CHAPTER 8

Section 8.1

1.

- **a.** Yes. It is an assertion about the value of a parameter.
- **b.** No. The sample median $\tilde{X}$ is not a parameter.
- **c.** No. The sample standard deviation s is not a parameter.
- **d.** Yes. The assertion is that the standard deviation of population #2 exceeds that of population #1
- **e.** No. $\overline{X}$ and $\overline{Y}$ are statistics rather than parameters, so cannot appear in a hypothesis.
- **f.** Yes. H is an assertion about the value of a parameter.

2.

- **a.** These hypotheses comply with our rules.
- **b.** H_o is not an equality claim (e.g. $s = 20$), so these hypotheses are not in compliance.
- **c.** H_o should contain the equality claim, whereas H_a does here, so these are not legitimate.
- **d.** The asserted value of $m_1 - m_2$ in H_o should also appear in H_a. It does not here, so our conditions are not met.
- **e.** Each S^2 is a statistic, so does not belong in a hypothesis.
- **f.** We are not allowing both H_o and H_a to be equality claims (though this is allowed in more comprehensive treatments of hypothesis testing).
- **g.** These hypotheses comply with our rules.
- **h.** These hypotheses are in compliance.

3. In this formulation, H_o states the welds do not conform to specification. This assertion will not be rejected unless there is strong evidence to the contrary. Thus the burden of proof is on those who wish to assert that the specification is satisfied. Using H_a: $m < 100$ results in the welds being believed in conformance unless provided otherwise, so the burden of proof is on the non-conformance claim.

4. When the alternative is H_a: $m < 5$, the formulation is such that the water is believed unsafe until proved otherwise. A type I error involved deciding that the water is safe (rejecting H_o) when it isn't (H_o is true). This is a very serious error, so a test which ensures that this error is highly unlikely is desirable. A type II error involves judging the water unsafe when it is actually safe. Though a serious error, this is less so than the type I error. It is generally desirable to formulate so that the type 1 error is more serious, so that the probability of this error can be explicitly controlled. Using H_a: $m > 5$, the type II error (now stating that the water is safe when it isn't) is the more serious of the two errors.

5. Let s denote the population standard deviation. The appropriate hypotheses are $H_o : s = .05$ vs $H_a : s < .05$. With this formulation, the burden of proof is on the data to show that the requirement has been met (the sheaths will not be used unless H_o can be rejected in favor of H_a. Type I error: Conclude that the standard deviation is < .05 mm when it is really equal to .05 mm. Type II error: Conclude that the standard deviation is .05 mm when it is really < .05.

6. $H_o : m = 40$ vs $H_a : m \neq 40$, where m is the true average burn-out amperage for this type of fuse. The alternative reflects the fact that a departure from $m = 40$ in either direction is of concern. Notice that in this formulation, it is initially believed that the value of m is the design value of 40.

7. A type I error here involves saying that the plant is not in compliance when in fact it is. A type II error occurs when we conclude that the plant is in compliance when in fact it isn't. Reasonable people may disagree as to which of the two errors is more serious. If in your judgement it is the type II error, then the reformulation $H_o : m = 150$ vs $H_a : m < 150$ makes the type I error more serious.

8. Let $m_1 =$ the average amount of warpage for the regular laminate, and $m_2 =$ the analogous value for the special laminate. Then the hypotheses are $H_o : m_1 = m_2$ vs $H_o : m_1 > m_2$. Type I error: Conclude that the special laminate produces less warpage than the regular, when it really does not. Type II error: Conclude that there is no difference in the two laminates when in reality, the special one produces less warpage.

9.

a. R_1 is most appropriate, because x either too large or too small contradicts p = .5 and supports $p \neq .5$.

b. A type I error consists of judging one of the tow candidates favored over the other when in fact there is a 50-50 split in the population. A type II error involves judging the split to be 50-50 when it is not.

c. X has a binomial distribution with n = 25 and p = 0.5. $\alpha =$ P(type I error) = $P(X \le 7 or X \ge 18$ when X ~ Bin(25, .5)) = B(7; 25,.5) + 1 – B(17; 25,.5) = .044

d. $\beta(.4) = P(8 \le X \le 17$ when p = .4) = B(17; 25,.5) – B(7, 25,.4) = 0.845, and $\beta(.6) = 0.845$ also. $\beta(.3) = B(17;25,.3) - B(7;25,.3) = .488 = \beta(.7)$

e. x = 6 is in the rejection region R_1, so H_o is rejected in favor of H_a.

10.

a. $H_o : \mu = 1300$ vs $H_a : \mu > 1300$

b. $\bar{x}$ is normally distributed with mean $E(\bar{x}) = \mu$ and standard deviation $\frac{\sigma}{\sqrt{n}} = \frac{60}{\sqrt{20}} = 13.416.$ When H_o is true, $E(\bar{x}) = 1300$. Thus $\alpha = P(\bar{x} \ge 1331.26$ when H_o is true) =

$$P\left(z \ge \frac{1331.26 - 1300}{13.416}\right) = P(z \ge 2.33) = .01$$

c. When $\mu = 1350$, $\bar{x}$ has a normal distribution with mean 1350 and standard deviation 13.416, so $\beta(1350) = P(\bar{x} < 1331.26$ when $\mu = 1350)$ =

$$P\left(z \le \frac{1331.26 - 1350}{13.416}\right) = P(z \le -1.40) = .0808$$

d. Replace 1331.26 by c, where c satisfies $\frac{c - 1300}{13.416} = 1.645$ (since $P(z \ge 1.645) = .05)$. Thus c = 1322.07. Increasing α gives a decrease in β; now $\beta(1350) = P(z \le -2.08) = .0188$.

e. $\bar{x} \ge 1331.26$ iff $z \ge \frac{1331.26 - 1300}{13.416}$ i.e. iff $z \ge 2.33$.

11.

a. $H_o: m = 10$ vs $H_a: m \neq 10$

b. $a = P($ rejecting H_o when H_o is true) = $P(\bar{x} \geq 10.1032$ or $\leq 9.8968 \text{ when } m = 10)$.
Since $\bar{x}$ is normally distributed with standard deviation
$\frac{s}{\sqrt{n}} = \frac{.2}{5} = .04,\ a = P(z \geq 2.58 \text{ or } \leq -2.58) = .005 + .005 = .01$

c. When $m = 10.1$, $E(\bar{x}) = 10.1$, so $b(10.1) = P(9.8968 < \bar{x} < 10.1032$ when $m = 10.1) = P(-5.08 < z < .08) = .5319$. Similarly,
$b(9.8) = P(2.42 < z < 7.58) = .0078$

d. $c = \pm 2.58$

e. Now $\frac{s}{\sqrt{n}} = \frac{.2}{3.162} = .0632$. Thus 10.1032 is replaced by c, where $\frac{c-10}{.0632} = 1.96$ and so c = 10.124. Similarly, 9.8968 is replaced by 9.876.

f. $\bar{x} = 10.020$. Since $\bar{x}$ is neither ≥ 10.124 nor ≤ 9.876, it is not in the rejection region. H_o is not rejected; it is still plausible that $m = 10$.

g. $\bar{x} \geq 10.1032$ or ≤ 9.8968 iff $z \geq 2.58$ or ≤ -2.58.

12.

a. Let $m =$ true average braking distance for the new design at 40 mph. The hypotheses are $H_o : m = 120$ vs $H_a : m < 120$.

b. R_2 should be used, since support for H_a is provided only by an $\bar{x}$ value substantially smaller than 120. ($E(\bar{x}) = 120$ when H_o is true and , 120 when H_a is true).

c. $s_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{10}{6} = 1.6667$, so $a = P(\bar{x} \geq 115.20$ when $m = 120) =$ $P\left(z \leq \frac{115.20 - 120}{1.6667}\right) = P(z \leq -2.88) = .002$. To obtain $a = .001$, replace 115.20 by $c = 120 - 3.08(1.6667) = 114.87$, so that $P(\bar{x} \leq 114.87$ when $m = 120) = P(z \leq -3.08) = .001$.

d. $b(115) = P(\bar{x} > 115.2$ when $m = 115) = P(z > .12) = .4522$

e. $a = P(z \leq -2.33) = .01$, because when H_o is true Z has a standard normal distribution ($\bar{X}$ has been standardized using 120). Similarly $P(z \leq -2.88) = .002$, so this second rejection region is equivalent to R_2.

13.

a. $P(\bar{x} \geq m_o + 2.33\frac{s}{\sqrt{n}}\, when m = m_o) = P\left(z \geq \frac{\left(m_o + 2.33\frac{s}{\sqrt{n}}\right)}{\frac{s}{\sqrt{n}}}\right)$

$= P(z \geq 2.33) = .01$, where Z is a standard normal r.v.

b. P(rejecting H_o when $m = 99) = P(\bar{x} \geq 102.33$ when $m = 99)$ $= P\left(z \geq \frac{102 - 99}{1}\right) = P(z \geq 3.33) = .0004$. Similarly, $a(98) = P(\bar{x} \geq 102.33$ when $m = 98) = P(z \geq 4.33) = 0$. In general, we have P(type I error) < .01 when this probability is calculated for a value of m less than 100. The boundary value $m = 100$ yields the largest a .

14.

a. $s_{\bar{x}} = .04$, so $P(\bar{x} \geq 10.1004 or \leq 9.8940$ when $m = 10) = P(z \geq 2.51 or \leq -2.65) = .006 + .004 = .01$

b. $b(10.1) = P(9.8940 < \bar{x} < 10.1004$ when $m = 10.1) = P(-5.15 < z < .01) = .5040$, whereas $b(9.9) = P(-.15 < z < 5.01) = .5596.$ Since $m = 9.9$ and $m = 10.1$ represent equally serious departures from H_o, one would probably want to use a test procedure for which $b(9.9) = b(10.1)$. A similar result and comment apply to any other pair of alternative values symmetrically placed about 10.

Section 8.2

15.

a. $a = P(z \geq 1.88$ when z has a standard normal distribution) $= 1 - \Phi(1.88) = .0301$

b. $a = P(z \leq -2.75$ when z ~ N(0, 1) $= \Phi(-2.75) = .003$

c. $a = \Phi(-2.88) + (1 - \Phi(2.88)) = .004$

16.

a. $a = P(t \geq 3.733$ when t has a t distribution with 15 d.f.) =.001, because the 15 d.f. row of Table A.5 shows that $t_{.001,15} = .3733$

b. d.f. = n – 1 = 23, so $a = P(t \leq -2.500) = .01$

c. d.f. = 30, and $a = P(t \geq 1.697) + P(t \leq -1.697) = .05 + .05 = .10$

17.

a. $z = \dfrac{20{,}960 - 20{,}000}{1500/\sqrt{16}} = 2.56 > 2.33$ so reject H_o.

b. $b(20{,}500) : \Phi\left(2.33 + \dfrac{20{,}000 - 20{,}500}{1500/\sqrt{16}}\right) = \Phi(1.00) = .8413$

c. $b(20{,}500) = .05 : n = \left[\dfrac{1500(2.33 + 1.645)}{20{,}000 - 20{,}500}\right]^2 = 142.2$, so use n = 143

d. $a = 1 - \Phi(2.56) = .0052$

18.

a. $\dfrac{72.3 - 75}{1.8} = -1.5$ so 72.3 is 1.5 SD's (of $\overline{x}$) below 75.

b. H_o is rejected if $z \le -2.33$; since $z = -1.5$ is not ≤ -2.33, don't reject H_o.

c. $a =$ area under standard normal curve below –2.88 $= \Phi(-2.88) = .0020$

d. $\Phi\left(-2.88 + \dfrac{75 - 70}{9/5}\right) = \Phi(-.1) = .4602$ so $b(70) = .5398$

e. $n = \left[\dfrac{9(2.88 + 2.33)}{75 - 70}\right]^2 = 87.95$, so use n = 88

f. $a(76) = P(Z < -2.33$ when $m = 76) = P(\overline{X} < 72.9$ when $m = 76)$

$= \Phi\left(\dfrac{72.9 - 76}{.9}\right) = \Phi(-3.44) = .0003$

19.

a. Reject H_o if either $z \ge 2.58$ or $z \le -2.58$; $\frac{s}{\sqrt{n}} = 0.3$, so

$z = \frac{94.32 - 95}{0.3} = -2.27$. Since –2.27 is not < -2.58, don't reject H_o.

b. $b(94) = \Phi\left(2.58 - \frac{1}{0.3}\right) - \Phi\left(-2.58 - \frac{1}{0.3}\right) = \Phi(-.75) - \Phi(-5.91) = .2266$

c. $n = \left[\frac{1.20(2.58 + 1.28)}{95 - 94}\right]^2 = 21.46$, so use n = 22.

20. With H_o: $m = 750$, and H_a: $m < 750$ and a significance level of .05, we reject H_o if z < -1.645; z = -2.14 < -1.645, so we reject the null hypothesis and do not continue with the purchase. At a significance level of .01, we reject H_o if z < -2.33; z = -2.14 > -2.33, so we don't reject the null hypothesis and thus continue with the purchase.

21. With H_o: $m = .5$, and H_a: $m \ne .5$ we reject H_o if $t > t_{a/2,n-1}$ or $t < -t_{a/2,n-1}$

a. $1.6 < t_{.025,12} = 2.179$, so don't reject H_o

b. $-1.6 > -t_{.025,12} = -2.179$, so don't reject H_o

c. $-2.6 > -t_{.005,24} = -2.797$, so don't reject H_o

d. –3.9 < the negative of all t values in the df = 24 row, so we reject H_o in favor of H_a.

22.

a. It appears that the true average weight could be more than the production specification of 200 lb per pipe.

b. H_o: $m = 200$, and H_a: $m > 200$ we reject H_o if $t > t_{.05,29} = 1.699$.

$t = \frac{206.73 - 200}{6.35/\sqrt{30}} = \frac{6.73}{1.16} = 5.80 > 1.699$, so reject H_o. The test appears to substantiate the statement in part **a**.

23. H_o: $m = 360$ vs. H_a: $m > 360$; $t = \frac{\bar{x} - 360}{s/\sqrt{n}}$; reject H_o if $t > t_{.05,25} = 1.708$;

$t = \frac{370.69 - 360}{24.36/\sqrt{26}} = 2.24 > 1.708$. Thus H_o should be rejected. There appears to be a contradiction of the prior belief.

24. H$_o$: $m = 3000$ vs. H$_a$: $m \neq 3000$; $t = \dfrac{\bar{x} - 3000}{s / \sqrt{n}}$; reject H$_o$ if $|t| > t_{.025,4} = 2.776$;
$t = \dfrac{2887.6 - 3000}{84 / \sqrt{5}} = -2.99 < -2.776$, so we reject H$_o$. This requirement is not satisfied.

25.

a. H$_o$: $m = 5.5$ vs. H$_a$: $m \neq 5.5$; for a level .01 test, (not specified in the problem description), reject H$_o$ if either $z \geq 2.58$ or $z \leq -2.58$. Since $z = \dfrac{5.25 - 5.5}{.075} = -3.33 \leq -2.58$, reject H$_o$.

b. $$1 - b(5.6) = 1 - \Phi\left(2.58 + \frac{(-.1)}{.075}\right) + \Phi\left(-2.58 - \frac{(-.1)}{.075}\right)$$
$$= 1 - \Phi(1.25) + \Phi(-3.91) = .105$$

c. $n = \left[\dfrac{.3(2.58 + 2.33)}{-.1}\right]^2 = 216.97$, so use n = 217.

26. Reject H$_o$ if $z \geq 1.645$; $\dfrac{s}{\sqrt{n}} = .7155,$ so $z = \dfrac{52.7 - 50}{.7155} = 3.77$. Since 3.77 is ≥ 1.645 , reject H$_o$ at level .05 and conclude that true average penetration exceeds 50 mils.

27. We wish to test H$_o$: $m = 75$ vs. H$_a$: $m < 75$; Using $a = .01$, H$_o$ is rejected if $t \leq -t_{.01,41} \approx -2.423$ (from the df 40 row of the t-table). Since $t = \dfrac{73.1 - 75}{5.9 / \sqrt{42}} = -2.09$, which is not ≤ -2.423, H$_o$ is not rejected. The alloy is not suitable.

28. With $m =$ true average recumbency time, the hypotheses are H$_o$: $m = 20$ vs H$_a$: $m < 20$. The test statistic value is $z = \dfrac{\bar{x} - 20}{s / \sqrt{n}}$, and H$_o$ should be rejected if $z \leq -z_{.10} = -1.28$
Since $z = \dfrac{18.86 - 20}{8.6 / \sqrt{73}} = -1.13$, which is not ≤ -1.28 , H$_o$ is not rejected. The sample data does not strongly suggest that true average time is less than 20.

29.

a. For n = 8, n – 1 = 7, and $t_{.05,7} = 1.895$, so H_o is rejected at level .05 if $t \geq 1.895$.

Since $\frac{s}{\sqrt{n}} = \frac{1.25}{\sqrt{8}} = .442$, $t = \frac{3.72 - 3.50}{.442} = .498$; this does not exceed 1.895, so H_o is not rejected.

b. $d = \frac{|m_o - m|}{s} = \frac{|3.50 - 4.00|}{1.25} = .40$, and n = 8, so from table A.17, $b(4.0) \approx .72$

30. n = 115, $\bar{x} = 11.3$, $s = 6.43$

1 Parameter of Interest: $m =$ true average dietary intake of zinc among males aged 65 – 74 years.

2 Null Hypothesis: H_o: $m = 15$

3 Alternative Hypothesis: H_a: $m < 15$

4 $z = \frac{\bar{x} - m_o}{s/\sqrt{n}} = \frac{\bar{x} - 15}{s/\sqrt{n}}$

5 Rejection Region: No value of α was given, so select a reasonable level of significance, such as α= .05. $z \leq z_a$ or $z \leq -1.645$

6 $z = \frac{11.3 - m_o}{6.43/\sqrt{115}} = -6.17$

7 –6.17 < -1.645, so reject H_o. The data does support the claim that average daily intake of zinc for males aged 65 - 74 years falls below the recommended daily allowance of 15 mg/day.

31. The hypotheses of interest are H_o: $m = 7$ vs H_a: $m < 7$, so a lower-tailed test is appropriate; H_o should be rejected if $t \leq -t_{.1,8} = -1.397$. $t = \frac{6.32 - 7}{1.65/\sqrt{9}} = -1.24$. Because -1.24 is not ≤ -1.397, H_o (prior belief) is not rejected (contradicted) at level .01.

32. n = 12, $\bar{x} = 98.375$, $s = 6.1095$

a.

1 Parameter of Interest: $m =$ true average reading of this type of radon detector when exposed to 100 pCi/L of radon.

2 Null Hypothesis: H_o: $m = 100$

3 Alternative Hypothesis: H_a: $m \neq 100$

4 $t = \dfrac{\bar{x} - m_o}{s/\sqrt{n}} = \dfrac{\bar{x} - 100}{s/\sqrt{n}}$

5 $t \le -2.201$ or $t \ge 2.201$

6 $t = \dfrac{98.375 - 100}{6.1095/\sqrt{12}} = -.9213$

7 Fail to reject H_o. The data does not indicate that these readings differ significantly from 100.

b. $\sigma = 7.5$, $\beta = 0.10$. From table A.17, df ≈ 29, thus n ≈30.

33. $b(m_o - \Delta) = \Phi\left(z_{a/2} + \Delta\sqrt{n}/s\right) - \Phi\left(-z_{a/2} - \Delta\sqrt{n}/s\right)$

$= 1 - \left[\Phi\left(-z_{a/2} - \Delta\sqrt{n}/s\right) + \Phi\left(z_{a/2} - \Delta\sqrt{n}/s\right)\right] = b(m_o + \Delta)$

(since 1 - Φ(c) = Φ(-c)).

34. For an upper-tailed test, $= b(m) = \Phi\left(z_a + \sqrt{n}(m_o - m)/s\right)$. Since in this case we are considering $m > m_o$, $m_o - m$ is negative so $\sqrt{n}(m_o - m)/s \to -\infty$ as n $\to \infty$. The desired conclusion follows since $\Phi(-\infty) = 0$. The arguments for a lower-tailed and tow-tailed test are similar.

Section 8.3

35.

1 Parameter of interest: p = true proportion of cars in this particular county passing emissions testing on the first try.

2 H_o: p = .70

3 H_a: p ≠ .70

4 $z = \dfrac{\hat{p} - p_o}{\sqrt{p_o(1 - p_o)/n}} = \dfrac{\hat{p} - .70}{\sqrt{.70(.30)/n}}$

5 either z ≥ 1.96 or z ≤ -1.96

6 $z = \dfrac{124/200 - .70}{\sqrt{.70(.30)/200}} = -2.469$

7 Reject H_o. The data indicates that the proportion of cars passing the first time on emission testing or this county differs from the proportion of cars passing statewide.

36.

a.

1 p = true proportion of all nickel plates that blister under the given circumstances.

2 H_o: p = .10

3 H_a: p > .10

4 $$z = \frac{\hat{p} - p_o}{\sqrt{p_o(1-p_o)/n}} = \frac{\hat{p} - .10}{\sqrt{.10(.90)/n}}$$

5 Reject H_o if $z \geq 1.645$

6 $$z = \frac{14/100 - .10}{\sqrt{.10(.90)/100}} = 1.33$$

7 Fail to Reject H_o. The data does not give compelling evidence for concluding that more than 10% of all plates blister under the circumstances.

The possible error we could have made is a Type II error: Failing to reject the null hypothesis when it is actually true.

b. $b(.15) = \Phi\left[\frac{.10 - .15 + 1.645\sqrt{.10(.90)/100}}{\sqrt{.15(.85)/100}}\right] = \Phi(-.02) = .4920$. When n = 200, $b(.15) = \Phi\left[\frac{.10 - .15 + 1.645\sqrt{.10(.90)/200}}{\sqrt{.15(.85)/200}}\right] = \Phi(-.60) = .2743$

c. $n = \left[\frac{1.645\sqrt{.10(.90)} + 1.28\sqrt{.15(.85)}}{.15 - .10}\right]^2 = 19.01^2 = 361.4$, so use n = 362

37.

1 p = true proportion of the population with type A blood

2 H_o: p = .40

3 H_a: $p \neq .40$

4 $$z = \frac{\hat{p} - p_o}{\sqrt{p_o(1-p_o)/n}} = \frac{\hat{p} - .40}{\sqrt{.40(.60)/n}}$$

5 Reject H_o if $z \geq 2.58$ or $z \leq -2.58$

6 $$z = \frac{82/150 - .40}{\sqrt{.40(.60)/150}} = \frac{.147}{.04} = 3.667$$

7 Reject H_o. The data does suggest that the percentage of the population with type A blood differs from 40%. (at the .01 significance level). Since the z critical value for a significance level of .05 is less than that of .01, the conclusion would not change.

38.

a. We wish to test H_o: p = .02 vs H_a: p < .02; only if H_o can be rejected will the inventory be postponed. The lower-tailed test rejects H_o if $z \le -1.645$. With $\hat{p} = \frac{15}{1000} = .015$, z = -1.01, which is not $\le$ -1.645. Thus, H_o cannot be rejected, so the inventory should be carried out.

b. $b(.01) = \Phi\left[\frac{.02 - .01 + 1.645\sqrt{.02(.98)/1000}}{\sqrt{.01(.99)/1000}}\right] = \Phi(5.49) \approx 1$

c. $b(.05) = \Phi\left[\frac{.02 - .05 + 1.645\sqrt{.02(.98)/1000}}{\sqrt{.05(.95)/1000}}\right] = \Phi(-3.30) = .0005$, so is p = .05 it is highly unlikely that H_o will be rejected and the inventory will almost surely be carried out.

39. Let p denote the true proportion of those called to appear for service who are black. We wish to test H_o: p = .25 vs H_a: p < .25. We use $z = \frac{\hat{p} - .25}{\sqrt{.25(.75)/n}}$, with the rejection region $z \le -z_{.01} = -2.33$. We calculate $\hat{p} = \frac{177}{1050} = .1686$, and $z = \frac{.1686 - .25}{.0134} = -6.1$. Because $-6.1 < -2.33$, H_o is rejected. A conclusion that discrimination exists is very compelling.

40.

a. P = true proportion of current customers who qualify. H_o: p = .05 vs H_a: p $\ne$ .05, $z = \frac{\hat{p} - .05}{\sqrt{.05(.95)/n}}$, reject H_o if $z \ge 2.58$ or $z \le -2.58$. $\hat{p} = .08$, so $z = \frac{.03}{.00975} = 3.07 \ge 2.58$, so H_o is rejected. The company's premise is not correct.

b. $b(.10) = \Phi\left[\frac{.05 - .10 + 2.58\sqrt{.05(.95)/500}}{\sqrt{.10(.90)/500}}\right] = \Phi(-1.85) = .0332$

41.

a. The alternative of interest here is H_a: p > .50 (which states that more than 50% of all enthusiasts prefer gut), so the rejection region should consist of large values of X (an upper-tailed test). Thus (15, 16, 17, 18, 19, 20) is the appropriate region.

b. $a = P(15 \le X \text{ when } p = .5)$ = 1 – B(14; 20, .05) = .021, so this is a level .05 test. For R = {14, 15, …, 20}, α = .058, so this R does not specify a level .05 test and the region of **a** is the best level .05 test. (α ≤ .05 along with smallest possible β).

c. β(.6) = B(14; 20, .6) = .874, and β(.8) = B(14; 20, .8) = .196.

d. The best level .10 test is specified by R = (14, …, 20} (with α = .052) Since 13 is not in R, H_o is not rejected at this level.

42. The hypotheses are H_o: p = .10 vs. H_a: p > .10, so R has the form {c, …, n}. For n = 10, c = 3 (i.e. R = {3, 4, …, 10}) yields α = 1 – B(2; 10, .1) = .07 while no larger R has α ≤ .10; however β(.3) = B(2; 10, .3) = .383. For n = 20, c = 5 yields α = 1 – B(4; 20, .1) = .043, but again β(.3) = B(4; 20, .3) = .238. For n = 25, c = 5 yields α = 1 – B(4; 25, .1) = .098 while β(.7) = B(4; 25, .3) = .090 ≤ .10, so n = 25 should be used.

43. H_o: p = .035 vs H_a: p < .035. We use $z = \frac{\hat{p} - .035}{\sqrt{.035(.965)/n}}$, with the rejection region z ≤ -$z_{.01}$ = -2.33. With $\hat{p} = \frac{15}{500} = .03$, $z = \frac{-.005}{\sqrt{.0082}} = -.61$. Because -.61 isn't ≤ -2.33, H_o is not rejected. Robots have not demonstrated their superiority.

Section 8.4

44. Using α = .05, H_o should be rejected whenever p-value < .05.

a. P-value = .001 < .05, so reject H_o

b. .021 < .05, so reject H_o.

c. .078 is not < .05, so don't reject H_o.

d. .047 < .05, so reject H_o (a close call).

e. .148 > .05, so H_o can't be rejected at level .05.

45.

a. p-value = .084 > .05 = α, so don't reject H_o.

b. p-value = .003 < .001 = α, so reject H_o.

c. .498 >> .05, so H_o can't be rejected at level .05

d. 084 < .10, so reject H_o at level .10

e. .039 is not < .01, so don't reject H_o.

f. p-value = .218 > .10, so H_o cannot be rejected.

46. In each case the p-value = $1-\Phi(z)$

a. .0778

b. .1841

c. .0250

d. .0066

e. .4562

47.

a. .0358

b. .0802

c. .5824

d. .1586

e. 0

48.

a. In the df = 8 row of table A.5, t = 2.0 is between 1.860 and 2.306, so the p-value is between .025 and .05: .025 < p-value < .05.

b. 2.201 < | -2.4 | < 2.718, so .01 < p-value < .025.

c. 1.341 < | -1.6 | < 1.753, so .05 < P(t < -1.6) < .10. Thus a two-tailed p-value: 2(.05 < P(t < -1.6) < .10), or .10 < p-value < .20

d. With an upper-tailed test and t = -.4, the p-value = P(t > -.4) > .50.

e. 4.032 < t=5 < 5.893, so .001 < p-value < .005

f. 3.551 < | -4.8 |, so P(t < -4.8) < .0005. A two-tailed p-value = 2[P(t < -4.8)] < 2(.0005), or p-value < .001.

49. An upper-tailed test

a. Df = 14, α=.05; $t_{.05,14} = 1.761$: 3.2 > 1.761, so reject H_o.

b. $t_{.01,18} = 2.896$; 1.8 is not > 2.896, so don't reject H_o.

c. Df = 23, p-value > .50, so fail to reject H_o at any significance level.

50. The p-value is greater than the level of significance α = .01, therefore fail to reject H_o that $m = 5.63$. The data does not indicate a difference in average serum receptor concentration between pregnant women and all other women.

51. Here we might be concerned with departures above as well as below the specified weight of 5.0, so the relevant hypotheses are H_o: $m = 5.0$ vs H_a: $m \neq 5.0$. At level .01, reject H_o if either $z \geq 2.58$ or $z \leq -2.58$. Since $\frac{s}{\sqrt{n}} = .035$, $z = \frac{-.13}{.035} = -3.71$, which is ≤ -2.58 , so H_o should be rejected. Because 3.71 is "off" the z-table, p-value < 2(.0002) = .0004 (.0002 corresponds to z = -3.49).

52.

a. For testing H_o: p = .2 vs H_a: p > .2, an upper-tailed test is appropriate. The computed Z is z = .97, so p-value = $1 - \Phi(.97) = .166$. Because the p-value is rather large, H_o would not be rejected at any reasonable α (it can't be rejected for any α < .166), so no modification appears necessary.

b. With p = .5, $1 - b(.5) = 1 - \Phi[(-.3 + 2.33(.0516))/.0645] = 1 - \Phi(-2.79) = .9974$

53. p = proportion of all physicians that know the generic name for methadone.
H_o: p = .50 vs H_a: p < .50; We can use a large sample test if both $np_0 \geq 10$ and $n(1 - p_0) \geq 10$; 102(.50) = .51, so we can proceed. $\hat{p} = \frac{47}{102}$, so

$$z = \frac{\frac{47}{102} - .50}{\sqrt{\frac{(.50)(.50)}{102}}} = \frac{-.039}{.050} = -.79$$

We will reject H_0 if the p-value < .01. For this lower tailed test, the p-value = Φ(z) = Φ(-.79) =.2148, which is not < .01, so we do not reject H_0 at significance level .01.

54. $m =$ the true average percentage of organic matter in this type of soil, and the hypotheses are H_o: $m = 3$ vs H_a: $m \neq 3$. With n = 30, and assuming normality, we use the t test:

$t = \dfrac{\bar{x}-3}{s/\sqrt{n}} = \dfrac{2.481-3}{.295} = \dfrac{-.519}{.295} = -1.759$. The p-value = 2[P(t > 1.759)] = 2(.041) = .082. At significance level .10, since .082 = .10, we would reject H_0 and conclude that the true average percentage of organic matter in this type of soil is something other than 3. At significance level .05, we would not have rejected H_0.

55. The hypotheses to be tested are H_o: $m = 25$ vs H_a: $m > 25$, and H_o should be rejected if $t \geq t_{.05,12} = 1.782$. The computed summary statistics are $\bar{x} = 27.923$, $s = 5.619$, so $\dfrac{s}{\sqrt{n}} = 1.559$ and $t = \dfrac{2.923}{1.559} = 1.88$. From table A.8, P(t > 1.88) ˜ .041, which is less than .05, so H_o is rejected at level .05.

56.

a. The appropriate hypotheses are H_o: $m = 10$ vs H_a: $m < 10$

b. P-value = P(t > 2.3) = .017, which is = .05, so we would reject H_o. The data indicates that the pens do not meet the design specifications.

c. P-value = P(t > 1.8) = .045, which is not = .01, so we would not reject H_o. There is not enough evidence to say that the pens don't satisfy the design specifications.

d. P-value = P(t > 3.6) ˜ .001, which gives strong evidence to support the alternative hypothesis.

57. $m =$ true average reading, H_o: $m = 70$ vs H_a: $m \neq 70$, and

$t = \dfrac{\bar{x}-70}{s/\sqrt{n}} = \dfrac{75.5-70}{7/\sqrt{6}} = \dfrac{5.5}{2.86} = 1.92$. From table A.8, df = 5, p-value = 2[P(t> 1.92)] ˜ 2(.058) = .116. At significance level .05, there is not enough evidence to conclude that the spectrophotometer needs recalibrating.

58. With H_o: $m = .60$ vs H_a: $m \neq .60$, and a two-tailed p-value of .0711, we fail to reject H_o at levels .01 and .05 (thus concluding that the amount of impurities need not be adjusted) , but we would reject H_o at level .10 (and conclude that the amount of impurities does need adjusting).

Section 8.5

59.

a. The formula for b is $1-\Phi\left(-2.33+\frac{\sqrt{n}}{9.4}\right)$, which gives .8980 for n = 100, .1049 for n = 900, and .0014 for n = 2500.

b. Z = -5.3, which is "off the z table," so p-value < .0002; this value of z is quite statistically significant.

c. No. Even when the departure from H_o is insignificant from a practical point of view, a statistically significant result is highly likely to appear; the test is too likely to detect small departures from H_o.

60.

a. Here $b = \Phi\left(\frac{-.01+.9320/\sqrt{n}}{.4073/\sqrt{n}}\right) = \Phi\left(\frac{\left(-.01\sqrt{n}+.9320\right)}{.4073}\right)$ = .9793, .8554, .4325, .0944, and 0 for n = 100, 2500, 10,000, 40,000, and 90,000, respectively.

b. Here $z = .025\sqrt{n}$ which equals .25, 1.25, 2.5, and 5 for the four n's, whence p-value = .4213, .1056, .0062, .0000, respectively.

c. No; the reasoning is the same as in 54 (c).

Supplementary Exercises

61. Because n = 50 is large, we use a z test here, rejecting H_o: $m = 3.2$ in favor of H_a: $m \neq 3.2$ if either $z \geq z_{.025} = 1.96$ or $z \leq -1.96$. The computed z value is $z = \frac{3.05-3.20}{.34/\sqrt{50}} = -3.12$. Since –3.12 is ≤ -1.96, H_o should be rejected in favor of H_a.

62. Here we assume that thickness is normally distributed, so that for any *n* a t test is appropriate, and use Table A.17 to determine n. We wish $p(3) = .95$ when $d = \frac{|3.2-3|}{.3} = .667$. By inspection, n = 20 satisfies this requirement, so n = 50 is too large.

63.

a. H$_o$: $m = 3.2$ vs H$_a$: $m \neq 3.2$ (Because H$_a$: $m > 3.2$ gives a p-value of roughly .15)

b. With a p-value of .30, we would reject the null hypothesis at any reasonable significance level, which includes both .05 and .10.

64.

a. H$_o$: $m = 2150$ vs H$_a$: $m > 2150$

b. $t = \dfrac{\bar{x} - 2150}{s/\sqrt{n}}$

c. $t = \dfrac{2160 - 2150}{30/\sqrt{16}} = \dfrac{10}{7.5} = 1.33$

d. Since $t_{.10,15} = 1.341$, p-value > .10 (actually $\approx .10$)

e. From **d**, p-value > .05, so H$_o$ cannot be rejected at this significance level.

65.

a. The relevant hypotheses are H$_o$: $m = 548$ vs H$_a$: $m \neq 548$. At level .05, H$_o$ will be rejected if either $t \geq t_{.025,10} = 2.228$ or $t \leq -t_{.025,10} = -2.228$. The test statistic value is $t = \dfrac{587 - 548}{10/\sqrt{11}} = \dfrac{39}{3.02} = 12.9$. This clearly falls into the upper tail of the two-tailed rejection region, so H$_o$ should be rejected at level .05, or any other reasonable level).

b. The population sampled was normal or approximately normal.

66. $n = 8, \bar{x} = 30.7875, s = 6.5300$

1 Parameter of interest: $m =$ true average heat-flux of plots covered with coal dust

2 H$_o$: $m = 29.0$

3 H$_a$: $m > 29.0$

4 $t = \dfrac{\bar{x} - 29.0}{s/\sqrt{n}}$

5 RR: $t \geq t_{a,n-1}$ or $t \geq 1.895$

6 $t = \dfrac{30.7875 - 29.0}{6.53/\sqrt{8}} = .7742$

7 Fail to reject H$_o$. The data does not indicate the mean heat-flux for pots covered with coal dust is greater than for plots covered with grass.

67. N = 47, $\bar{x} = 215$ mg, s = 235 mg. Range 5 mg to 1,176 mg.

a. No, the distribution does not appear to be normal, it appears to be skewed to the right. It is not necessary to assume normality if the sample size is large enough due to the central limit theorem. This sample size is large enough so we can conduct a hypothesis test about the mean.

b.

1 Parameter of interest: $m =$ true daily caffeine consumption of adult women.

2 H_o: $m = 200$

3 H_a: $m > 200$

4 $z = \dfrac{\bar{x} - 200}{s / \sqrt{n}}$

5 RR: $z \geq 1.282$ or if p-value $\leq .10$

6 $z = \dfrac{215 - 200}{235 / \sqrt{47}} = .44$; p-value $= 1 - \Phi(.44) = .33$

7 Fail to reject H_o. because .33 > .10. The data does not indicate that daily consump tion of all adult women exceeds 200 mg.

68. At the .05 significance level, reject H_o because .043 < .05. At the level .01, fail to reject H_o because .043 > .01. Thus the data contradicts the design specification that sprinkler activation is less than 25 seconds at the level .05, but not at the .01 level.

69.

a. From table A.17, when $m = 9.5$, d = .625, df = 9, and $b \approx .60$, when $m = 9.0$, d = 1.25, df = 9, and $b \approx .20$.

b. From Table A.17, $b = .25$, d = .625, n ≈ 28

70. A normality plot reveals that these observations could have come from a normally distributed population, therefore a t-test is appropriate. The relevant hypotheses are H_o: $m = 9.75$ vs H_a: $m > 9.75$. Summary statistics are n = 20, $\bar{x} = 9.8525$, and s = .0965, which leads to a test statistic $t = \dfrac{9.8525 - 9.75}{.0965 / \sqrt{20}} = 4.75$, from which the p-value = .0001. (From MINITAB output). With such a small p-value, the data strongly supports the alternative hypothesis. The condition is not met.

71.

a. With H_o: p = $\frac{1}{75}$ vs H_o: p ≠ $\frac{1}{75}$, we reject H_o if either $z \ge 1.96$ or $z \le -1.96$.

With $\hat{p} = \frac{16}{800} = .02$, $z \frac{.02 - .01333}{\sqrt{\frac{.01333(.98667)}{800}}} = 1.645$, which is not in either rejection region. Thus, we fail to reject the null hypothesis. There is not evidence that the incidence rate among prisoners differs from that of the adult population. The possible error we could have made is a type II.

b. P-value = $2[1-\Phi(1.645)] = 2[.05] = .10$. Yes, since .10 < .20, we could reject H_o.

72. A t test is appropriate; H_o: $m = 1.75$ is rejected in favor of H_a: $m \ne 1.75$ if the p-value >.05. The computed t is $t = \frac{1.89 - 1.75}{.42/\sqrt{26}} = 1.70$. Since $1.70 \approx 1.708 = t_{.025,25}$, $P \approx 2(.05) = .10$ (since for a two-tailed test, $.05 = a/2$.), do not reject H_o; the data does not contradict prior research. We assume that the population from which the sample was taken was approximately normally distributed.

73. Even though the underlying distribution may not be normal, a z test can be used because n is large. H_o: $m = 3200$ should be rejected in favor of H_a: $m < 3200$ if $z \le -z_{.001} = -3.08$. The computed z is $z = \frac{3107 - 3200}{188/\sqrt{45}} = -3.32 \le -3.08$, so H_o should be rejected at level .001.

74. Let p = the true proportion of mechanics who could identify the problem. Then the appropriate hypotheses are H_o: p = .75 vs H_a: p < .75, so a lower-tailed test should be used. With p_o= .75 and $\hat{p} = \frac{42}{72} = .583$, z = -3.28 and $P = \Phi(-3.28) = .0005$. Because this p-value is so small, the data argues strongly against H_o, so we reject it in favor of H_a.

75. We wish to test H_o: $l = 4$ vs H_a: $l > 4$ using the test statistic $z = \frac{\bar{x} - 4}{\sqrt{4/n}}$. For the given sample, n = 36 and $\bar{x} = \frac{160}{36} = 4.444$, so $z = \frac{4.444 - 4}{\sqrt{4/36}} = 1.33$. At level .02, we reject H_o if $z \ge z_{.02} \approx 2.05$ (since $1 - \Phi(2.05) = .0202$). Because 1.33 is not ≥ 2.05, H_o should not be rejected at this level.

76. H$_o$: $m = 15$ vs H$_a$: $m > 15$. Because the sample size is less than 40, and we can assume the distribution is approximately normal, the appropriate statistic is $t = \frac{\bar{x} - 15}{s/\sqrt{n}} = \frac{17.5 - 15}{2.2/\sqrt{32}} = \frac{2.5}{.390} = 6.4$. Thus the p-value is "off the chart" in the 20 df column of Table A.8, and so is approximately $0 < .05$, so H$_o$ is rejected in favor of the conclusion that the true average time exceeds 15 minutes.

77. H$_o$: $s^2 = .25$ vs H$_a$: $s^2 > .25$. The chi-squared critical value for 9 d.f. that captures upper-tail area .01 is 21.665. The test statistic value is $\frac{9(.58)^2}{.25} = 12.11$. Because 12.11 is not ≥ 21.665, H$_o$ cannot be rejected. The uniformity specification is not contradicted.

78. The 20 df row of Table A.7 shows that $c^2_{.99,20} = 8.26 < 8.58$ (H$_o$ not rejected at level .01) and $8.58 < 9.591 = c^2_{.975,20}$ (H$_o$ rejected at level .025). Thus .01 < p-value < .025 and H$_o$ cannot be rejected at level .01 (the p-value is the smallest alpha at which rejection can take place, and this exceeds .01).

79.

a. $E\left(\overline{X} + 2.33S\right) = E(\overline{X}) + 2.33E(S) = m + 2.33s$, so $\hat{q} = \overline{X} + 2.33S$ is approximately unbiased.

b. $V\left(\overline{X} + 2.33S\right) = V(\overline{X}) + 2.33^2 V(S) = \frac{s^2}{n} + 5.4289\frac{s^2}{2n}$. The estimated standard error (standard deviation) is $1.927\frac{s}{\sqrt{n}}$.

c. More than 99% of all soil samples have pH less than 6.75 iff the 95th percentile is less than 6.75. Thus we wish to test H$_o$: $m + 2.33s = 6.75$ vs H$_a$: $m + 2.33s < 6.75$. H$_o$ will be rejected at level .01 if $z \leq 2.33$. Since $z = \frac{-.047}{.0385} < 0$, H$_o$ clearly cannot be rejected. The 95th percentile does not appear to exceed 6.75.

80.

a. When H_o is true, $2l_o\Sigma X_i = 2\sum \frac{X_i}{m_o}$ has a chi-squared distribution with df = 2n. If the alternative is H_a: $m > m_o$, large test statistic values (large Σx_i, since $\overline{x}$ is large) suggest that H_o be rejected in favor of H_a, so rejecting when $2\sum \frac{X_i}{m_o} \geq c^2_{a,2n}$ gives a test with significance level a. If the alternative is H_a: $m < m_o$, rejecting when $2\sum \frac{X_i}{m_o} \leq c^2_{1-a,2n}$ gives a level a test. The rejection region for H_a: $m \neq m_o$ is either $2\sum \frac{X_i}{m_o} \geq c^2_{a/2,2n}$ or $\leq c^2_{1-a/2,2n}$.

b. H_o: $m = 75$ vs H_a: $m < 75$. The test statistic value is $\frac{2(737)}{75} = 19.65$. At level .01, H_o is rejected if $2\sum \frac{X_i}{m_o} \leq c^2_{.99,20} = 8.260$. Clearly 19.65 is not in the rejection region, so H_o should not be rejected. The sample data does not suggest that true average lifetime is less than the previously claimed value.

81.

a. P(type I error) = P(either $Z \geq z_g$ or $Z \leq z_{a-g}$) (when Z is a standard normal r.v.) = $\Phi(-z_{a-g}) + 1 - \Phi(z_g) = a - g + g = a$.

b. $b(m) = P(\overline{X} \geq m_o + \frac{sz_g}{\sqrt{n}} \; or \overline{X} \leq m_o - \frac{sz_{a-g}}{\sqrt{n}}$ when the true value is μ) =

$$\Phi\left(z_g + \frac{m_o - m}{s/\sqrt{n}}\right) - \Phi\left(-z_{a-g} + \frac{m_o - m}{s/\sqrt{n}}\right)$$

c. Let $l = \sqrt{n}\frac{\Delta}{s}$; then we wish to know when $p(m_o + \Delta) = 1 - \Phi(z_g - l)$ $+ \Phi(-z_{a-g} - l) > 1 - \Phi(z_g + l) + \Phi(-z_{a-g} + l) = p(m_o - \Delta)$. Using the fact that $\Phi(-c) = 1 - \Phi(c)$, this inequality becomes $\Phi(z_g + l) - \Phi(z_g - l) > \Phi(z_{a-g} + l) - \Phi(z_{a-g} - l)$. The l.h.s. is the area under the Z curve above the interval $(z_g + l, z_g - l)$, while the r.h.s. is the area above $(z_{a-g} - l, z_{a-g} + l)$. Both intervals have width $2l$, but when $z_g < z_{a-g}$, the first interval is closer to 0 (and thus corresponds to the large area) than is the second. This happens when $g > a - g$, i.e., when $g > a/2$.

82.

a. $a = P(X \le 5$ when p = .9) = B(5; 10, .9) = .002, so the region (0, 1, …, 5) does specify a level .01 test.

b. The first value to be placed in the upper-tailed part of a two tailed region would be 10, but P(X = 10 when p = .9) = .349, so whenever 10 is in the rejection region, $a \ge .349$.

c. Using the two-tailed formula for ß(p') on p. 341, we calculate the value for the range of possible p' values. The values of p' we chose, as well as the associated ß(p') are in the table below, and the sketch follows. ß(p') seems to be quite large for a great range of p' values.

P'	Beta
0.01	0.0000
0.10	0.0000
0.20	0.0000
0.30	0.0071
0.40	0.0505
0.50	0.1635
0.60	0.3594
0.70	0.6206
0.80	0.8696
0.90	0.9900
0.99	1.0000