CHAPTER 8
Comparison of Paired Samples

8.1.1 (a) The 16 possible mean differences are {-0.675, -0.62, -0.305, -0.285, -0.25, -0.23, -0.14, -0.085, 0.085, 0.14, 0.23, 0.25, 0.285, 0.305, 0.62, 0.675}

(b) Only 2 of the 16 means are as extreme as 0.675.

(c) The P-value is 2/16 = 0.1250, which is not less than $\alpha = 0.10$. There is no significant evidence that progesterone affects cAMP.

• **8.2.1 (a)** The standard deviation of the four sample differences is given as 0.68. The standard error is

$$SE_{\bar{D}} = \frac{s_D}{\sqrt{n_D}} = \frac{0.68}{\sqrt{4}} = 0.34.$$

(b) H_0: The mean yields of the two varieties are the same ($\mu_1 = \mu_2$)
H_A: The mean yields of the two varieties are different ($\mu_1 \neq \mu_2$)

$t_s = -1.65/0.34 = -4.85$. With df = 3, Table 4 gives $t_{0.01} = 4.541$ and $t_{0.005} = 5.841$; thus, 0.01 < P-value < 0.02. At significance level $\alpha = 0.05$, we reject H_0 if P < 0.05. Since 0.01 < P < 0.02, we reject H_0. There is sufficient evidence (0.01 < P < 0.02) to conclude that Variety 2 has a higher mean yield than Variety 1.

(c) H_0: The mean yields of the two varieties are the same ($\mu_1 = \mu_2$)
H_A: The mean yields of the two varieties are different ($\mu_1 \neq \mu_2$)

$$SE_{(\bar{y}_1-\bar{y}_2)} = \sqrt{\frac{1.76^2}{4} + \frac{1.72^2}{4}} = 1.230.$$

$t_s = -1.65/1.230 = -1.34$. With df = 6, Table 4 gives $t_{0.20} = .906$ and $t_{0.10} = 1.440$. Thus, 0.20 < P-value < 0.40 and we do not reject H_0. There is insufficient evidence (0.20 < P < 0.40) to conclude that the mean yields of the two varieties are different. (By contrast, the correct test, in part (b), resulted in rejection of H_0.)

8.2.2 (a) The standard deviation of the nine sample differences is given as 59.3. The standard error is

$$SE_{\bar{D}} = \frac{s_D}{\sqrt{n_D}} = \frac{59.3}{\sqrt{9}} = 19.77.$$

(b) H_0: The mean weight gains on the two diets are the same ($\mu_1 = \mu_2$)
H_A: The mean weight gains on the two diets are different ($\mu_1 \neq \mu_2$)

$t_s = 22.9/19.77 = 1.158$. With df = 8, Table 4 gives $t_{0.20} = 0.889$ and $t_{0.10} = 1.397$. Thus, 0.20 < P-value < 0.40 and we do not reject H_0. There is insufficient evidence (0.20 < P < 0.40) to conclude that the mean weight gains on the two diets are different.

(c) $22.9 \pm (1.860)(19.77)$

(-13.9,59.7) or $-13.9 < \mu_D < 59.7$ lb.

(d) We are 90% confident that the average steer gains somewhere between 59.7 pounds more and 13.9 pounds less when on Diet 1 than when on Diet 2 (in a 140-day period).

• **8.2.3** Let 1 denote control and let 2 denote progesterone.

H_0: Progesterone has no effect on cAMP ($\mu_1 = \mu_2$)
H_A: Progesterone has some effect on cAMP ($\mu_1 \neq \mu_2$)

The standard error is

$$SE_{\bar{D}} = \frac{s_D}{\sqrt{n_D}} = \frac{0.40}{\sqrt{4}} = 0.20.$$

The test statistic is

$$t_s = \frac{\bar{d}}{SE_{\bar{D}}} = \frac{0.68}{0.20} = 3.4.$$

To bracket the P-value, we consult Table 4 with df = 4 - 1 = 3. Table 4 gives $t_{0.025} = 3.182$ and $t_{0.02} = 3.482$. Thus, the P-value is bracketed as

0.04 < P-value < 0.05.

At significance level $\alpha = 0.10$, we reject H_0 if P < 0.10. Since 0.04 < P < 0.05, we reject H_0. There is sufficient evidence (0.04 < P <0 .05) to conclude that progesterone decreases cAMP under these conditions.

8.2.4 (a) $t_s = 4.17$

(b) H_0: Weight change when taking mCPP is no different than when taking a placebo. H_A: Weight change when taking mCPP is different than when taking a placebo.

(c) We reject H_0 because the P-value is smaller than 0.10. We have strong evidence to say that weight change when taking mCPP is greater than when taking a placebo.

8.2.5 (a) $1.00 \pm (3.355)(0.24)$

(0.19, 1.81) or $0.19 < \mu_D < 1.81$ lb.

(b) We are 99% confident that mean weight loss on mCPP is between 0.19 and 1.81 lb greater than when on placebo.

• **8.2.6 (a)** Let 1 denote treated side and 2 denote control side. The standard error is

$$SE_{\bar{D}} = \frac{s_D}{\sqrt{n_D}} = \frac{1.118}{\sqrt{15}} = 0.2887.$$

The critical value $t_{0.025}$ is found from Student's t distribution with df = n_D - 1 = 15 - 1 = 14. From Table 4 we find that $t_{14,0.025} = 2.145$.

The 95% confidence interval is

$$\bar{d} \pm t_{0.025} SE_{\bar{D}}$$

$0.117 \pm (2.145)(0.2887)$

(-0.50,0.74) or $-0.50 < \mu_1 - \mu_2 < 0.74$ °C.

(b) $SE_{(\bar{Y}_1-\bar{Y}_2)} = \sqrt{\frac{1.217^2}{15} + \frac{1.302^2}{15}} = 0.460.$

$0.117 \pm (2.048)(0.460)$ (using df = 28)

(-0.83,1.06) or $-0.83 < \mu_1 - \mu_2 < 1.06$ °C.

This interval is wider than the one obtained in part (a).

8.2.7 Let 1 denote treated side and 2 denote control side.
H_0: The electrical treatment has no effect on collagen shrinkage temperature ($\mu_1 = \mu_2$)
H_A: The electrical treatment tends to reduce collagen shrinkage temperature ($\mu_1 < \mu_2$)

We note that $\bar{y}_1 > \bar{y}_2$, so the data do not deviate from H_0 in the direction specified by H_A. Thus, P > 0.50 and we do not reject H_0. There is no evidence (P > 0.50) that the electrical treatment tends to reduce collagen shrinkage temperature under these conditions.

8.2.8 The data provide fairly strong evidence (P = 0.03) that desipramine is more effective than clomipramine in reducing the compulsion to pull one's hair.

8.2.9 $SE_{\bar{D}} = \frac{s_D}{\sqrt{n_D}} = \frac{3}{\sqrt{28}} = 0.57$. The confidence interval is $10.9 \pm (2.052)(0.57)$ or (9.7,12.1).

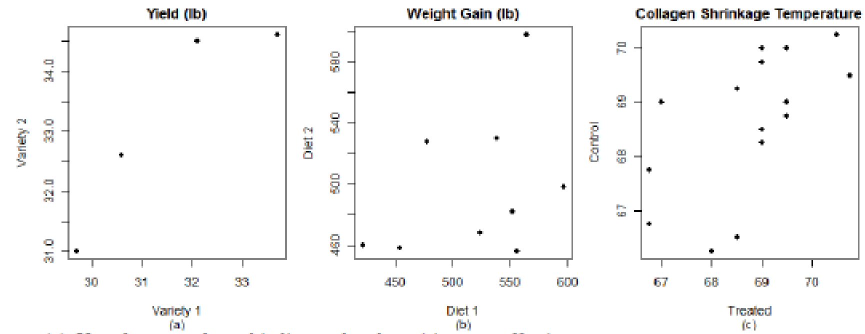
8.2.10 With the outliers deleted, the mean of the remaining 26 differences is 11.0 and the standard deviation is 2.1. $SE_{\bar{D}} = \frac{s_D}{\sqrt{n_D}} = \frac{2.1}{\sqrt{26}} = 0.41$. The confidence interval is $11.0 \pm (2.060)(0.41)$ or (10.1,11.8). This interval is more narrow than the previous interval that was based on all of the data, including the outliers, but the difference is not great.

8.2.11 There is no single correct answer. Any data set with Y_1 and Y_2 varying, but d not varying, is correct; for example:

Y_1	Y_2	d
5	3	2
6	4	2
3	1	2
4	2	2
5	3	2

8.3.1 – 8.3.3 See Section III of this Manual.

8.3.4 The following are plots for parts (a-c).



(a) Yes, the upward trend indicates that the pairing was effective.

(b) The upward trend here is rather weak, which indicates that the pairing was not especially effective.

(c) Yes, the upward trend indicates that the pairing was effective.

8.3.5 (a) Each of the subjects could be subjected to both the warm and the cold room and the pulse could be recorded under both conditions. To avoid bias, the 20 subjects could be randomly split so that ten would be subject to the cold room first while the other ten would be subject to the warm room first.

(b) The paired design would eliminate between subject variability in heart rates. Consequently there will be more power to detect differences due to temperature.

• **8.4.1** (a) $B_s = 6$. Looking under $n_D = 9$ in Table 7, we see that there is no entry less than or equal to 6. Therefore, P > 0.20.

(b) $B_s = 7$. Looking under $n_D = 9$ in Table 7, the P-value is 0.180.

(c) $B_s = 8$. Looking under $n_D = 9$ in Table 7, the P-value is 0.039.