Chapter 14
Two-Sample Inference

14.1 Compute the differences x – y

x	7	9	9	6	8	8
y	4	4	3	4	9	5
d = x – y	3	5	6	2	–1	3

$\bar{d} = 3.000,\ s_d = 2.449$

The 95% confidence interval for the difference x – y is

$C.I. = \overline{D} \pm t_{5,.975} \frac{s}{\sqrt{n}}$ or $0.43 < \mu_0 < 5.6$.

14.3 The two-sided 99% confidence interval for the average increase in sleeping time for patients using the sleeping pill is of the form C.I.: $\mu_D = \bar{d} \pm t_{15,.995} \frac{s}{\sqrt{16}}$

Here $\bar{d}$ = 1.58 and standard deviation s = 1.23 so the interval is

99% C.I.: μ_D = 1.58 hours ± 0.91 hours or (0.67 hours, 2.49 hours).

14.5 Here $n_x = 8,\ \bar{x} = 35,\ s_x = 4,\ n_y = 12,\ \bar{y} = 20,\ s_y = 3$

a. The standard error the difference in means is

$$SE_{(\bar{x}-\bar{y})} = \sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}} = \sqrt{\frac{16}{8} + \frac{9}{12}} = \sqrt{2.75} = 1.658$$

b. The smaller sample size is 8 so a conservative way to compute the degrees of freedom is df = 8 – 1 = 7 and $t_{7,0.975}$ = – 2.3646 and the 95% confidence interval for $\mu_x - \mu_y$ is

$(\bar{x} - \bar{y}) \pm t_{7,.975} SE_{(\bar{x}-\bar{y})}$ or 95%CI: 15.0 ± 3.9 or 11.1 ≤ $\mu_x - \mu_y$ ≤ 18.9 .

14.7 **a.** The standard error is

$$SE_{(\bar{X}-\bar{Y})} = \sqrt{\frac{6.4^2}{10}+\frac{5.0^2}{25}} = 2.2574$$

b. The smaller sample size is 10 so a conservative df = 10 – 1 = 9 and $t_{9,.95}$ = 1.8331

The test statistic is $t_s = \frac{\bar{x}-\bar{y}}{SE_{(\bar{x}-\bar{y})}}$ = 1.772

The null hypothesis is H_0: $\mu_x - \mu_y \le 0$ and the alternative hypothesis is H_a: $\mu_x - \mu_y > 0$

Reject H_0: if $T > t_{9,.05}$. Conclusion: Reject H_0:.

Alternatively, compute the P-value = $Pr(T_9 > t_s) = Pr(T_9 \ge 1.772) = 0.055 > \alpha = 0.05$.

There is insufficient evidence to conclude $\mu_x - \mu_y > 0$ at the 5% significance level.

14.9 **a.** H_0: $\mu_x - \mu_y = 0$; No difference in mean salaries.

b. H_a: $\mu_x - \mu_y \neq 0$; There is a difference in mean salaries.

c. $SE_{(\bar{x}-\bar{y})} = \sqrt{\frac{s_x^2}{n_x}+\frac{s_y^2}{n_y}} \approx 166.67$. Test statistic: $t_s = \frac{\bar{x}-\bar{y}}{SE_{(\bar{x}-\bar{y})}}$

Test statistic $t_s = (26240 - 25930)/166.67 = 1.86$

Rejection region: $|T| > t_{df,.975} \approx 1.99$, df = min($n_x$, n_y) – 1 (Conservative method).

Alternatively, P-value = $Pr(|T_{df}| \ge t_s) = Pr(|T| \ge 1.86) \approx 0.067 > \alpha$

d. Decision: There is insufficient evidence (test statistic is not in rejection region) to conclude that there is a difference between the average salary of union and nonunion workers.

14.11 Here $n_x = 46,\ \bar{x} = 17.8\%,\ s_x = 3.06\%,\ n_y = 46,\ \bar{y} = 15.1\%,\ s_y = 1.89\%$

H_0: $\mu_x - \mu_y = 0$ H_a: $\mu_x - \mu_y > 0$,

$SE_{(\bar{x}-\bar{y})} = \sqrt{\frac{s_x^2}{n_x}+\frac{s_y^2}{n_y}} = 0.53029$, so test statistic $t_s = \frac{\bar{x}-\bar{y}}{SE_{(\bar{x}-\bar{y})}} = 5.09152$

Rejection Region: $t_s > t_{45,.99} = 2.4121$

Conclusion: Since 5.0915 >2.4121, reject the null hypothesis. There is sufficient evidence to decide there is a difference in tip rates.

Alternatively, compute the P-value = $Pr(T_{45} \ge t_s) = Pr(T_{45} \ge 5.0915) \approx 3\cdot 10^{-6}$.

Since the P-value is less than α = 0.01, you reach the same conclusion.

14.13 **a.** Here $\frac{s_x^2}{n_x} = 4.7^2/25 = 0.88360$ and $\frac{s_y^2}{n_y} = 2.3^2/45 = 0.11756$

Welch's formula: $df = \dfrac{(0.88360 + 0.11756)^2}{\dfrac{.883660^2}{24} + \dfrac{.11756^2}{44}} = 30$

b. The 95% confidence interval for $\mu_x - \mu_y$ is of the form $(\bar{x} - \bar{y}) \pm t_{df,.975} SE_{(\bar{x}-\bar{y})}$

95%C.I.: $-4 \pm 2.0423\sqrt{0.88360 + 0.11756} = -4 \pm 2.0435$ or

95%CI: $-6.044 \leq \mu_x - \mu_y \leq -1.956$

14.15 **a.** Here $SE_{\bar{X}}^2 = \frac{s_x^2}{n_x} = 6.4^2/10 = 4.096$ and $SE_{\bar{Y}}^2 = \frac{s_y^2}{n_y} = 5^2/25 = 1.000$ and

$SE_{\bar{X}-\bar{Y}}^2 = SE_{\bar{X}}^2 + SE_{\bar{Y}}^2 = 4.096 + 1 = 5.096$

Welch's formula: $df = \dfrac{SE_{\bar{X}-\bar{Y}}^4}{\dfrac{SE_{\bar{X}}^4}{n_x - 1} + \dfrac{SE_{\bar{Y}}^4}{n_y - 1}} = \dfrac{(5.096)^2}{\dfrac{4.096^2}{9} + \dfrac{1^2}{24}} = 13.6$, rounded down to df = 13.

b. Test statistic $t_s = \dfrac{\bar{x} - \bar{y}}{SE_{(\bar{x}-\bar{y})}} = 1.7719.$

P-value = $Pr(T_{13} \geq t_s) = Pr(T_{13} \geq 1.7719) = 0.049$. Since the P-value is less than the significance level $\alpha = 0.05$, reject Ho in favor of the hypothesis H_a: $\mu_x - \mu_y > 0$

14.17 **a.** Here $SE_{\bar{X}}^2 = \frac{s_x^2}{n_x} = 119^2/28 = 505.75$ and $SE_{\bar{Y}}^2 = \frac{s_y^2}{n_y} = 87^2/28 = 270.32$ and

$SE_{\bar{X}-\bar{Y}} = \sqrt{SE_{\bar{X}}^2 + SE_{\bar{Y}}^2} = 27.858$. Welch's formula:

$$df = \dfrac{SE_{\bar{X}-\bar{Y}}^4}{\dfrac{SE_{\bar{X}}^4}{n_x - 1} + \dfrac{SE_{\bar{Y}}^4}{n_y - 1}} = \dfrac{602286.86}{\dfrac{505.75^2}{27} + \dfrac{270.32^2}{27}} = 49.5$$, rounded down to 49.

95% CI: $(\bar{x} - \bar{y}) \pm t_{49,.975} SE_{(\bar{x}-\bar{y})} = 29 \pm 2.0096(27.858) = 29 \pm 55.98$ or

95%CI: $-27 \leq \mu_x - \mu_y \leq 85$

b. Using the conservative df = 27, 95% CI: $(\bar{x} - \bar{y}) \pm t_{27,.975} SE_{(\bar{x}-\bar{y})} =$

29 ±2.0518(27.858) = 29 ± 57.16 or $-28.16 \leq \mu_x - \mu_y \leq 86.16$

14.19 Let control = x and treatment = y: Here

$n_x = 3,\ \bar{x} = 15,\ s_x = 3.50555,\ n_y = 5,\ \bar{y} = 20,\ s_y = 3.08221$

a. The 90% confidence interval is of the form $(\bar{x} - \bar{y}) \pm t_{df,.95} SE_{(\bar{x}-\bar{y})}$

$SE^2_{\bar{X}} = 4.33 \qquad SE^2_{\bar{Y}} = 1.9 \qquad SE^2_{\bar{X}-\bar{Y}} = 6.2333$

Welch's formula $df = \dfrac{(6.2333)^2}{\dfrac{4.33^2}{2} + \dfrac{1.9^2}{4}} = 3.87$ rounded down to 3.

$t_{3,.95}$ = 2.3534 so the 90% CI is $(\bar{x} - \bar{y}) \pm t_{3,.95} SE_{(\bar{x}-\bar{y})}$ = –5 ± 2.3534(2.2411) = –5 ± 5.87

or 90%CI: $-10.4 \le \mu_x - \mu_y \le 0.4$

b. The test statistic is $t_s = \dfrac{\bar{x} - \bar{y}}{SE_{(\bar{x}-\bar{y})}}$ = 5/2.4967 = 2.00267

P-value = $Pr(T_3 \ge t_s) = Pr(T_3 \ge 2.00267) = 0.0695$

14.21 **a.** Similar to 14.20a.

$$s_p = \sqrt{\frac{(n_x - 1)S_x^2 + (n_y - 1)S_y^2}{(n_x + n_y - 2)}} = \sqrt{\frac{5 \cdot 2.7^2 + 7 \cdot 2.3^2}{12}} = \sqrt{\frac{73.48}{12}} = 2.4745$$

b. The 95% confidence interval for $\mu_x - \mu_y$ is $(\bar{x} - \bar{y}) \pm t_{12,.975} s_p \sqrt{\dfrac{1}{n_x} + \dfrac{1}{n_y}} =$

(16 – 20) ± 2.11788(2.4745)(0.54006) = –4 ± 2.9 or $-6.9 \le \mu_x - \mu_y \le -1.1$

14.23 **a.** The pooled standard deviation is

$$s_p = \sqrt{\frac{(n_x - 1)S_x^2 + (n_y - 1)S_y^2}{(n_x - 1) + (n_y - 1)}} = \sqrt{\frac{862.44}{33}} = 5.1122$$

b. The test statistic is $t_s = \dfrac{\bar{x} - \bar{y}}{SE_{(\bar{x}-\bar{y})}}$ = 4/1.9128 = 2.0912. To test the one-sided alternative hypothesis of $H_a: \mu_x - \mu_y > 0$ use the P-value = $Pr(T_{33} > 2.0912) = 0.0222$.

Since P-value < significance level = α = 0.05, reject the null hypothesis in favor of the alternative hypothesis.

14.25 Here $n_x = 25,\ \bar{x} = 129.87,\ s_x = 15.32,\ n_y = 30,\ \bar{y} = 119.71,\ s_y = 18.23$.

The pooled estimate of the common standard deviation is

$$s_p = \sqrt{\frac{24 \cdot 15.32^2 + 29 \cdot 18.23^2}{24 + 29}} = 16.9742$$

and the standard error is

$$SE_{(\bar{x}-\bar{y})} = s_p \sqrt{\frac{1}{n_x} + \frac{1}{n_y}} = 4.5966.$$

The decision rule is: Retain H_0: if $|t_s| \le t_{53,.995}$ and reject H_o if $|t_s| > t_{53,.995}$.

Here $t_s = \dfrac{\bar{x} - \bar{y}}{SE_{(\bar{x}-\bar{y})}}$ = 10.16/4.5966 = 2.2103 and $t_{53,.995}$. = *2.6718.*

So the null hypothesis is retained at the 1% significance level.

The P-value is $Pr(|T_{53}| \ge t_s) = 0.03$.

Conclude that there is no significant difference in the mean systolic blood pressure of these groups (at the 1% significance level).

14.27 Company A = x and company B = y:

$$n_x = 10,\ \bar{x} = 35.3,\ s_x^2 = 5.8,\ n_y = 9,\ \bar{y} = 31.7,\ s_y^2 = 11.6$$

A 95% confidence interval of the true difference in means is $(\bar{x} - \bar{y}) \pm t_{17,.975} SE_{(\bar{x}-\bar{y})}$.

Here $t_{17,.975}$= *2.1098* and $SE_{(\bar{x}-\bar{y})} = s_p \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}$ = 1.3419 with

$$s_p = \sqrt{\frac{(n_x - 1)S_x^2 + (n_y - 1)S_y^2}{(n_x - 1) + (n_y - 1)}} = 2.9205$$

So the 95% confidence interval bounds are 3.6 $\pm$ 2.831 screws per container.

14.29 For football: $n_1 = 93,\ k_1 = 72,\ p_1 = 72/93 = 0.7742$

For basketball: $n_2 = 189,\ k_2 = 150,\ p_2 = 150/189 = 0.7937$

The null and alternative hypothesis are H_0: $\pi_1 - \pi_2 = 0$ and H_0: $\pi_1 - \pi_2 \neq 0$ with $\alpha = 5\%$.
The pooled estimate of the common success rate under the null hypothesis H_o is

$$p = \frac{k_1 + k_2}{n_1 + n_2} = \frac{222}{282} = 0.78723 \quad \text{and} \quad s_p = \sqrt{p(1-p)} = 0.40926.$$

So the standard error of $p_1 - p_2$ as an estimate of $\pi_1 - \pi_2$ is

$$SE_{(p_1 - p_2)} = s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 0.05184$$

and the test statistic is

$$z_s = \frac{(p_1 - p_2)}{SE_{(p_1 - p_2)}} = \frac{-0.01946}{0.05184} \approx -0.375$$

The decision rule is: Retain H_o if $|z_s| \leq z_{0.975} = 1.96$ and reject H_o if $|z_s| > z_{0.975} = 1.96$.
So the null hypothesis is retained at the 5% significance level. The evidence does not suggest that there is a difference in success rates for these sports (where success is defined as winning the game when leading after 3 quarters).

Another approach is to compute a P-value based on the standard normal curve:

$$\Pr(|Z| > |z_s|) = \Pr(|Z| > 0.3753) = 2\Pr(z > 0.3753) \approx 0.70 > 0.05$$

Since the P-value exceeds the significance level, the null hypothesis is retained.

14.31 For the new drug: $n_1 = 42,\ k_1 = 38,\ p_1 = 38/42 = 0.9048$

For the standard: $n_2 = 21,\ k_2 = 14,\ p_2 = 14/21 = 0.6667$

The null and alternative hypothesis are H_0: $\pi_1 - \pi_2 = 0$ and H_0: $\pi_1 - \pi_2 > 0$ with $\alpha = 1\%$.
The pooled estimate of the common success rate under the null hypothesis H_o is

$$p = \frac{k_1 + k_2}{n_1 + n_2} = \frac{52}{63} \quad \text{and} \quad s_p = \sqrt{\frac{52}{63} \cdot \frac{11}{63}} = 0.37963$$

So the standard error of $p_1 - p_2$ as an estimate of $\pi_1 - \pi_2$ is

$$SE_{(p_1 - p_2)} = s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 0.10146$$

and the test statistic is

$$z_s = \frac{(p_1 - p_2)}{SE_{(p_1 - p_2)}} \approx 2.3467$$

The decision rule is: Retain H_o if Retain H_o if $z_s \leq z_{0.99} = 2.3264$ and reject H_o if $z_s > z_{0.99} = 2.3264$. Since $z_s = 2.3467 > z_{0.99} = 2.3264$, the null hypothesis is rejected.
Also, the P-value = Pr(Z > 2.3467) = 0.009 < α.
The evidence supports the conclusion (at the 1% significance level) that the new drug does out-perform the standard treatment.

Comment. This is a large sample method, but one of the samples sizes is small. So we do not recommend actually this method here but recommend using the Fisher's exact method (where you would obtain a P-value ≈ 0.025). To see why n may be too small, refer to the general rule for approximating a binomial random variable by a normal random variable in Section 8.3. Note $n_1(1 - p) < 5$ so the normal approximation is not considered good here.

14.33 **a.** 95% CI:

$$\mu_X - \mu_Y = (\bar{x} - \bar{y}) \pm t_{4,0.95} SE_{(\bar{x}-\bar{y})} = (70 - 60) \pm 2.13185\sqrt{\frac{100}{5} + \frac{64}{12}} = 10.00 \pm 10.73$$

b. 95% CI:

$$\mu_X - \mu_Y = (\bar{x} - \bar{y}) \pm t_{6,0.95} SE_{(\bar{x}-\bar{y})} = (70 - 60) \pm 1.94318\sqrt{\frac{100}{5} + \frac{64}{12}} = 10.00 \pm 9.78\,.$$

Here df = $\left(\frac{100}{5} + \frac{64}{12}\right)^2 \left(\frac{\left(\frac{100}{5}\right)^2}{4} + \frac{\left(\frac{64}{12}\right)^2}{11} \right)^{-1} \approx 6.3$, which is rounded down to df = 6.

14.35 $H_0 : \pi_s - \pi_{ns} = 0,\ H_a : \pi_s - \pi_{ns} > 0.\ p_s = \frac{12}{15} = 0.8000,\ p_{ns} = \frac{15}{33} = 0.4545.$

The test statistic is $z_s = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{SE_{(p_1 - p_2)}} = \frac{19/55}{0.15448} = 2.23626$.

Under the null hypothesis Z_s would be approximately standard normal if the sample sizes were large, so the P-value would be: $P - value = \Pr(Z_s > 2.23626) = 0.013$. This P-value is close to 1%, so it is somewhat unreasonable to attribute the observed difference in flu rates to chance variation and we would have evidence that the flu shot was effective. Note: the normal approximation may be inaccurate for small samples, however the Fisher's exact test P-value is 0.0254 and our conclusions would have been roughly the same if we had used this test.

14.37 $H_0: \pi_{CSP} - \pi_{TES} = 0, H_a: \pi_{CSP} - \pi_{TES} \neq 0.$ $p_{CSP} = \frac{34}{56} = 0.60714, p_c = \frac{25}{30} = 0.83\overline{3}.$

The test statistic is $z_s = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{SE_{(p_1 - p_2)}} = \frac{\frac{34}{56} - \frac{25}{30}}{0.10500} = \text{-}2.15412.$

Under the null hypothesis Z_s would be approximately standard normal, so the P-value would be: $P - value = \Pr(|Z_s| > 2.15412) = 0.03123$. This P-value is less than 5% so the null hypothesis is rejected. Note: the normal approximation may be inaccurate for small samples. In this case the success rates are quite high, so the sample sizes may be considered to be small. If we had used the Fisher's exact test, the P-value is 0.0499 we would have rejected the null hypotheses and concluded the underlying rates of identifying programmed guilty subjects were different.

14.39 95% CI:

$$\mu_{After} - \mu_{Before} = (\bar{y} - \bar{x}) \pm t_{7,0.975} SE_{(\bar{y} - \bar{x}))} = (550 - 530) \pm 2.3646\sqrt{\frac{4900}{5} + \frac{2750}{5}} =$$

$$20 \pm 92.49 \;\; or \;\; -72.5 \leq \mu_{After} - \mu_{Before} \leq 112.5$$

Here df = $\left(\frac{4900}{5} + \frac{2750}{5}\right)^2 \left(\frac{\left(\frac{4900}{5}\right)^2}{4} + \frac{\left(\frac{2750}{5}\right)^2}{4} \right)^{-1} \approx 7.414.$

Note: This interval is too wide because the positive correlation between before and after scores within students was ignored.

14.43 **a.** Use the result from Exercise 9.46. Let $f_Z(z)$ denote the standard normal density. Then

$$f_Q(q) = -\frac{f_Z(z_q)}{Q'(x_q)} = \frac{f_Z(z_q)}{f_Z(z_q)} = 1.$$

b. Let $f_Z(z)$ denote the standard normal density and let $f_Z(z|1)$ let denote the normal density with mean 1 and standard deviation 1. Then $f_Q(q) = -\frac{f_Z(z_q)}{Q'(x_q)} = \frac{f_Z(z_q)}{f_Z(z_q|1)} = e^{(z_q - 0.5)}.$

Here z_q denotes the qth quantile of a standard normal random variable.