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 they cannot appear in a hypothesis.

f. Yes. H is an assertion about the value of a parameter.

3. We reject H_0 iff P-value $\leq \alpha = .05$.
a. Reject H_0 **b.** Reject H_0 **c.** Do not reject H_0 **d.** Reject H_0 **e.** Do not reject H_0

5. In this formulation, H_0 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: $\mu < 100$ results in the welds being believed in conformance unless proved otherwise, so the burden of proof is on the non-conformance claim.

7. Let σ denote the population standard deviation. The appropriate hypotheses are H_0: $\sigma = .05$ v. H_a: $\sigma < .05$. 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_0 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.

9. 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 judgment it is the type II error, then the reformulation H_0: $\mu = 150$ v. H_a: $\mu < 150$ makes the type I error more serious.

11.

a. A type I error consists of judging one of the two companies 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.

b. We expect 25(.5) = 12.5 "successes" when H_0 is true. So, any X-values less than 6 are at least as contradictory to H_0 as $x = 6$. But since the alternative hypothesis states $p \neq .5$, X-values that are just as far away on the <u>high</u> side are equally contradictory. Those are 19 and above.
So, values at least as contradictory to H_0 as $x = 6$ are {0,1,2,3,4,5,6,19,20,21,22,23,24,25}.

c. When H_0 is true, X has a binomial distribution with $n = 25$ and $p = .5$.
From part (b), P-value $= P(X \leq 6 \text{ or } X \geq 19) = B(6; 25,.5) + [1 - B(18; 25,.5)] = .014$.

d. Looking at Table A.1, a two-tailed P-value of .044 ($2 \times .022$) occurs when $x = 7$. That is, saying we'll reject H_0 iff P-value $\leq .044$ must be equivalent to saying we'll reject H_0 iff $X \leq 7$ or $X \geq 18$ (the same distance from 12.5, but on the high side). Therefore, for any value of $p \neq .5$, $\beta(p) = P(\text{do not reject } H_0 \text{ when } X \sim \text{Bin}(25, p)) = P(7 < X < 18 \text{ when } X \sim \text{Bin}(25, p)) = B(17; 25, p) - B(7; 25, p)$.
$\beta(.4) = B(17; 25,.4) - B(7, 25,.4) = .845$, while $\beta(.3) = B(17; 25, .3) - B(7; 25, .3) = .488$.
By symmetry (or re-computation), $\beta(.6) = .845$ and $\beta(.7) = .488$.

e. From part (c), the P-value associated with $x = 6$ is .014. Since $.014 \leq .044$, the procedure in (d) leads us to reject H_0.

13.

a. H_0: $\mu = 10$ v. H_a: $\mu \neq 10$.

b. Since the alternative is two-sided, values at least as contradictory to H_0 as $\bar{x} = 9.85$ are not only those less than 9.85 but also those equally far from $\mu = 10$ on the high side: i.e., $\bar{x}$ values ≥ 10.15.

When H_0 is true, $\bar{X}$ has a normal distribution with mean $\mu = 10$ and sd $\dfrac{\sigma}{\sqrt{n}} = \dfrac{.200}{\sqrt{25}} = .04$. Hence,

P-value $= P(\bar{X} \leq 9.85 \text{ or } \bar{X} \geq 10.15 \text{ when } H_0 \text{ is true}) = 2P(\bar{X} \leq 9.85 \text{ when } H_0 \text{ is true})$ by symmetry
$= 2P\left(Z < \dfrac{9.85-10}{.04}\right) = 2\Phi(-3.75) \approx 0$. (Software gives the more precise P-value .00018.)

In particular, since P-value $\approx 0 < \alpha = .01$, we reject H_0 at the .01 significance level and conclude that the true mean measured weight differs from 10 kg.

c. To determine $\beta(\mu)$ for any $\mu \neq 10$, we must first find the threshold between P-value $\leq \alpha$ and P-value $> \alpha$ in terms of $\bar{x}$. Parallel to part (b), proceed as follows:

$$.01 = P(\text{reject } H_0 \text{ when } H_0 \text{ is true}) = 2P(\bar{X} \leq \bar{x} \text{ when } H_0 \text{ is true}) = 2\Phi\left(\frac{\bar{x}-10}{.04}\right) \Rightarrow$$

$\Phi\left(\dfrac{\bar{x}-10}{.04}\right) = .005 \Rightarrow \dfrac{\bar{x}-10}{.04} = -2.58 \Rightarrow \bar{x} = 9.8968$. That is, we'd reject H_0 at the $\alpha = .01$ level iff the observed value of $\bar{X}$ is ≤ 9.8968 — or, by symmetry, $\geq 10 + (10 - 9.8968) = 10.1032$. Equivalently, we do not reject H_0 at the $\alpha = .01$ level if $9.8968 < \bar{X} < 10.1032$.
Now we can determine the chance of a type II error:
$\beta(10.1) = P(9.8968 < \bar{X} < 10.1032 \text{ when } \mu = 10.1) = P(-5.08 < Z < .08) = .5319$.
Similarly, $\beta(9.8) = P(9.8968 < \bar{X} < 10.1032 \text{ when } \mu = 9.8) = P(2.42 < Z < 7.58) = .0078$.

c. df = $n - 1 = 24$; the area to the left of -2.6 = the area to the right of $2.6 = .008$ according to Table A.8. Hence, the two-tailed P-value is $2(.008) = .016$. Since $.016 > .01$, we do not reject H_0 in this case.

d. Similar to part (c), Table A.8 gives a one-tail area of .000 for $t = \pm 3.9$ at df = 24. Hence, the two-tailed P-value is $2(.000) = .000$, and we reject H_0 at any reasonable α level.

31. This is an upper-tailed test, so the P-value in each case is $P(T \geq \text{observed } t)$.

a. P-value = $P(T \geq 3.2 \text{ with df} = 14) = .003$ according to Table A.8. Since $.003 \leq .05$, we reject H_0.

b. P-value = $P(T \geq 1.8 \text{ with df} = 8) = .055$. Since $.055 > .01$, do not reject H_0.

c. P-value = $P(T \geq -.2 \text{ with df} = 23) = 1 - P(T \geq .2 \text{ with df} = 23)$ by symmetry $= 1 - .422 = .578$. Since .578 is quite large, we would not reject H_0 at any reasonable α level. (Note that the sign of the observed t statistic contradicts H_a, so we know immediately not to reject H_0.)

33.

a. It appears that the true average weight could be significantly off from the production specification of 200 lb per pipe. Most of the boxplot is to the right of 200.

b. Let μ denote the true average weight of a 200 lb pipe. The appropriate null and alternative hypotheses are H_0: $\mu = 200$ and H_a: $\mu \neq 200$. Since the data are reasonably normal, we will use a one-sample t procedure. Our test statistic is $t = \dfrac{206.73 - 200}{6.35/\sqrt{30}} = \dfrac{6.73}{1.16} = 5.80$, for a P-value of ≈ 0. So, we reject H_0.

At the 5% significance level, the test appears to substantiate the statement in part **a**.

35.

a. The hypotheses are H_0: $\mu = 200$ versus H_a: $\mu > 200$. With the data provided,

$t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} = \dfrac{249.7 - 200}{145.1/\sqrt{12}} = 1.2$; at df = 12 − 1 = 11, P-value = .128. Since $.128 > .05$, H_0 is not rejected at the $\alpha = .05$ level. We have insufficient evidence to conclude that the true average repair time exceeds 200 minutes.

b. With $d = \dfrac{|\mu_0 - \mu|}{\sigma} = \dfrac{|200 - 300|}{150} = 0.67$, df = 11, and $\alpha = .05$, software calculates power $\approx .70$, so $\beta(300) \approx .30$.

41. μ = true average reading, H_0: $\mu = 70$ v. H_a: $\mu \neq 70$, 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 = $2[P(T > 1.92)] \approx 2(.058) = .116$. At significance level .05, there is not enough evidence to conclude that the spectrophotometer needs recalibrating.

Section 8.4

43.

a. The parameter of interest is p = the proportion of the population of female workers that have BMIs of at least 30 (and, hence, are obese). The hypotheses are H_0: $p = .20$ versus H_a: $p > .20$.
With $n = 541$, $np_0 = 541(.2) = 108.2 \geq 10$ and $n(1 - p_0) = 541(.8) = 432.8 \geq 10$, so the "large-sample" z procedure is applicable.

From the data provided, $\hat{p} = \dfrac{120}{541} = .2218$, so $z = \dfrac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} = \dfrac{.2218-.20}{\sqrt{.20(.80)/541}} = 1.27$ and P-value $= P(Z \geq 1.27) = 1 - \Phi(1.27) = .1020$. Since $.1020 > .05$, we fail to reject H_0 at the $\alpha = .05$ level. We do not have sufficient evidence to conclude that more than 20% of the population of female workers is obese.

b. A Type I error would be to incorrectly conclude that more than 20% of the population of female workers is obese, when the true percentage is 20%. A Type II error would be to fail to recognize that more than 20% of the population of female workers is obese when that's actually true.

c. The question is asking for the chance of committing a Type II error when the true value of p is .25, i.e. $\beta(.25)$. Using the textbook formula,

$$\beta(.25) = \Phi\left[\frac{.20-.25+1.645\sqrt{.20(.80)/541}}{\sqrt{.25(.75)/541}}\right] = \Phi(-1.166) \approx .121.$$

45. Let p = true proportion of all donors with type A blood. The hypotheses are H_0: $p = .40$ versus H_a: $p \neq .40$.
Using the one-proportion z procedure, the test statistic is $z = \dfrac{82/150-.40}{\sqrt{.40(.60)/150}} = \dfrac{.147}{.04} = 3.667$, and the corresponding P-value is $2P(Z \geq 3.667) \approx 0$. Hence, we reject H_0. The data does suggest that the percentage of all donors with type A blood differs from 40%. (at the .01 significance level). Since the P-value is also less than .05, the conclusion would not change.

47.

a. The parameter of interest is p = the proportion of all wine customers who would find screw tops acceptable. The hypotheses are H_0: $p = .25$ versus H_a: $p < .25$.
With $n = 106$, $np_0 = 106(.25) = 26.5 \geq 10$ and $n(1 - p_0) = 106(.75) = 79.5 \geq 10$, so the "large-sample" z procedure is applicable.

From the data provided, $\hat{p} = \dfrac{22}{106} = .208$, so $z = \dfrac{.208-.25}{\sqrt{.25(.75)/106}} = -1.01$ and P-value $= P(Z \leq -1.01) = \Phi(-1.01) = .1562$.
Since $.1562 > .10$, we fail to reject H_0 at the $\alpha = .10$ level. We do not have sufficient evidence to suggest that less than 25% of all customers find screw tops acceptable. Therefore, we recommend that the winery <u>should</u> switch to screw tops.

65.

a. From Table A.17, when $\mu = 9.5$, $d = .625$, and df = 9, $\beta \approx .60$.
When $\mu = 9.0$, $d = 1.25$, and df = 9, $\beta \approx .20$.

b. From Table A.17, when $\beta = .25$ and $d = .625$, $n \approx 28$.

67.

a. With H_0: $p = 1/75$ v. H_a: $p \neq 1/75$, $\hat{p} = \frac{16}{800} = .02$, $z = \frac{.02 - .01333}{\sqrt{\frac{.01333(.98667)}{800}}} = 1.645$, and P-value = .10, we fail to reject the null hypothesis at the $\alpha = .05$ level. There is no significant 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_0.

69. Even though the underlying distribution may not be normal, a z test can be used because n is large. The null hypothesis H_0: $\mu = 3200$ should be rejected in favor of H_a: $\mu < 3200$ if the P-value is less than .001. The computed test statistic is $z = \frac{3107 - 3200}{188/\sqrt{45}} = -3.32$ and the P-value is $\Phi(-3.32) = .0005 < .001$, so H_0 should be rejected at level .001.

71. We wish to test H_0: $\mu = 4$ versus H_a: $\mu > 4$ 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 $z = \frac{4.444 - 4}{\sqrt{4/36}} = 1.33$.

The P-value is $P(Z \geq 1.33) = 1 - \Phi(1.33) = .0918$. Since .0918 > .02, H_0 should not be rejected at this level. We do not have significant evidence at the .02 level to conclude that the true mean of this Poisson process is greater than 4.

73. The parameter of interest is p = the proportion of all college students who have maintained lifetime abstinence from alcohol. The hypotheses are H_0: $p = .1$, H_a: $p > .1$.
With $n = 462$, $np_0 = 462(.1) = 46.2 \geq 10$ $n(1 - p_0) = 462(.9) = 415.8 \geq 10$, so the "large-sample" z procedure is applicable.

From the data provided, $\hat{p} = \frac{51}{462} = .1104$, so $z = \frac{.1104 - .1}{\sqrt{.1(.9)/462}} = 0.74$.

The corresponding one-tailed P-value is $P(Z \geq 0.74) = 1 - \Phi(0.74) = .2296$.
Since .2296 > .05, we fail to reject H_0 at the $\alpha = .05$ level (and, in fact, at any reasonable significance level). The data does not give evidence to suggest that more than 10% of all college students have completely abstained from alcohol use.

75. Since n is large, we'll use the one-sample z procedure. With μ = population mean Vitamin D level for infants, the hypotheses are H_0: $\mu = 20$ v. H_a: $\mu > 20$. The test statistic is $z = \dfrac{21-20}{11/\sqrt{102}} = 0.92$, and the upper-tailed P-value is $P(Z \geq 0.92) = .1788$. Since $.1788 > .10$, we fail to reject H_0. It cannot be concluded that $\mu > 20$.

77. The 20 df row of Table A.7 shows that $\chi^2_{.99,20} = 8.26 < 8.58$ (H_0 not rejected at level .01) and $8.58 < 9.591 = \chi^2_{.975,20}$ (H_0 rejected at level .025). Thus $.01 < P\text{-value} < .025$, and H_0 cannot be rejected at level .01 (the P-value is the smallest α at which rejection can take place, and this exceeds .01).

79.

a. When H_0 is true, $2\lambda_0 \Sigma X_i = \dfrac{2}{\mu_0}\sum X_i$ has a chi-squared distribution with df = $2n$. If the alternative is H_a: $\mu < \mu_0$, then we should reject H_0 in favor of H_a when the sample mean $\bar{x}$ is small. Since $\bar{x}$ is small exactly when Σx_i is small, we'll reject H_0 when the test statistic is small. In particular, the P-value should be the area to the <u>left</u> of the observed value $\dfrac{2}{\mu_0}\sum x_i$.

b. The hypotheses are H_0: $\mu = 75$ versus H_a: $\mu < 75$. The test statistic value is $\dfrac{2}{\mu_0}\sum x_i = \dfrac{2}{75}(737) =$ 19.65. At df = 2(10) = 20, the P-value is the area to the left of 19.65 under the χ^2_{20} curve. From software, this is about .52, so H_0 clearly should not be rejected (the P-value is very large). The sample data do not suggest that true average lifetime is less than the previously claimed value.