002.

33. Hypotheses: $H_0: \mu = 24.1$ million shares versus $H_1: \mu \neq 24.1$ million shares
From Problem 29, $\bar{x} = 21.97$ million shares, $\sigma = 7.6$ million shares, and $n = 40$. Using $\alpha = 0.05$, we have $z_{\alpha/2} = z_{0.025} = 1.96$.

$$\text{Lower bound} = \bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 21.97 - 1.96 \cdot \frac{7.6}{\sqrt{40}} \approx 19.615 \text{ million shares}$$

$$\text{Upper bound} = \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 21.97 + 1.96 \cdot \frac{7.6}{\sqrt{40}} \approx 24.325 \text{ million shares}$$

Since 24.1 lies in this interval, we do not reject the null hypothesis.
Conclusion: There is not sufficient evidence to indicate that the volume of Dell stock is different from what it was in 2002.

34. Hypotheses: $H_0: \mu = 10.7$ million shares versus $H_1: \mu \neq 10.7$ million shares
From Problem 30, $\bar{x} = 5.97$ million shares, $\sigma = 5.7$ million shares, and $n = 35$. Using $\alpha = 0.05$, we have $z_{\alpha/2} = z_{0.025} = 1.96$.

$$\text{Lower bound } = \bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 5.97 - 1.96 \cdot \frac{5.7}{\sqrt{35}} \approx 4.082 \text{ million shares}$$

$$\text{Upper bound } = \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 5.97 + 1.96 \cdot \frac{5.7}{\sqrt{35}} \approx 7.858 \text{ million shares}$$

Since 10.7 does not lie in this interval, we reject the null hypothesis.
Conclusion: There is sufficient evidence to indicate that the volume of Google stock is different from what it was in 2005.

35. **(a)** $H_0: \mu = 515$ versus $H_1: \mu > 515$

(b) The test statistic is

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{519 - 515}{114/\sqrt{1800}} = 1.49$$

Classical approach:
This is a right-tailed test, so our critical value is $z_\alpha = z_{0.10} = 1.28$. Since $z_0 = 1.49 > z_{0.10} = 1.28$, the test statistic falls in the critical region. Thus, we reject the null hypothesis.

P-value approach:

$$P\text{-value} = P(Z > 1.49) = 1 - P(Z \le 1.49) = 1 - 0.9319 = 0.0681 \text{ [Tech: 0.0683]}$$

Since $P\text{-value} = 0.0681 < \alpha = 0.10$, we reject the null hypothesis.
Conclusion: There is sufficient evidence to conclude that the mean score of students taking this review is greater than 515.

(c) Answers will vary. In some states this would be regarded as a highly significant increase, although in most states it is not likely to be thought of as an increase that has any practical significance.

(d) The test statistic is now

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{519 - 515}{114/\sqrt{400}} = 0.70.$$

Classical approach:
This is a right-tailed test, so our critical value is $z_\alpha = z_{0.10} = 1.28$. Since

$z_0 = 0.70 < z_{0.10} = 1.28$, the test statistic does not fall in the critical region. Thus, we do not reject the null hypothesis.

P-value approach:

$$\begin{aligned} P\text{-value} &= P(Z > 0.70) \\ &= 1 - P(Z \le 0.70) \\ &= 1 - 0.7580 \\ &= 0.2420 \quad \text{[Tech: 0.2414]} \end{aligned}$$

Since $P\text{-value} = 0.2420 > \alpha = 0.10$, we do not reject the null hypothesis.

Conclusion: There is not sufficient evidence to conclude that the mean score of students taking this review is greater than 515.

The sample size can dramatically alter the conclusion of a hypothesis test. It is often possible to make even slight changes statistically significant by selecting a large enough sample size.

36. (a) $H_0 : \mu = 0$ pounds vs $H_1 : \mu > 0$ pounds

(b) The test statistic is

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{0.9 - 0}{7.2 / \sqrt{950}} = 3.85 .$$

Classical approach:

This is a right-tailed test, so our critical value is $z_\alpha = z_{0.10} = 1.28$. Since $z_0 = 3.85 > z_{0.10} = 1.28$, the test statistic falls in the critical region. So we reject the null hypothesis.

P-value approach:

$$\begin{aligned} P\text{-value} &= P(Z > 3.85) \\ &= 1 - P(Z \le 3.85) \\ &= 1 - 0.9999 \\ &= 0.0001 \end{aligned}$$

Since $P\text{-value} = 0.0001 < \alpha = 0.10$, we reject the null hypothesis.

Conclusion: There is sufficient evidence to conclude that the dietary supplement helps people lose weight.

(c) Answers will vary. There does not appear to be any practical significance in weight loss. A mean weight loss of less than one pound could be attributed to other factors such as water weight or some other change in health habits (e.g. taking stairs instead of elevators).

(d) The test statistic is now

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{0.9 - 0}{7.2 / \sqrt{40}} = 0.79 .$$

Classical approach:

This is a right-tailed test, so our critical value is $z_\alpha = z_{0.10} = 1.28$. Since $z_0 = 0.79 < z_{0.10} = 1.28$, the test statistic does not fall in the critical region and the null hypothesis is not rejected.

P-value approach:

$$\begin{aligned} P\text{-value} &= P(Z > 0.79) \\ &= 1 - P(Z \le 0.79) \\ &= 1 - 0.7852 \\ &= 0.2148 \quad \text{[Tech: 0.2146]} \end{aligned}$$

Since $P\text{-value} = 0.2148 > \alpha = 0.10$, we do not reject the null hypothesis.

Conclusion: There is not sufficient evidence to conclude that the dietary supplement helps people lose weight.

The sample size can dramatically alter the conclusion of a hypothesis test. It is often possible to make even slight changes statistically significant by selecting a large enough sample size.

37. (a) Answers will vary.

(b) Since the samples all come from a population that has mean equal to 80, approximately 10% of them, that is 5 out of the 50 samples, should give a sample mean that is in the critical region at a 10% level of significance.

(c) Answers will vary.

(d) Answers will vary. The true mean is 80, so the null hypothesis is correct. Thus, to reject it would be to commit a Type I error.

38. (a) Answers will vary.

(b) Since the samples all come from a population that has mean equal to 8, approximately 5% of them, that is 2 out of the 40 samples, should give a sample mean that is in the critical region at a 5% level of significance.

(c) Answers will vary.

(d) Answers will vary. The true mean is 8, so the null hypothesis is correct. Thus, to reject it would be to commit a Type I error.

39. Yes. Because the head of institutional research has access to the entire population, inference is unnecessary. He can say with 100% confidence that the mean age decreased because the mean age in the current semester is less than the mean age is 1995.

40. (a) *Step 1: Research objective.* To determine whether there is an association between distance of home address at birth from high voltage power lines and the incidence of leukemia and other cancers in children in England and Wales.

Step 2: Collect information. This is a case-control study, which is an observational study in which two groups are compared. Here the two groups are children who were born near high-voltage lines and those who were not born near high-voltage lines.

Step 3: Organize and summarize the information. The statistics based on the sample are based on relative risk. The relative risk of leukemia for children born within 200 meters of high-voltage power lines was 1.69 (these children were 1.69 times more likely to contract cancer than those who lived more than 600 meters from the power lines). The relative risk of leukemia of children who lived between 200 and 600 meters from high-voltage lines was 1.23 when compared with children who lived more than 600 meters from the power line.

Step 4: Draw conclusions. The 95% confidence interval for relative risk of children who lived within 200 meters of the high-voltage power lines compared to those who lived more than 600 meters from the power lines was lower bound: 1.13, upper bound: 2.53. The researchers are 95% confident that the relative risk is between 1.13 and 2.53. Because a relative risk of 1 would imply no association, and this interval does not contain 1, we would reject the null hypothesis of no association. The 95% confidence interval for relative risk for children who lived between 200 and 600 meters of high-voltage power lines compared to those who lived more than 600 meters from the power lines was lower bound: 1.02, upper bound: 1.49. The researchers are 95% confident that the relative risk is between 1.02 and 1.49. Because a relative risk of 1 would imply no association, and this interval does not contain 1, we would reject the null hypothesis of no association. There is a relation between the distance one is from the power lines and the relative risk of contracting leukemia with a P-value < 0.01. The likelihood of observing the relation observed when there is no relation between proximity to the power lines and leukemia is 0.01. There is an association between childhood leukemia and proximity of home address at birth to high-voltage power lines.

(b) It would be highly unethical to conduct an experiment such as this on human subjects, so the researchers conducted an observational study.

A case-control study is a retrospective study in which individuals with a certain characteristic are matched with those that do not. In this situation, we need a case-control study to help diminish the likelihood of lurking variables. We must assume that the subjects are similar in all other traits that might affect the likelihood of contracting leukemia.

Section 10.3

1. A hypothesis regarding a population mean with σ unknown can be tested provided that the sample is obtained using simple random sampling, and the population from which the sample is drawn is either normally distributed with no outliers or the sample size n is larger than 30.

2. For $\alpha = 0.01$ in a right-tailed test when σ is unknown with 15 degrees of freedom, the critical value is $t_{0.01} = 2.602$.

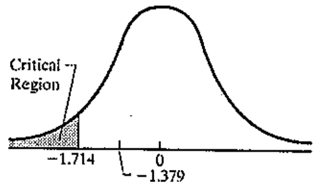
3. For $\alpha = 0.05$ in a two-tailed test with 12 degrees of freedom when σ is unknown, the critical values are $\pm t_{\alpha/2} = \pm t_{0.025} = \pm 2.179$.

4. For $\alpha = 0.05$ in a left-tailed test with 19 degrees of freedom when σ is unknown, the critical value is $-t_{0.05} = -1.729$.

5. (a) $t_0 = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} = \dfrac{47.1 - 50}{10.3/\sqrt{24}} = -1.379$

(b) This is a left-tailed test with $24 - 1 = 23$ degrees of freedom, so the critical value is $-t_{0.05} = -1.714$.

(c)

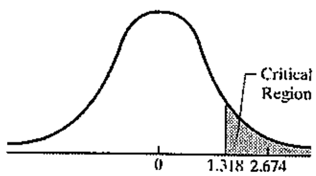


(d) Since the test statistic is not in the critical region $(-1.379 > -1.714)$, the researcher will not reject the null hypothesis.

6. (a) $t_0 = \dfrac{\overline{x} - \mu_0}{s / \sqrt{n}} = \dfrac{42.3 - 40}{4.3 / \sqrt{25}} = 2.674$

(b) This is a right-tailed test with $25 - 1 = 24$ degrees of freedom, so the critical value is $t_{0.10} = 1.318$

(c)

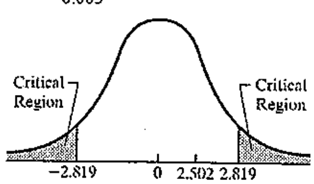


(d) Since the test statistic is in the critical region $(2.674 > 1.318)$, the researcher will reject the null hypothesis.

7. (a) $t_0 = \dfrac{\overline{x} - \mu_0}{s / \sqrt{n}} = \dfrac{104.8 - 100}{9.2 / \sqrt{23}} = 2.502$

(b) This is a two-tailed test with $23 - 1 = 22$ degrees of freedom, so the critical values are $\pm t_{0.005} = \pm 2.819$.

(c)



(d) Since the test statistic is not in the critical region $(-2.819 < 2.502 < 2.819)$, the researcher will not reject the null hypothesis.

(e) Using $\alpha = 0.01$ with 22 degrees of freedom, we have $t_{\alpha/2} = \pm t_{0.005} = \pm 2.819$.

$$\text{Lower bound} = \overline{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$
$$= 104.8 - 2.819 \cdot \frac{9.2}{\sqrt{23}}$$
$$\approx 99.39$$

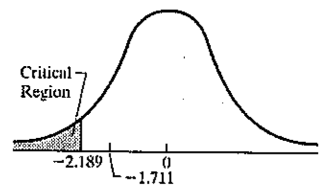
$$\text{Upper bound} = \overline{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$
$$= 104.8 + 2.819 \cdot \frac{9.2}{\sqrt{23}}$$
$$\approx 110.21$$

Because this confidence interval includes the hypothesized mean of 100, the researcher will not reject the null hypothesis.

8. (a) $t_0 = \dfrac{\overline{x} - \mu_0}{s / \sqrt{n}} = \dfrac{76.9 - 80}{8.5 / \sqrt{22}} = -1.711$

(b) This is a left-tailed test with $22 - 1 = 21$ degrees of freedom, so the critical value is $-t_{0.02} = -2.189$.

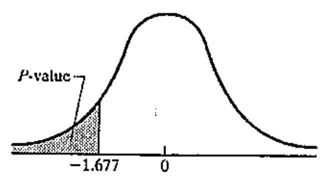
(c)



(d) Since the test statistic is not in the critical region $(-1.711 > -2.189)$, the researcher will not reject the null hypothesis.

9. (a) $t_0 = \dfrac{\overline{x} - \mu_0}{s / \sqrt{n}} = \dfrac{18.3 - 20}{4.3 / \sqrt{18}} = -1.677$.

(b)



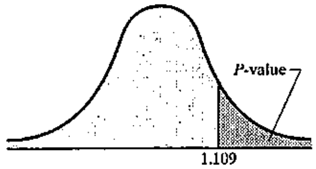
(c) This is a left-tailed test with $18 - 1 = 17$ degrees of freedom. The P-value is the area under the t-distribution to the left of the test statistic, $t_0 = -1.677$. Because of symmetry, the area under the distribution to the left of -1.677 equals the area under the distribution to the right of 1.677. From the t-distribution table (Table VI) in the row corresponding to 17 degrees of freedom, since $t = 1.677$ is between 1.333 and 1.740, whose right-tail areas are 0.10 and 0.05, respectively. So, $0.05 < P\text{-value} < 0.10$ [Tech: 0.0559].

Interpretation: If we obtain 100 random samples of size 18, we would expect about 6 of the samples to result in a sample mean of 18.3 or less *if* the population mean is $\mu = 20$.

(d) Since $P\text{-value} > \alpha = 0.05$, the researcher will not reject the null hypothesis.

10. (a) $t_0 = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} = \dfrac{4.9 - 4.5}{1.3/\sqrt{13}} = 1.109$.

(b)



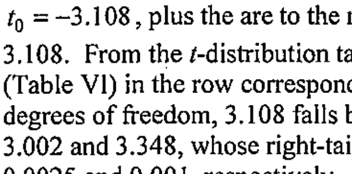
(c) This is a right-tailed test with $13 - 1 = 12$ degrees of freedom. The *P*-value is the area under the *t*-distribution to the right of the test statistic, $t_0 = 1.109$. From the *t*-distribution table (Table VI) in the row corresponding to 12 degrees of freedom, since 1.109 is between 1.083 and 1.356, whose right-tail areas are 0.15 and 0.10, respectively. So, $0.10 < P\text{-value} < 0.15$ [Tech: $P\text{-value} = 0.1445$].
Interpretation: If we obtain 100 random samples of size $n = 13$, we would expect about 15 of the samples to result in a sample mean of 4.9 or higher *if* the population mean is $\mu = 4.5$.

(d) Since $P\text{-value} > 0.10 = \alpha$, the researcher will not reject the null hypothesis.

11. (a) No, because this sample is large ($n \geq 30$).

(b) $t_0 = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} = \dfrac{101.9 - 105}{5.9/\sqrt{35}} = -3.108$.

(c)



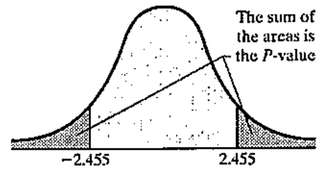
(d) This is a two-tailed test with $35 - 1 = 34$ degrees of freedom. The *P*-value of this two tailed test is the area to the left of $t_0 = -3.108$, plus the are to the right of 3.108. From the *t*-distribution table (Table VI) in the row corresponding to 34 degrees of freedom, 3.108 falls between 3.002 and 3.348, whose right-tail areas are 0.0025 and 0.001, respectively. We must double these values in order to get the total area in both tails: 0.005 and 0.002. Thus, $0.002 < P\text{-value} < 0.005$ [Tech: $P\text{-value} = 0.0038$].
Interpretation: If we obtain 1000 random samples of size 35, we would expect about 4 of the samples to result in a sample mean as extreme or more extreme than the one observed *if* the population mean is $\mu = 105$.

(e) Since $P\text{-value} < \alpha = 0.010$, we reject the null hypothesis.

12. (a) No, because this sample is large ($n \geq 30$).

(b) $t_0 = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} = \dfrac{48.3 - 45}{8.5/\sqrt{40}} = 2.455$.

(c)



(d) This is a two-tailed test with $40 - 1 = 39$ degrees of freedom. The *P*-value of this two tailed test is the area to the right of $t_0 = 2.455$, plus the are to the left of -2.455. From the *t*-distribution table (Table VI) in the row corresponding to 39 degrees of freedom, 2.455 falls between 2.426 and 2.708, whose right-tail areas are 0.01 and 0.005, respectively. We must double these values in order to get the total area in both tails: 0.02 and 0.01. Thus, $0.01 < P\text{-value} < 0.02$ [Tech: $P\text{-value} = 0.0186$].
Interpretation: If we obtain 100 random samples of size $n = 40$, we would expect about 2 of the samples to result in a sample mean as extreme or more extreme than the one observed *if* the population mean is $\mu = 45$.

(e) Since $P\text{-value} > 0.01 = \alpha$, we do not reject the null hypothesis.

(f) Using $\alpha = 0.01$ with 39 degrees of freedom, we have $t_{\alpha/2} = \pm t_{0.005} = \pm 2.708$.

$$\text{Lower bound} = \bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 48.3 - 2.708 \cdot \frac{8.5}{\sqrt{40}} \approx 44.66$$

$$\text{Upper bound} = \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 48.3 + -2.708 \cdot \frac{8.5}{\sqrt{40}} \approx 51.94$$

Because this confidence interval includes the hypothesized mean of 45, the researcher will not reject the null hypothesis.

13. Hypotheses: $H_0: \mu = 9.02 \text{ cm}^3$ versus $H_1: \mu < 9.02 \text{ cm}^3$

Test Statistic:

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{8.10 - 9.02}{0.7/\sqrt{12}} = -4.553$$

$\alpha = 0.01$; d.f. $= n-1 = 12-1 = 11$

Classical approach:
Because this is a left-tailed test with 11 degrees of freedom, the critical value is $-t_{0.01} = -2.718$. Since $t_0 = -4.553 < -t_{0.01} = -2.718$, the test statistic falls within the critical region. We reject the null hypothesis.

P-value approach:
This is a left-tailed test with 11 degrees of freedom. The P-value is the area under the t-distribution to the left of the test statistic, $t_0 = -4.553$. Because of symmetry, the area under the distribution to the left of -4.553 equals the area under the distribution to the right of 4.553. From the t-distribution table (Table VI) in the row corresponding to 11 degrees of freedom, 4.553 falls to the right of 4.437, whose right-tail area is 0.0005. So, P-value < 0.0005 [Tech: P-value= 0.0004]. Since P-value $< \alpha = 0.01$, we reject the null hypothesis.

Conclusion: There is sufficient evidence to conclude that the mean hippocampal volume in alcoholic adolescents is less than the normal volume of 9.02 cm^3.

14. Hypotheses: $H_0: \mu = 355.7$ mg versus $H_1: \mu > 355.7$ mg

Test Statistic

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{396.9 - 355.7}{45.4/\sqrt{12}} = 3.144$$

$\alpha = 0.05$; d.f. $= n-1 = 12-1 = 11$

Classical approach:
Because this is a right-tailed test with 11 degrees of freedom, the critical value is $t_{0.05} = 1.796$. Since $t_0 = 3.144 > t_{0.05} = 1.796$, the test statistic falls within the critical region. We reject the null hypothesis.

P-value approach:
This is a right-tailed test with 11 degrees of freedom. The P-value is the area under the t-distribution to the right of the test statistic, $t_0 = 3.144$. From the t-distribution table in the row corresponding to 11 degrees of freedom, 3.144 falls between 3.106 and 3.497, whose right-tail areas are 0.005 and 0.0025, respectively. So, $0.0025 < P$-value < 0.005 [Tech: 0.0047]. Since P-value $< \alpha = 0.05$, we reject the null hypothesis.

Conclusion: There is sufficient evidence to conclude that the mean weight of kidneys of first-generation mice whose parents were both exposed to para-nonylphenol is greater than 355.7 mg, the mean weight of the kidneys of unexposed mice.

15. Hypotheses: $H_0: \mu = 703.5$ versus $H_1: \mu > 703.5$

Test Statistic

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{714.2 - 703.5}{83.2/\sqrt{40}} = 0.813$$

$\alpha = 0.05$; d.f. $= n-1 = 40-1 = 39$

Classical approach:
Because this is a right-tailed test with 39 degrees of freedom, the critical value is $t_{0.05} = 1.685$. Since $t_0 = 0.813 < t_{0.05} = 1.685$, the test statistic does not fall within the critical region. We do not reject the null hypothesis.

P-value approach:
This is a right-tailed test with 39 degrees of freedom. The P-value is the area under the t-distribution to the right of the test statistic, $t_0 = 0.813$. From the t-distribution table in the row corresponding to 39 degrees of freedom, 0.813 falls between 0.681 and 0.851, whose right-tail areas are 0.25 and 0.20, respectively. So, $0.20 < P$-value < 0.25 [Tech: P-value = 0.2105]. Since P-value $> \alpha = 0.05$, we do not reject the null hypothesis.

Conclusion: There is not sufficient evidence to conclude that the mean FICO score of high-income individuals is greater than that of the general population. In other words, it is not unlikely to obtain a mean credit score of 714.2 even though the true population mean credit score is 703.4.

16. Hypotheses: $H_0: \mu = 154.8$ minutes versus $H_1: \mu < 154.8$ minutes

Test Statistic:

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{128.7 - 154.8}{46.5/\sqrt{50}} = -3.969$$

$\alpha = 0.05$; d.f. $= n-1 = 50-1 = 49$

Classical approach:
This is a left-tailed test with 49 degrees of freedom. However, since our t-distribution table does not contain a row for 49, we use df = 50. The critical value is $-t_{0.05} = -1.676$. Since $t_0 = -3.969 < -t_{0.05} = -1.676$, the test statistic falls within the critical region. We reject the null hypothesis.

P-value approach:
This is a left-tailed test with 49 degrees of freedom. The P-value is the area under the t-distribution to the left of the test statistic $t_0 = -3.969$. Because of symmetry, the area under the distribution to the left of -3.969 equals the area under the distribution to the right of 3.969. From the t-distribution table (Table VI) in the row corresponding to 50 degrees of freedom (since the table does not contain a row for df = 49), 3.969 falls to the right of 3.496, whose right-tail area is 0.0005. So, P-value < 0.0005 [Tech: P-value= 0.0001]. Since $P\text{-value} < \alpha = 0.05$, we reject the null hypothesis.

Conclusion: There is sufficient evidence to conclude that Internet users watch less television per day than the typical American.

17. (a) Hypotheses: $H_0: \mu = 98.6°\text{F}$ versus $H_1: \mu < 98.6°\text{F}$

Test statistic:

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{98.2 - 98.6}{0.7/\sqrt{700}} = -15.119.$$

This is a left-tailed test with $n - 1$ = 700 – 1 = 699 degrees of freedom. However, since our t-distribution table does not contain a row for 699, we use df = 1000 (the closest one to 699). So, the critical value is $-t_{0.01} = -2.330$. Since $t_0 = -15.119 < \alpha = -2.330$, the test statistic falls in the critical region and we reject the null hypothesis. There is sufficient evidence to conclude that the mean temperature of humans is less than 98.6° F.

(b) This is a left-tailed test with 699 degrees of freedom. The P-value is the area under the t-distribution to the left of the test statistic $t_0 = -15.119$. Because of symmetry, the area under the distribution to the left of -15.119 equals the area under the distribution to the right of 15.119. From the t-distribution table (Table VI) in the row corresponding to 1000 degrees of freedom (the row closest to df = 699), 15.119 falls to the right of 3.300, whose right-tail area is 0.0005. So, P-value < 0.0005 [Tech: P-value < 0.0001].

Interpretation: If we obtain 10,000 random samples of size n = 700, we would expect only about 1 of the samples to result in a sample mean of 98.2° F or less *if* the population mean is $\mu = 98.6°$ F.

18. (a) Hypotheses: $H_0: \mu = 98.6°\text{F}$ versus $H_1: \mu < 98.6°\text{F}$

Test statistic:

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{98.4 - 98.6}{0.7/\sqrt{123}} = -3.169.$$

This is a left-tailed test with $n - 1$ = 123 – 1 = 122 degrees of freedom. However, since our t-distribution table does not contain a row for 122, we use df = 100. So, the critical value is $-t_{0.01} = -2.364$. Since $t_0 = -3.169 < -t_{0.01} = -2.364$, the test statistic is in the critical region. Thus, we reject the null hypothesis. There is sufficient evidence to conclude that the mean temperature of women is less than 98.6 F.

(b) This is a left-tailed test with 122 degrees of freedom. The P-value is the area under the t-distribution to the left of the test statistic $t_0 = -3.169$. Because of symmetry, the area under the distribution to the left of -3.169 equals the area under the distribution to the right of 3.169. From the t-distribution table (Table VI) in the row corresponding to 100 degrees of freedom (the row closest to df = 122), 3.169 falls between 2.871 and 3.174, whose right-tail areas are 0.0025 and 0.001. Thus, $0.001 < P\text{-value} < .0025$ [Tech: 0.0010].

Interpretation: If we obtain 1000 random samples of size n = 123, we would expect about 1 of the samples to result in a sample mean of 98.4° F or less *if* the population mean is $\mu = 98.6°$ F.

19. Hypotheses: $H_0: \mu = 40.7$ years versus $H_1: \mu \neq 40.7$ years

Using $\alpha = 0.05$ with $n-1 = 32-1 = 31$ degrees of freedom, we have $t_{\alpha/2} = t_{0.025} = 2.037$.

$$\begin{aligned}\text{Lower bound} &= \bar{x} - t_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \\ &= 38.9 - 2.037 \cdot \frac{9.6}{\sqrt{32}} \\ &= 35.44 \text{ years}\end{aligned}$$

$$\begin{aligned}\text{Upper bound} &= \bar{x} + t_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \\ &= 38.9 + 2.037 \cdot \frac{9.6}{\sqrt{32}} \\ &= 42.36 \text{ years}\end{aligned}$$

Because this interval includes the hypothesized mean 40.7 years, we do not reject the null hypothesis. Thus, there is not sufficient evidence to conclude that the mean age of death-row inmates has change since 2002.

20. Hypotheses: $H_0: \mu = \$1493$ versus $H_1: \mu \neq \$1493$

Using $\alpha = 0.05$ with $n-1 = 35-1 = 34$ degrees of freedom, we have $t_{\alpha/2} = t_{0.025} = 2.032$.

$$\begin{aligned}\text{Lower bound} &= \bar{x} - t_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \\ &= 1618 - 2.032 \cdot \frac{321}{\sqrt{35}} \\ &= \$1507.75\end{aligned}$$

$$\begin{aligned}\text{Upper bound} &= \bar{x} + t_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \\ &= 1618 + 2.032 \cdot \frac{321}{\sqrt{35}} \\ &= \$1728.25\end{aligned}$$

Because this interval does not include the hypothesized mean \$1493, we reject the null hypothesis. Thus, there is significant evidence to conclude that the mean expenditure (adjusted for inflation) for energy has changed since 2001.

21. (a) The plotted data 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. Therefore, the conditions for testing the hypothesis are satisfied.

(b) Hypotheses: $H_0: \mu = 1.68$ inches versus $H_1: \mu \neq 1.68$ inches.

We compute the sample mean and sample standard deviation to be $\bar{x} = 1.681$ inches and $s \approx 0.0045$ inches.

The test statistic is

$$\begin{aligned}t_0 &= \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \\ &= \frac{1.681 - 1.68}{0.0045/\sqrt{12}} \\ &= 0.770 \quad \text{[Tech: 0.778]}\end{aligned}$$

<u>Classical approach</u>:
This is a two-tailed test with $n-1 = 12-1 = 11$ degrees of freedom, so the critical values are $\pm t_{0.025} = \pm 2.201$. Since $t_0 = 0.770$ is between -2.201 and 2.201, the test statistic does not fall in the critical region. Therefore, we do not reject the null hypothesis.

<u>*P*-value approach</u>:
This is a two-tailed test with 11 degrees of freedom. The *P*-value of this two tailed test is the area to the right of $t_0 = 0.770$, plus the area to the left of -0.770. From the *t*-distribution table (Table VI) in the row corresponding to 11 degrees of freedom, 0.770 falls between 0.697 and 0.876, whose right-tail areas are 0.25 and 0.20, respectively. We must double these values in order to get the total area in both tails: 0.50 and 0.40. Thus, $0.40 < P\text{-value} < 0.50$ [Tech: *P*-value = 0.4529].

<u>Conclusion</u>: There is not sufficient evidence to conclude that the mean diameter of Maxfli XS golf balls is different from 1.68 inches. In other words, there is not sufficient evidence to conclude that the golf balls do not conform. Therefore, we will give Maxfli the benefit of the doubt and assume the balls are conforming.

22. **(a)** The plotted data 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. Therefore, the conditions for testing the hypothesis are satisfied.

(b) Hypotheses: $H_0: \mu = 198$ words versus $H_1: \mu > 198$ words

We compute the sample mean and sample standard deviation to be $\bar{x} = 208.4$ words and $s \approx 9.3832$ words.

The test statistic is

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{208.4 - 198}{9.3832/\sqrt{10}} = 3.505$$

Classical approach:
This is a right-tailed test with $n-1 = 10-1 = 9$ degrees of freedom, so the critical value is $t_{0.1} = 1.383$. Since $t_0 = 3.505 > t_{0.1} = 1.383$, the test statistic is in the critical region. Therefore, we reject the null hypothesis.

P-value approach:
This is a right-tailed test with 9 degrees of freedom. The *P*-value is the area to the right of $t_0 = 3.505$. From the *t*-distribution table (Table VI) in the row corresponding to 9 degrees of freedom, 3.505 falls between 3.250 and 3.690, whose right-tail areas are 0.005 and 0.0025, respectively. So, $0.0025 < P\text{-value} < 0.005$ [Tech: 0.0033]. Since $P\text{-value} < \alpha = 0.1$, we reject the null hypothesis.

Conclusion: There is sufficient evidence to conclude that the mean number of words read per minute is now more than 198. In other words, there is sufficient evidence to conclude that the class was effective.

23. **(a)** The plotted data are all within the bounds of the normal probability plot, and the boxplot shows that there are no outliers. Therefore, the conditions for a hypothesis test are satisfied.

(b) Hypotheses: $H_0: \mu = 84.3$ seconds versus $H_1: \mu < 84.3$ seconds

We compute the sample mean and sample standard deviation to be $\bar{x} = 78$ seconds and $s = 15.21$ seconds.

The test statistic is

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{78 - 84.3}{15.21/\sqrt{10}} = -1.310$$

Classical approach:
This is a left-tailed test with $n-1 = 10-1 = 9$ degrees of freedom, so the critical value is $-t_{0.10} = -1.383$. Since $t_0 = -1.310 > -t_{0.10} = -1.383$, the test statistic is not in the critical region, so we do not reject the null hypothesis.

P-value approach:
From the *t*-distribution table (Table VI) in the row corresponding to 9 degrees of freedom, 1.310 falls between 1.100 and 1.383, whose right-tail areas are 0.15 and 0.1, respectively. So, $0.1 < P\text{-value} < 0.15$ [Tech: 0.1113]. Since $P\text{-value} > \alpha = 0.1$, we do not reject the null hypothesis.

Conclusion: There is not sufficient evidence to conclude that the new system results in a mean wait time that is less than 84.3 seconds. In other words, there is not sufficient evidence to conclude that the new system is effective.

24. **(a)** The plotted data are all within the bounds of the normal probability plot, and the boxplot shows that there are no outliers. Therefore, the conditions for a hypothesis test are satisfied.

(b) Hypotheses: $H_0: \mu = 7.0$ versus $H_1: \mu \neq 7.0$

We compute the sample mean and sample standard deviation to be $\bar{x} = 7.01$ and $s \approx 0.0316$.

The test statistic is

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{7.01 - 7.0}{0.0316/\sqrt{14}} = 1.184$$

[Tech: 1.183]

Classical approach:
This is a two-tailed test with $n-1 = 14-1 = 13$ degrees of freedom, so the critical values are $\pm t_{\alpha/2} = \pm t_{0.025} = \pm 2.160$. Since $t_0 = 1.184$ falls between $-t_{0.025} = -2.160$ and $t_{0.025} = 2.160$, the test statistic is not in the critical region, and we do not reject the null hypothesis.

P-value approach:
This is a two-tailed test with 13 degrees of freedom. The P-value of this two tailed test is the area to the right of $t_0 = 1.184$, plus the area to the left of -1.184. From the t-distribution table (Table VI) in the row corresponding to 13 degrees of freedom, 1.184 falls between 1.083 and 1.356, whose right-tail areas are 0.15 and 0.10, respectively. We must double these values in order to get the total area in both tails: 0.30 and 0.20. Thus, $0.20 < P\text{-value} < 0.30$ [Tech: P-value = 0.2579]. Since $P\text{-value} > \alpha = 0.05$, we do not reject the null hypothesis.

Conclusion: There is not enough evidence to conclude that the pH meter is incorrectly calibrated. In other words, the pH meter appears to be correctly calibrated.

25. (a) We assume that there is no difference between actual and predicted earnings, so the null hypothesis is $H_0: \mu = 0$. We want to gather evidence that shows that the predictions are off target, but we are not indicating whether the analysts, on average, had predictions that were too high or too low, so the alternative hypothesis is $H_1: \mu \neq 0$.

(b) Since $P\text{-value} = 0.046 < \alpha = 0.05$, we reject the statement in the null hypothesis. The evidence suggests that the analyst's earning predictions are not true.

(c) The researcher is supposed to select the level of significance and direction of the alternative hypothesis prior to gathering evidence. This removes the possibility of any researcher bias.

26. Hypotheses: $H_0: \mu = 0$ versus $H_1: \mu \neq 0$
The P-value is 0.44, which indicates that we should not reject the null hypothesis (we would expect to get the results we obtained, or more extreme, in about 44 samples out of 100 samples if the null hypothesis were true). Therefore, there is not sufficient evidence to conclude that the weather forecasts are inaccurate, on average.

27. The farmer's analysis may be correct, but his data come from only a small sample in a localized area. His farm is probably not representative of the entire country. In other words, the analysis conducted by the farmer only applies to his farm, not the entire United States.

28. (a) Answers will vary. We would expect 50 samples out of 1000 (i.e. 5%) to result in a rejection of the null hypothesis at $\alpha = 0.05$. The probability of a Type I error is $\alpha = 0.05$.

(b) Answers will vary. We would expect 50 samples out of 1000 (i.e. 5%) to result in a rejection of the null hypothesis at $\alpha = 0.05$. The probability of a Type I error is $\alpha = 0.05$.

(c) Answers will vary.

29. (a) Answers will vary. We would expect 50 samples out of 1000 (i.e. 5%) to result in a rejection of the null hypothesis at $\alpha = 0.05$. The probability of a Type I error is $\alpha = 0.05$. Discrepancies might occur if the requirements for hypothesis testing of the mean (e.g. normality) are not met.

(b) Answers will vary. We would expect 50 samples out of 1000 (i.e. 5%) to result in a rejection of the null hypothesis at $\alpha = 0.05$.

30. (a), (c) Answers will vary.

(b) We would expect 2 samples out of 40 (i.e. 5%) to result in a rejection of the null hypothesis at $\alpha = 0.05$. The probability of a Type I error is $\alpha = 0.05$.

(d) We established the parameters of the distribution at the outset of the simulation. For example, we fixed the population mean to be $\mu = 50$. We are then drawing samples from a population with known mean $\mu = 50$. Rejecting the null hypothesis $H_0: \mu = 50$ is incorrect because the mean is 50. Therefore, rejecting the null is a Type I error.

Consumer Reports®: Eyeglass Lenses

(a) Denoting the population mean haze difference as μ_{hdiff}, the hypotheses are:

$H_0 : \mu_{\text{hdiff}} = 0.6$ versus $H_1 : \mu_{\text{hdiff}} < 0.6$.

We compute the sample mean haze difference and sample standard deviation of haze difference to be $\bar{x}_{\text{hdiff}} \approx 0.4733$ and $s_{\text{hdiff}} \approx 0.1665$.

The test statistic is

$$t_0 = \frac{\bar{x}_{\text{hdiff}} - \mu_{\text{hdiff}}}{s_{\text{hdiff}} / \sqrt{n}} = \frac{0.4733 - 0.6}{0.1665/\sqrt{6}} = -1.864$$

Classical approach:
This is a left-tailed test with $n-1 = 6-1 = 5$ degrees of freedom. Using $\alpha = 0.10$, the critical value is $-t_{0.10} = -1.476$. Since $t_0 = -1.864 < -t_{0.10} = -1.476$, the test statistic is in the critical region, so we reject the null hypothesis.

P-value approach:
From the *t*-distribution table in the row corresponding to 5 degrees of freedom, 1.864 falls between 1.476 and 2.015, whose right-tail areas are 0.10 and 0.05, respectively. So, $0.05 < P\text{-value} < 0.10$ [Tech: 0.0607]. Since $P\text{-value} < \alpha = 0.10$, we reject the null hypothesis.

Conclusion: There is sufficient evidence, at the $\alpha = 0.10$ level of significance, to conclude that the mean haze difference for this manufacturer's lenses is below 0.6. In other words, there is significant evidence to conclude that this manufacturer's lenses are more scratch resistant than its closest competitor.

(b) Although it is difficult to verify that the manufacturer's lenses are "the most scratch-resistant plastic lenses ever made", based on our tumbling test, there is significant evidence to conclude that the manufacturer's lenses have better scratch resistance than the closest competitor.

Section 10.4

1. A hypothesis about a population proportion can be tested if the sample is obtained using simple random sampling and $np(1-p) \geq 10$ with $n \leq 0.05N$ (the sample size is no more than 5% of the population size).

2. Using the point estimate 55% and the margin of error 3% to create a confidence interval, we obtain: Lower Bound: $55\% - 3\% = 52\%$; Upper Bound: $55\% + 3\% = 58\%$. Because a majority would be any percentage greater than 50%, the headline is accurate.

3. $np_0(1-p_0) = 200 \cdot 0.3(1-0.3) = 42 \geq 10$, so the requirements of the hypothesis test are satisfied.

   (a) $\hat{p} = \frac{75}{200} = 0.375$

   The test statistic is

   $$z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.375 - 0.3}{\sqrt{\frac{0.3(1-0.3)}{200}}} \approx 2.31.$$

   This is a right-tailed test, so the critical value is $z_{0.05} = 1.645$. Since $z_0 = 2.31 > z_{0.05} = 1.645$, the test statistic is in the critical region, so we reject the null hypothesis.

   (b) $$\begin{aligned} P\text{-value} &= P(Z > 2.31) \\ &= 1 - P(Z \leq 2.31) \\ &= 1 - 0.9896 \\ &= 0.0104 \text{ [Tech: 0.0103]} \end{aligned}$$

   Since $P\text{-value} = 0.0104 < \alpha = 0.05$, we reject the null hypothesis.

4. $np_0(1-p_0) = 250 \cdot 0.6(1-0.6) = 60 \geq 10$, so the requirements of the hypothesis test are satisfied.

   (a) $\hat{p} = \frac{124}{250} = 0.496$

   The test statistic is

   $$z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.496 - 0.6}{\sqrt{\frac{0.6(1-0.6)}{250}}} \approx -3.36.$$

   This is a left-tailed test so the critical value is $-z_{0.01} = -2.33$. Since $z_0 = -3.36 < -z_{0.01} = -2.33$, the test statistic is in the critical region, so we reject the null hypothesis.

   (b) $P\text{-value} = P(Z < -3.36) = 0.0004$

   Since $P\text{-value} = 0.0004 < \alpha = 0.01$, we reject the null hypothesis.