10.6 A large sample 98% confidence interval for $\mu_1 - \mu_2$ is given by

$$(\bar{x} - \bar{y}) \pm z_{0.01} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = (76.4 - 81.2) \pm 2.33 \sqrt{\frac{8.2^2}{90} + \frac{7.6^2}{100}} = -4.8 \pm 2.68$$

or $(-7.48, -2.12)$. We are 98% confident that urban students have a mean score 2.12 to 7.48 points higher than that of the rural students.

10.14 (a) We first obtain:

MALE $\bar{x} = 6$

$$s_1^2 = \frac{2^2 + (-2)^2 + 0}{2} = 4$$

FEMALE $\bar{y} = 3$

$$s_2^2 = \frac{2^2 + (-2)^2 + 0}{2} = 4$$

Consequently, the pooled variance is given by

$$s_{\text{pooled}}^2 = \frac{\sum (x_i - \bar{x})^2 + \sum (y_i - \bar{y})^2}{n_1 + n_2 - 2} = \frac{8 + 8}{3 + 3 - 2} = 4$$

(b) We estimate the common sigma by $s_{\text{pooled}} = \sqrt{4} = 2$.

(c) The $t$-statistic is $T = \frac{\bar{x} - \bar{y}}{s_{\text{pooled}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ with d.f. $= n_1 + n_2 - 2$. In the present problem, we have $T = \frac{6 - 6}{2 \sqrt{\frac{1}{3} + \frac{1}{3}}} = 0$, with d.f. $= 4$.

10.22 The summary statistics are:

Isometric Method: $n_1 = 10$, $\bar{x} = 2.4$, $s_1 = 0.8$

Isotonic Method: $n_2 = 10$, $\bar{y} = 3.2$, $s_2 = 1.0$

(a) Denote by $\mu_1$ and $\mu_2$ the population mean decrease in abdomen measurements under the Isometric Method and Isotonic Method, respectively. Since the intent is to demonstrate that the Isotonic Method is more effective (that is, $\mu_1$ is smaller than $\mu_2$), we formulate the hypotheses

$$H_0: \mu_1 - \mu_2 = 0 \text{ versus } H_1: \mu_1 - \mu_2 < 0$$

We assume normal populations with equal variance, so the test statistic is

$$T = \frac{\bar{x} - \bar{y}}{s_{\text{pooled}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \text{ with d.f. } = n_1 + n_2 - 2$$
Since $H_1$ is left-sided, the rejection region is $R: T \leq -t_{0.05} = -1.734$ for d.f. $= 10 + 10 - 2 = 18$. Using the summary statistics, we obtain
\[ \bar{x} - \bar{y} = 2.4 - 3.2 = -0.8 \]
\[ s_{\text{pooled}} = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{9(0.8)^2 + 9(1.0)^2}{18}} = 0.906 \]
Hence, the observed value of $t$ is given by
\[ t = -0.8 \cdot \sqrt{\frac{1}{10} + \frac{1}{10}} = -0.8 \cdot \frac{1}{0.405} = -1.98, \]
which lies in $R$. So, $H_0$ is rejected at $\alpha = 0.05$. As such, the superiority of the Isotonic Method is demonstrated by the data.

(b) Using the calculations from part (a), a corresponding 95% confidence interval has the form
\[ (\bar{x} - \bar{y}) \pm t_{0.05, \frac{1}{2}} s_{\text{pooled}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = -0.8 \pm 2.101(0.405) = -0.8 \pm 0.85 \]
or $(-1.65, 0.05)$ centimeters.

10.26 The summary statistics are:
\[ n_1 = 40, \quad \bar{x} = 14.38, \quad s_1 = 7.34 \]
\[ n_2 = 43, \quad \bar{y} = 23.72, \quad s_2 = 5.78 \]
Since the sample sizes are large, neither the assumption of normal populations nor the assumption of equal variances is needed. A large sample 99% confidence interval for $\mu_1 - \mu_2$ is given by
\[ (\bar{x} - \bar{y}) \pm z_{0.005} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \]
Observe that
\[ z_{0.005} = 2.58 \]
\[ \bar{x} - \bar{y} = 14.38 - 23.72 = -9.34 \]
\[ \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{7.34^2}{40} + \frac{5.78^2}{43}} = 1.457 \]
So, the 99% confidence interval becomes
\[ -9.34 \pm 2.58(1.457) = -9.34 \pm 3.759 \quad \text{or} \quad (-13.099, -5.581). \]
10.38 (a) The paired differences $d = (A - B)$ are 0.9, -0.1, 0.8, 0, -0.1, 0.3. Assume that these constitute a random sample from a population with distribution $N(\delta, \sigma^2)$. Observe that

\[ n = 6, \bar{d} = 0.3, s_d = 0.452. \]

\[ \frac{s_d}{\sqrt{d}} = \frac{0.452}{\sqrt{6}} = 0.1845. \]

For $\alpha = 0.05$ and d.f. = 5, we have $t_{0.025} = 2.571$. So, the 95% confidence interval for $\delta$ is given by

\[ \bar{d} \pm t_{0.025} \frac{s_d}{\sqrt{d}} = 0.3 \pm 2.571 \left( \frac{0.452}{\sqrt{6}} \right) = 0.3 \pm 0.47 \text{ or } (-0.17, 0.77). \]

(b) For each person, the assignment of A and B should be made at random between the two nights. This can be done, for instance, by tossing a coin for each person. If this coin falls Heads, then give A the first night and B the second; if Tails appear, then give B the first night and A the second.

10.40 (a) Observe that

\[ n = 10, \bar{d} = 42.9, \text{ and } s_d = 34.346. \]

For $\alpha = 0.05$ and d.f. = 9, we have $t_{0.025} = 2.262$. So, the 95% confidence interval for $\delta$ (the mean of $D = \text{previous year usage} - \text{experimental year usage}$) is given by

\[ \bar{d} \pm t_{0.025} \frac{s_d}{\sqrt{d}} = 42.9 \pm 2.262 \left( \frac{34.346}{\sqrt{9}} \right) \text{ or } (18.3, 67.5). \]

(b) We reject $H_0 : \delta = 0$ for any value of the observed $t$ which lies in the rejection region $R : |T| \geq t_{0.025} = 2.262$. Here, $t = \frac{\bar{d} - \mu}{S_d} = \frac{42.9 - 0}{34.346/\sqrt{9}} = 3.915$, which lies in $R$. Hence, we reject $H_0$ in favor of $H_1 : \delta \neq 0$ at $\alpha = 0.05$.

(c) Some persons would have to be under the pricing policy the first year, and others on it the second year. The allocation would be made at random.

(d) If July the previous year was much hotter than July in the experimental year, then air conditioner electrical power requirements could be the cause of the drop in usage. Without randomizing the price scheme over the two summers, we cannot separate the two causes: temperature and price.

10.42 Here we have a situation of independent random samples because 16 subjects (farms) were randomly divided into two groups of 8 each to be assigned to the two treatments. We calculate the following summary statistics:

Strain A: \( n_1 = 8, \bar{x} = 32.0, s_1 = 8.83 \)

Strain B: \( n_2 = 8, \bar{y} = 27.5, s_2 = 8.37 \)

Denote by $\mu_1$ and $\mu_2$ the population mean yields of strain A and strain B, respectively. Since the conjecture is that $\mu_1$ is larger than $\mu_2$, we formulate the hypotheses

\[ H_0 : \mu_1 - \mu_2 = 0 \text{ versus } H_1 : \mu_1 - \mu_2 > 0. \]
We assume normal populations with equal variance, so the test statistic is

\[ T = \frac{\bar{X} - \bar{Y}}{s_{\text{pooled}}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \]

with d.f. = \( n_1 + n_2 - 2 \)

Since \( H_1 \) is right-sided, the rejection region is \( R : T \geq t_{0.05} = 1.761 \) for d.f. = 8 + 8 - 2 = 14. Using the summary statistics, we obtain

\[ \bar{X} - \bar{Y} = 4.50 \]

\[ s_{\text{pooled}} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{7(18.33)^2 + 7(8.37)^2}{14}} = 8.60 \]

Hence, the observed value of \( t \) is given by

\[ t = \frac{4.50}{8.60} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 1.05 \]

which does not lie in \( R \). So, \( H_0 \) is not rejected at \( \alpha = 0.05 \). We conclude that the claim of a higher mean yield for strain A fails to be demonstrated by the data.

(b) The selection of 8 farms for assigning each treatment should be randomized.

Tag the 16 farms 1, 2, ..., 16, place 16 numbered marbles in an urn, and randomly draw 8 marbles. Assign the farms corresponding to the selected marbles to strain A and the others to strain B.

(c) Comparing the calculations in Exercises 10.41 and 10.42, we find that \( d = \bar{X} - \bar{Y} = 4.50 \). So, the numerators of the two \( t \)-statistics are the same.

However, the estimated S.E.s are quite different:

For matched pair: \( \frac{t_{0.05}}{s_{\text{pooled}}} = 4.417 = 1.453 \)

For independent samples: \( s_{\text{pooled}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 4.30 \)

A considerable reduction of variability in the yield data has been accomplished by means of the matched pair design.

10.46 (a) We formulate the hypotheses:

\[ H_0 : p_A = p_B \quad \text{versus} \quad H_1 : p_A < p_B \]

The test statistic is \( Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_q \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}} \), where \( \hat{p}_1 \) is identified as \( \hat{p}_A \) and \( \hat{p}_2 \) is identified as \( \hat{p}_B \). Since \( H_1 \) is left-sided and \( \alpha = 0.05 \), the rejection region is \( R : Z \leq -z_{0.05} = -1.645 \). We calculate the following:

\[ \hat{p}_1 = \frac{10}{20} = 0.50, \quad \hat{p}_2 = \frac{18}{20} = 0.90 \]

\[ \hat{p}_q = \frac{\hat{p}_1 + \hat{p}_2}{2} = 0.70 \]

\[ Z = \frac{0.50 - 0.90}{\sqrt{0.70 \left( \frac{1}{20} + \frac{1}{20} \right)}} = -1.645 \]

So, \( H_0 \) is rejected at \( \alpha = 0.05 \). We conclude that there is evidence to support the claim that strain A has a higher mean yield than strain B.
Pooled estimate \( \hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \frac{120(0.417) + 150(0.587)}{120 + 150} = 0.511 \),

Observed \( z \) is \( \frac{0.417 - 0.587}{\sqrt{\frac{1}{120} + \frac{1}{150}}} = -0.17 \), \( E = -2.778 \),

which lies in \( R \). Hence, \( H_0 \) is rejected at \( \alpha = 0.05 \). Furthermore, the associated \( p \)-value is \( P[Z \leq -2.778] = 0.0028 \). So, there is very strong evidence that drug B has a higher cure rate than drug A. The rejection region and \( p \)-value are illustrated below (in that order):

\[
(b) \quad \hat{p}_2 - \hat{p}_1 = 0.587 - 0.417 = 0.17
\]

Estimated S.E. \( \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = \sqrt{\frac{(0.417)(0.583) + (0.587)(0.413)}{120 + 150}} = 0.060 \)

So, a 95% confidence interval for \( p_1 - p_2 \) is given by

\[
\left( \hat{p}_1 - \hat{p}_2 \right) \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = 0.17 \pm 1.96(0.060) = 0.17 \pm 0.118
\]
or \((0.05, 0.29)\).

For testing the hypotheses \( H_0: p_1 = p_2 \) versus \( H_1: p_1 \neq p_2 \), we use the test statistic \( Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q} \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}} \).

(a) Let

\( p_1 \) = proportion for 25-34 year olds, and
\( p_2 \) = proportion for 35-44 year olds.

Since \( H_1 \) is two-sided and \( \alpha = 0.05 \), the rejection region is

\( R: |Z| \geq z_{\alpha/2} = 1.96 \). We calculate the following:

Pooled estimate \( \hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \frac{47(0.0638) + 64(0.2031)}{47 + 64} = 0.1719 \),

Observed \( z \) is \( \frac{0.2031 - 0.1505}{\sqrt{(0.1719)(0.8281)}} = 0.858 \),

which does not lie in \( R \). Hence, \( H_0 \) is not rejected at \( \alpha = 0.05 \).

(b) Let

\( p_1 \) = proportion for 19-24 year olds, and
\( p_2 \) = proportion for 25-34 year olds.

Again, since \( H_1 \) is two-sided and \( \alpha = 0.05 \), the rejection region is

\( R: |Z| \geq z_{\alpha/2} = 1.96 \). We calculate the following:

Pooled estimate \( \hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \frac{47(0.0638) + 64(0.2031)}{47 + 64} = 0.1441 \),

Observed \( z \) is \( \frac{-0.1393}{\sqrt{(0.1441)(0.8559)}} = 2.065 \),

which lies in \( R \). Hence, \( H_0 \) is rejected at \( \alpha = 0.05 \).
10.58 (a) We formulate the hypotheses:

\[ H_0 : \mu_1 = \mu_2 \quad \text{versus} \quad H_1 : \mu_1 > \mu_2 . \]

The test statistic is

\[ Z = \frac{\hat{\mu}_1 - \hat{\mu}_2}{\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}}, \]

where we identify \( \hat{\mu}_1 \) and \( \hat{\mu}_2 \), respectively. Since \( H_1 \) is right-sided and \( \alpha = 0.10 \), the rejection region is \( R : Z \geq z_{0.10} = 1.28 \). We calculate the following:

\[ \hat{\sigma}_1 = \frac{41}{160} = 0.256, \quad \hat{\sigma}_2 = \frac{43}{200} = 0.215 \]

Pooled estimate \( \hat{\mu} = \frac{n_1 \hat{\mu}_1 + n_2 \hat{\mu}_2}{n_1 + n_2} = \frac{78 + 43}{150 + 200} = 0.346, \]

Observed \( z \) is

\[ \frac{0.520 - 0.215}{\sqrt{(0.346)(0.654)\left(\frac{1}{150} + \frac{1}{200}\right)}} = 5.94, \]

which does not lie in \( R \). Hence, \( H_0 \) is not rejected at \( \alpha = 0.10 \).

(b) We formulate the hypotheses:

\[ H_0 : \mu_1 = \mu_2 \quad \text{versus} \quad H_1 : \mu_1 > \mu_2 . \]

The test statistic is

\[ Z = \frac{\hat{\mu}_1 - \hat{\mu}_2}{\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}}, \]

where we identify \( \hat{\mu}_1 \) and \( \hat{\mu}_2 \), respectively. Since \( H_1 \) is right-sided, the rejection region is of the form \( R : Z \geq c \). We calculate the following:

\[ \hat{\sigma}_1 = \frac{41}{160} = 0.256, \quad \hat{\sigma}_2 = \frac{43}{200} = 0.215 \]

Pooled estimate \( \hat{\mu} = \frac{n_1 \hat{\mu}_1 + n_2 \hat{\mu}_2}{n_1 + n_2} = \frac{78 + 43}{150 + 200} = 0.346, \]

Observed \( z \) is

\[ \frac{0.520 - 0.215}{\sqrt{(0.346)(0.654)\left(\frac{1}{150} + \frac{1}{200}\right)}} = 5.94, \]

The associated \( p \)-value \( P(Z \geq 5.94) \) is less than 0.0002. Hence, the data strongly substantiate \( H_1 \).

10.64 (a) Since the assertion is that \( \mu_a > \mu_b \), we formulate the hypotheses

\[ H_0 : \mu_a - \mu_b = 0 \quad \text{versus} \quad H_1 : \mu_a - \mu_b > 0 . \]

(b) Since the sample sizes \( n_1 = 55 \) and \( n_2 = 58 \) are large, we employ the \( Z \)-test, and so, the test statistic is

\[ Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} . \]

\( \bar{X} \) and \( \bar{Y} \) are the sample means and \( \alpha = 0.10 \), the rejection region is \( R : Z \geq z_{0.10} = 1.28 \).

(c) From the sample data, we calculate the value of the observed \( z \) to be

\[ z = \frac{4.64 - 4.03}{\sqrt{(1.25)^2 + (1.82)^2}} = 2.09 , \]

which lies in \( R \). Hence, \( H_0 \) is rejected at \( \alpha = 0.10 \). Furthermore, the associated \( p \)-value is \( P(Z \geq 2.09) = 0.0183 \), so the evidence is support of \( H_1 \) is strong.
10.84 (a) Let \( p_F \) and \( p_M \) denote the true proportion of females and males who donate to charity respectively. We formulate the hypotheses:
\[
H_0 : p_F = p_M \quad \text{versus} \quad H_1 : p_F > p_M.
\]
The test statistic is
\[
Z = \frac{\hat{p}_F - \hat{p}_M}{\sqrt{\hat{p}_F \hat{q}_F \frac{1}{n_F} + \hat{p}_M \hat{q}_M \frac{1}{n_M}}},
\]
where \( \hat{p}_F \) and \( \hat{p}_M \) are identified with \( \hat{p}_F \) and \( \hat{p}_M \). Since \( H_1 \) is right-sided and \( \alpha = 0.05 \), the rejection region is \( R: Z \geq z_{0.05} = 1.96 \). Using the given sample statistics, we have
\[
\text{Pooled estimate } \hat{p} = \frac{n_F \hat{p}_F + n_M \hat{p}_M}{n_F + n_M} = \frac{195(0.2462) + 294(0.1769)}{195 + 294} = 0.205.
\]
Observed \( z \) is
\[
\frac{0.2462 - 0.1769}{\sqrt{(0.205)(0.795) \left( \frac{1}{195} + \frac{1}{294} \right)}} = 1.858,
\]
which does not lie in \( R \). So, \( H_0 \) is not rejected at \( \alpha = 0.05 \).

(b) \( \hat{p}_F - \hat{p}_M = 0.0693 \)

Estimated S.E.
\[
\sqrt{\frac{\hat{p}_F \hat{q}_F + \hat{p}_M \hat{q}_M}{n_F} + \frac{\hat{p}_F \hat{q}_M + \hat{p}_M \hat{q}_F}{n_M}} = \sqrt{\frac{(0.2462)(0.7538) + (0.1769)(0.8231)}{195} + \frac{(0.2462)(0.1769) + (0.1769)(0.8231)}{294}} = 0.0380
\]
So, a 95% confidence interval for \( p_F - p_M \) is given by
\[
(\hat{p}_F - \hat{p}_M) \pm z_{0.05} \sqrt{\frac{\hat{p}_F \hat{q}_F + \hat{p}_M \hat{q}_M}{n_F} + \frac{\hat{p}_F \hat{q}_M + \hat{p}_M \hat{q}_F}{n_M}} = 0.0693 \pm 1.96(0.0380) = 0.0693 \pm 0.07448
\]
or \((-0.00518, 0.14378)\).