

(1)

## Problems Set 11

8.61 From 8.41,  $\hat{p} = \frac{2}{3}$ . It is given that  $B = .02$ ,  $\alpha = .01$ , and  $z_{\alpha/2} = z_{.005} = 2.576$ . Then

$$z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = B$$

or

$$2.576 \sqrt{\frac{(\frac{2}{3})(\frac{1}{3})}{n}} = .02.$$

Implying

$$n = \frac{(2.576)^2 (\frac{2}{3})(\frac{1}{3})}{(.02)^2} = 3686.54.$$

Hence we round  $n = 3687$ .

8.62 This is similar to previous exercises, with  $B = .1$  and  $\sigma = .5$ . The required sample size is obtained by solving

$$.1 = 1.96 \left( \frac{.5}{\sqrt{n}} \right) \quad \text{so that} \quad \sqrt{n} = 5(1.96) \quad \text{or} \quad n = 97$$

Notice that water specimens should be selected randomly and not from the same rainfall, in order that all observations be independent.

8.70 a.  $n = 20$ ,  $\bar{x} = 419$ ,  $s = 57$ . Then the 90% confidence interval for the mean SAT scores for urban high school seniors is

$$\bar{y} \pm t_{.05} \left( \frac{s}{\sqrt{n}} \right)$$

where  $t_{.05}$  is based on  $n - 1 = 19$  degrees of freedom. From the Appendix, this is  $t_{.05} = 1.729$ . Then the confidence interval is

$$419 \pm 1.729 \left( \frac{57}{\sqrt{20}} \right) = 419 \pm 22.04 = (396.96, 441.04).$$

- b. The interval does include 422. Thus 422 is a believable value for  $\mu$  at the 90% confidence level. However, numbers such as 397, 410, and 441, for example, are also believable values for  $\mu$ .
- c. Given  $n = 20$ ,  $\bar{x} = 455$ ,  $s = 69$ , the 90% confidence interval for the mean mathematics SAT score is

$$\bar{y} \pm t_{.05} \left( \frac{s}{\sqrt{n}} \right) = 455 \pm 1.729 \left( \frac{69}{\sqrt{20}} \right) = 455 \pm 26.67 = (428.33, 481.67).$$

The interval does include 474. We would conclude, based on our 90% confidence interval, that the true mean mathematics SAT score is not different from 474.

2

8.74 For the  $n = 12$  measurements given here, calculate  $\sum y_i = 108$  and  $\sum y_i^2 = 1426$ . Then

$$\bar{y} = \frac{108}{12} = 9 \quad \text{and} \quad s^2 = \frac{1426 - \frac{(108)^2}{12}}{11} = 41.2727$$

The 90% confidence interval is then

$$\bar{y} \pm t_{.05} \left( \frac{s}{\sqrt{n}} \right) = \pm 1.796 \sqrt{\frac{41.2727}{12}} = 9 \pm 3.33 \quad \text{or} \quad (5.67, 12.33).$$

8.75 Refer to Exercise 8.74, where  $\bar{y}_1 = 9$  and  $(n_1 - 1)s_1^2 = \sum y_1^2 - \left[ \frac{(\sum y_1)^2}{n} \right] = 454$ . From this,

$$\begin{aligned} \sum y_2 &= 10.7 & \sum y_2^2 &= 65.09 \\ \sum y_2^2 - \frac{(\sum y_2)^2}{n_2} &= 26.92667 & s_2^2 &= \frac{26.92667}{2} = 13.4633 \end{aligned}$$

a. The 90% confidence interval for  $\mu_2$  is

$$\begin{aligned} \bar{y}_2 \pm t_{.05} \left( \frac{s_2}{\sqrt{n_2}} \right) &= 3.57 \pm 2.92 \sqrt{\frac{13.4633}{3}} \\ &= 3.57 \pm 6.19 = (-2.62, 9.76). \end{aligned}$$

b. Calculate

$$S_p^2 = \frac{454 + 26.92667}{13} = 36.9944$$

The 90% confidence interval for  $\mu_1 - \mu_2$  is

$$\begin{aligned} (\bar{y}_1 - \bar{y}_2) \pm t_{.05} \sqrt{S_p^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)} &= (9 - 3.57) \pm 1.771 \sqrt{36.9944 \left( \frac{1}{12} + \frac{1}{3} \right)} = 5.43 \pm 6.95 \\ \text{or } (-1.52, 12.38). &\text{ We must assume that the LC50 measurements are normally distributed} \\ \text{and independent and that } \sigma_1^2 &= \sigma_2^2. \end{aligned}$$

8.85 a. A  $100(1 - \alpha)\%$  upper confidence bound for  $\sigma$  is

$$\sqrt{\frac{(n-1)s^2}{\chi_{1-\alpha}^2}}$$

where  $\chi_{1-\alpha}^2$  is chosen such that

$$P \left( \frac{(n-1)s^2}{\sigma^2} \geq \chi_{1-\alpha}^2 \right) = 1 - \alpha.$$

b. A  $100(1 - \alpha)\%$  lower confidence bound for  $\sigma$  is

$$\sqrt{\frac{(n-1)s^2}{\chi_{\alpha}^2}}$$

where  $\chi_{1-\alpha}^2$  is chosen so that

$$P \left( \frac{(n-1)s^2}{\sigma^2} \leq \chi_{\alpha}^2 \right) = 1 - \alpha.$$

8.86 Exercise 8.85(a) gives the  $100(1 - \alpha)\%$  upper confidence bound for  $\sigma$  as

$$\sqrt{\frac{(n-1)s^2}{\chi_{1-\alpha}^2}}$$

where  $\chi_{1-\alpha}^2$  is chosen so that

$$P \left( \frac{(n-1)s^2}{\sigma^2} \geq \chi_{1-\alpha}^2 \right) = 1 - \alpha.$$

With  $(n - 1) = 19$  degrees of freedom,  $\chi_{.99}^2 = 7.6327$ . Then

$$(n - 1)s^2 = \sum x_i^2 - \frac{(\sum x_i)^2}{n} = 92,305,600 - \frac{(42,812)^2}{20} = 662,232.8$$

and the bound is

$$\sqrt{\frac{662,232.8}{7.6327}} = 294.55.$$

At the 99% confidence level, the population standard deviation could be less than 150 because the bound is greater than 150. This means that values less than 150 are just as believable as values above 150.

8.87 Calculate

$$\sum y_i = 56.91 \quad \sum y_i^2 = 539.9341 \quad \sum y_i^2 - \frac{(\sum y_i)^2}{n} = .14275$$

Thus,  $s^2 = .02855$ . Then the 90% confidence interval for  $\sigma^2$  is

$$\frac{(n-1)s^2}{\chi_{.05}^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_{.95}^2} \quad \text{or} \quad \frac{.14275}{11.0705} < \sigma^2 < \frac{.14275}{1.145476} \quad \text{or} \quad (.0129, .1246)$$

10.2 The test statistic  $Y$  has a binomial distribution with  $n = 20$  and  $p$ .

- A Type I error occurs if the experimenter concluded that the drug dosage level induces sleep in less than 80% of the people suffering from insomnia when, in fact, drug dosage level does induce sleep in 80% of insomniacs.
- $\alpha = P(\text{reject } H_0 | H_0 \text{ true}) = P(Y \leq 12 | p = .8) = .032$ , using Table I, Appendix III.
- A Type II error would occur if the experimenter concluded that the drug dosage level induces sleep in 80% of the people suffering from insomnia when, in fact, fewer than 80% experience relief.
- If  $p = .6$ ,  

$$\beta = P(\text{accept } H_0 | H_0 \text{ false}) = P(Y > 12 | p = .6) = 1 - P(Y \leq 12 | p = .6) = 1 - .584 = .416$$
- If  $p = .4$ , then  

$$\beta = P(Y > 12 | p = .4) = 1 - P(Y \leq 12 | p = .4) = 1 - .979 = .021.$$

- 10.3
- With  $n = 20$  and  $p = .8$ , it is necessary to find  $c$  such that  $\alpha = P(Y \leq c | p = .8) = .01$ . From Table I, Appendix III, this value is  $c = 11$ .
  - With the rejection region given as  $Y \leq 11$ ,  

$$\beta = P(Y > 11 | p = .6) = 1 - P(Y \leq 11 | p = .6) = 1 - .404 = .596$$
  - $$\beta = P(Y > 11 | p = .4) = 1 - P(Y \leq 11 | p = .4) = 1 - .943 = .057$$

- 10.4
- A Type I error occurs if we conclude that the proportion of ledger sheets with errors is larger than .05 when, in fact, the proportion is .05.
  - By the scheme being used, we will reject for the following situations:  
 (NOTE: NE = no error, E = error)

Sheet 1	Sheet 2	Sheet 3
NE	NE	.
NE	E	NE
E	NE	NE
E	E	NE

- Thus,  $\alpha = (.95)^2 + 2(.05)(.95)^2 + (.05)^2(.95) = .9025 + .09025 + .002375 = .995125$ .
- A Type II error occurs if we conclude that the proportion of ledger sheets with errors is .05 when, in fact, the proportion is larger than .05.
  - $$\beta = P(\text{accept } H_0 \text{ when } H_a \text{ is true}) = P(\text{accepting } H_0 | p = p_a)$$

$$= 2p_a^2(1 - p_a) + p_a^3.$$
 Since we reject if we observe E,E,E or NE,E,E or E,NE,E,

- 10.7 a. Since it is necessary to test a claim that the average amount saved,  $\mu$ , is \$900, the hypothesis to be tested is two-tailed:

$$H_0: \mu = 900 \quad \text{vs.} \quad H_a: \mu \neq 900$$

- b. The rejection region with  $\alpha = .01$  is determined by a critical value of  $z$  such that  $P[|z| > z_0] = .01$

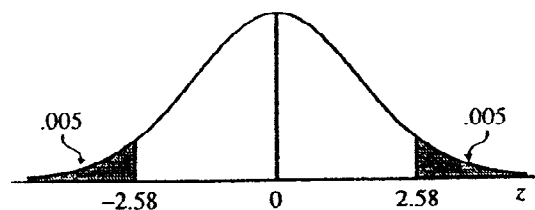


Figure 10.1

This value is  $z_0 = 2.58$  (see Figure 10.1) and the rejection region is  $|z| > 2.58$ .

- c. The test statistic is

$$z = \frac{\bar{y} - \mu}{\frac{\sigma}{\sqrt{n}}} \approx \frac{\bar{y} - \mu}{\frac{s}{\sqrt{n}}} = \frac{885 - 900}{\frac{30}{\sqrt{35}}} = -1.77$$

- d. The observed value,  $z = -1.77$ , does not fall in the rejection region, and  $H_0$  is not rejected. We cannot conclude that the average savings is different than claimed.

- 10.10 Let  $\mu$  be the average hardness index. We are to test  $H_0: \mu \geq 64$  vs.  $H_a: \mu < 64$ . The test statistic is

$$z = \frac{\bar{y} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{62 - 64}{\frac{8}{\sqrt{50}}} = -1.77$$

The rejection region is RR: Reject  $H_0$  if  $z < -z_{.01} = -2.326$ .

Conclusion: Do not reject  $H_0$  at  $\alpha = .01$ . There is insufficient evidence to reject the manufacturer's claim.

- 10.14a. If we define  $p$  as the proportion of college students aged 30 years or more, then we test

$$H_0: p = .25 \quad \text{vs.} \quad H_a: p \neq .25$$

The test statistic is

$$z = \frac{\bar{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{\frac{98}{300} - .25}{\sqrt{\frac{(.25)(.75)}{300}}} = 3.07$$

and the rejection region, with  $\alpha = .05$  is  $|z| > 1.96$ .  $H_0$  is rejected and we conclude that the 25% figure is not accurate.

- b. Yes, the results do give evidence that the columnist's claim is too low.

(5)

10.24 If  $\mu$  is the average length of stay the hypothesis of interest is

$$H_0: \mu = 5 \quad \text{vs.} \quad H_a: \mu > 5$$

and the test statistic is approximately

$$z = \frac{\bar{y} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{5.4 - 5}{\frac{1.1}{\sqrt{500}}} = 2.89$$

The rejection region with  $\alpha = .05$  is  $z > 1.645$ . Hence we reject  $H_0$  and support the agency's hypothesis.

10.25 Let  $p_1$  be the proportion of homeless men currently working and  $p_2$  be the proportion of domiciled men currently working. The hypothesis of interest is

$$H_0: p_1 - p_2 = 0 \quad \text{vs.} \quad H_a: p_1 - p_2 < 0.$$

Calculate

$$\hat{p}_1 = \frac{34}{112} = .30, \hat{p}_2 = .38, \text{ and } \hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{34 + 98}{112 + 260} = .355$$

The test statistic is then

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{.30 - .38}{\sqrt{(.355)(.645)\left(\frac{1}{112} + \frac{1}{260}\right)}} = -1.48.$$

The rejection region with  $\alpha = .01$  is  $z < -2.326$  and  $H_0$  is not rejected. There is no evidence that the proportion of homeless men working is less than the proportion of domiciled men.

10.51 The hypothesis to be tested is

$$H_0: \mu = 800 \quad \text{vs.} \quad H_a: \mu < 800.$$

However, there are only five measurements on which to base the test, and a  $t$  statistic

must be employed. The test statistic is  $T = \frac{\bar{Y} - \mu}{\frac{s}{\sqrt{n}}}$

where

$$\bar{y} = \frac{\sum y_i}{n} = \frac{3975}{5} = 795$$

and

$$s^2 = \frac{\sum \left( \frac{\sum y_i}{n} \right)^2}{n-1} = \frac{3,160,403 - \frac{15,800,625}{5}}{4} = 69.5$$

Then

$$t = \frac{\bar{y} - \mu}{\frac{s}{\sqrt{n}}} = \frac{795 - 800}{\frac{\sqrt{69.5}}{\sqrt{5}}} = \frac{-5}{3.728} = -1.341$$

The rejection region is determined by a  $t$  value based on  $(n - 1) = 4$  degrees of freedom.

Indexing  $t_{.05}$  in Table 5, the rejection region is  $t < -2.132$ . Since the observed value of the test statistic does not fall in the rejection region, we do not reject  $H_0$ . Refer to Figure 10.6 at the right and notice that the  $t$  distribution is similar to the  $z$  distribution. The shaded area constitutes the size of the rejection region.

To find bounds on the  $p$ -value, note in Table 5 that a value of  $t = -1.553$  would yield a  $p$ -value of .10. The observed value of  $t$  is  $-1.341 > -1.553$ . Thus,  $p\text{-value} > .10$ .

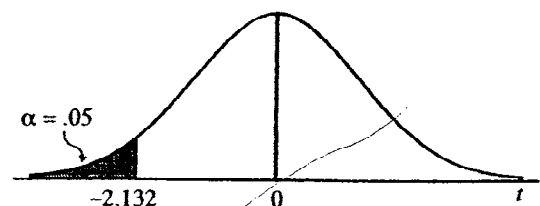


Figure 10.6

(6)

10.57 This is similar to Examples 10.14 and 10.15. The hypothesis to be tested is

$$H_0: \mu_1 - \mu_2 = 0 \quad \text{vs.} \quad H_a: \mu_1 - \mu_2 \neq 0.$$

We must assume that the data come from two normal populations with a common variance. We obtain an estimate for the common variance  $\sigma^2$  by calculating

$$s^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2} = \frac{10(52) + 13(71)}{11 + 14 - 2} = \frac{1443}{23} = 62.74$$

The test statistic is

$$t = \frac{\bar{y}_1 - \bar{y}_2 - D_0}{\sqrt{s^2 \left[ \left( \frac{1}{n_1} \right) + \left( \frac{1}{n_2} \right) \right]}} = \frac{64 - 69}{\sqrt{62.74 \left[ \left( \frac{1}{11} \right) + \left( \frac{1}{14} \right) \right]}} = -1.57$$

The rejection region is  $|t| > t_{.025, 23} = 2.069$ . Since the observed value of the test statistic does not exceed  $t = 2.069$  in absolute value, the null hypothesis is not rejected. Note in Table 5 that  $1.319 < 1.57 < 1.714$ . The  $p$ -values associated with 1.319 and 1.714 are 2(.10) and 2(.05). Thus,  $.10 < p\text{-value} < .20$ .

10.62 The hypothesis of interest is

$$H_0: \mu = 6 \quad \text{vs.} \quad H_a: \mu < 6,$$

and the test statistic is

$$t = \frac{\bar{y} - \mu_0}{\sqrt{\frac{s^2}{n}}} = \frac{9 - 6}{\sqrt{\frac{31.2727}{12}}} = 1.62$$

The rejection region is  $t < -t_{.05, 11} = -1.796$ , and the null hypothesis is not rejected. There is insufficient evidence to indicate that  $\mu$  is less than 6.