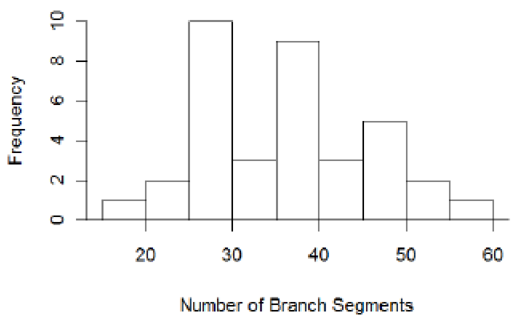
• **6.5.1** The fact that the mean is less than the SD casts doubt on the condition that the population is normal, for the following reason. In a normal population, about 15% of the observations fall more than one SD below the mean, whereas this sample cannot have any observations that far below the mean because $\bar{y} - s$ is negative and the observed variable (serum SGOT) cannot be negative.

6.5.2 **(a)** There were 36 cells, but only seven guinea pigs, so there is a hierarchical structure in the data, which suggests that the observations are not independent.

(b)

Number of branch segments	Frequency
15-19	1
20-24	2
25-29	10
30-34	3
35-39	9
40-44	3
45-49	5
50-54	2
55-59	1
Total	36



The distribution has two or perhaps three modes, which may reflect the hierarchical structure in the data (that is, different modes may represent different animals or groups of animals.)

6.5.3 The outlier (1,060) suggests that the population distribution is not normal but rather is skewed to the right or long-tailed. Because the sample size is small, Student's t method is not appropriate if the population is not normal.

6.5.4 **(a)** No. The observed frequency distribution is highly skewed, which suggests that the population distribution is skewed.

(b) Even if the population distribution is not normal, Student's t method is approximately valid if the sample size is large. In this case, the sample size is n = 242, which is quite large.

6.5.5 **(a)** $\bar{y} = 5.68$; $s = 1.54$; $n = 9$.

The 90% confidence interval for μ is

$$5.68 \pm 1.860\left(\frac{1.54}{\sqrt{9}}\right)$$

$$(4.73, 6.63) \text{ or } 4.73 < \mu < 6.63\sqrt{\text{cm}}.$$

(b) We are 90% confident that the average square root of the diameters of all American Sycamore trees in the population is between 4.73 and $6.63\sqrt{\text{cm}}$.

• **6.5.6** In this experiment the treatments have been assigned to the flasks, so the entire flasks are the independent units. Though randomly drawn from within each flask, the three aliquots are not independent from one another in the larger context of this study: the aliquots from the same flask will be more similar to each other than aliquots from different flasks. Due to this hierarchical sampling, the SE as described would be an underestimate of the true SE. The SE as computed violates Condition 1(b). (One correct analysis would be to compute the mean density of the three aliquots from each flask, and then use these values as the independent measurements for analysis)

6.5.7 In this experiment the independently observed units are the entire plants. The oil measurements from the seeds within each plant will likely be more similar than measurements across plants. Due to this hierarchical sampling (and lack of independence in the observations), the SE as described would be an underestimate of the true SE. The SE as computed violates Condition 1(b). (One correct analysis would be to compute the mean oil measurements for the five seeds from each plant, and then use these values as the independent measurements for analysis)

6.5.8 Nearly 12% [= 100 × (30/255)%] of the entire population has been sampled. For our SE calculations to be valid, the population must be large relative to the sample size. For our SE formula to be valid (without making special adjustments), we should sample no more than 5% of the population.

• **6.6.1** We first find the standard error of each mean.

$$SE_1 = SE_{\bar{y}_1} = \frac{s_1}{\sqrt{n_1}} = \frac{4.3}{\sqrt{6}} = 1.755.$$

$$SE_2 = SE_{\bar{y}_2} = \frac{s_2}{\sqrt{n_2}} = \frac{5.7}{\sqrt{12}} = 1.645.$$

$$SE_{(\bar{y}_1 - \bar{y}_2)} = \sqrt{SE_1^2 + SE_2^2} = \sqrt{1.755^2 + 1.645^2} = 2.41.$$

6.6.2 $SE_1 = \frac{44.2}{\sqrt{10}} = 13.977$; $SE_2 = \frac{28.7}{\sqrt{10}} = 9.076$.

$$\sqrt{13.977^2 + 9.076^2} = 16.7.$$

6.6.3 $SE_1 = \frac{6.5}{\sqrt{5}} = 2.907$; $SE_2 = \frac{8.4}{\sqrt{7}} = 3.175$.

$\sqrt{2.907^2 + 3.175^2} = 4.3$.

6.6.4 $SE_1 = \frac{6.5}{\sqrt{10}} = 2.055$; $SE_2 = \frac{8.4}{\sqrt{14}} = 2.245$.

$\sqrt{2.055^2 + 2.245^2} = 3.04$.

6.6.5 $\sqrt{3.7^2 + 4.6^2} = 5.90$.

6.6.6 $SE_{(\bar{Y}_1-\bar{Y}_2)} = \sqrt{SE_1^2+SE_2^2} = \sqrt{0.5^2 + 0.7^2} = 0.86$.

• **6.6.7** $SE_1 = \frac{2.4}{\sqrt{49}} = 0.343$; $SE_2 = \frac{2.0}{\sqrt{52}} = 0.277$.

$SE_{(\bar{Y}_1-\bar{Y}_2)} = \sqrt{SE_1^2+SE_2^2} = \sqrt{0.343^2 + 0.277^2} = 0.44$.

6.6.8 $SE_1 = \frac{0.400}{\sqrt{9}} = 0.133$; $SE_2 = \frac{0.220}{\sqrt{6}} = 0.090$.

$\sqrt{0.133^2 + 0.090^2} = 0.16$.

6.6.9 $\sqrt{5.5^2 + 8.6^2} = 10.2$.

• **6.7.1** Let 1 denote conventional treatment and let 2 denote Coblation.

$\bar{y}_1 = 4.3$; $SE_1 = 2.8 / \sqrt{49} = 0.4$.

$\bar{y}_2 = 1.9$; $SE_2 = 1.8 / \sqrt{52} = 0.25$.

The standard error of the difference is $SE_{(\bar{Y}_1-\bar{Y}_2)} = \sqrt{0.4^2 + 0.25^2} = 0.47$.

The critical value $t_{0.025}$ is determined from Student's t distribution with df = 81.1. Using df = 80 (the nearest value given in Table 4), we find that $t_{80,0.025} = 1.990$.

The 95% confidence interval is

$\bar{y}_1 - \bar{y}_2 \pm t_{0.025} SE_{(\bar{Y}_1-\bar{Y}_2)}$

$(4.3 - 1.9) \pm (1.990)(0.47)$.

So the confidence interval is (1.46,3.34) or $1.46 < \mu_1 - \mu_2 < 3.34$.

6.7.2 Let 1 denote dark and let 2 denote photoperiod.

$SE_{(\bar{Y}_1-\bar{Y}_2)} = \sqrt{\frac{13^2}{4} + \frac{13^2}{4}} = 9.192$.

(a) $(92 - 115) \pm (2.447)(9.192)$ (df = 6)

$(-45.5,-0.5)$ or $-45.5 < \mu_1 - \mu_2 < -0.5$ nmol/gm.

(b) $(92 - 115) \pm (1.943)(9.192)$ (df = 6)

$(-40.9,-5.1)$ or $-40.9 < \mu_1 - \mu_2 < -5.1$ nmol/gm.

6.7.3 0.5 nmol/gm

6.7.4 (a) Let 1 denote biofeedback and let 2 denote control.

$SE_{(\bar{Y}_1-\bar{Y}_2)} = \sqrt{1.34^2 + 1.30^2} = 1.867$.

$(13.8 - 4.0) \pm (1.977)(1.867)$ (using df = 140)

$(6.1,13.5)$ or $6.1 < \mu_1 - \mu_2 < 13.5$ mm Hg.

(b) We are 95% confident that the population mean reduction in systolic blood pressure for those who receive training for eight weeks (μ_1) is larger than that for others (μ_2) by an amount that might be as small as 6.1 mm Hg or as large as 13.5 mm Hg.

6.7.5 No. The confidence interval found in Exercise 6.7.3 is valid even if the distributions are not normal, because the sample sizes are large.

• **6.7.6** (a) Let 1 denote antibiotic and let 2 denote control.

$SE_{(\bar{Y}_1-\bar{Y}_2)} = \sqrt{\frac{10^2}{10} + \frac{8^2}{10}} = 4.050$.

$(25 - 23) \pm (1.740)(4.050)$ (using df = 17)

$(-5.0,9.0)$ or $-5 < \mu_1 - \mu_2 < 9$ sec.

(b) For the interval to be valid, the sampling distribution of $\bar{Y}_1 - \bar{Y}_2$ must be approximately normal. This will occur when either the data itself comes from a normally distributed population, or when the sample sizes are large (via the Central Limit Theorem). Since the sample sizes are not particularly large (10 and 10) in this case, we must trust that the data itself follows a normal distribution. (Note that such faith might not be wise. One should have a good reason for this belief, or else poor decisions based on the data are possible.)

(c) We are 90% confident that the population mean prothrombin time for rats treated with an antibiotic (μ_1) is smaller than that for control rats (μ_2) by an amount that might be as much as 5 seconds or is larger than that for control rats (μ_2) by an amount that might be as large as 9 seconds.

6.7.7 Let 1 denote control and let 2 denote Pargyline.

$$SE_{(\bar{Y}_1-\bar{Y}_2)} = \sqrt{\frac{5.4^2}{900} + \frac{11.7^2}{905}} = 0.4286.$$

(a) $(14.9 - 46.5) \pm (1.96)(.4286)$ (df = ∞)

(-32.4,-30.8) or $-32.4 < \mu_1 - \mu_2 < -30.8$ mg.

(b) $(14.9 - 46.5) \pm (2.576)(.4286)$ (df = ∞)

(-32.7,-30.5) or $-32.7 < \mu_1 - \mu_2 < -30.5$ mg.

6.7.8 (a) Let 1 denote successful and let 2 denote unsuccessful.

$$SE_{(\bar{Y}_1-\bar{Y}_2)} = \sqrt{\frac{0.283^2}{22} + \frac{0.262^2}{17}} = 0.08763.$$

$(8.498 - 8.440) \pm (2.021)(0.08763)$ (using df = 40)

(-0.12,0.24) or $-0.12 < \mu_1 - \mu_2 < 0.24$ mm.

(b) We are 95% confident that the population mean head width of all females who mate successfully (μ_1) is smaller than that for rejected females (μ_2) by an amount that might be as much as 0.12 mm or is larger than that for rejected females (μ_2) by an amount that might be as large as 0.24 mm.

(c) There is no compelling evidence that the mean head width for the successful and unsuccessful maters differs, as zero is contained in the confidence interval computed in part (a). That is, a difference in mean head width of zero (no difference) is consistent with this data.

6.7.9 We are 97.5% confident that the population mean drop in systolic blood pressure of adults placed on a diet rich in fruits and vegetables for eight weeks (μ_1) is larger than that for adults placed on a standard diet (μ_2) by an amount that might be as small as .9 mm Hg or as large as 4.7 mm Hg.

6.7.10 We are 97.5% confident that the population mean drop in diastolic blood pressure of adults placed on a diet rich in fruits and vegetables for eight weeks (μ_1) is smaller than that for adults placed on a standard diet (μ_2) by an amount that might be as much as 0.3 mm Hg or is larger than that for adults placed on a standard diet (μ_2) by an amount that might be as much as 2.4 mm Hg.

6.7.11 (a) Let 1 denote caffeine and let 2 denote decaf.

$$SE_{(\bar{Y}_1-\bar{Y}_2)} = \sqrt{3.7^2 + 3.4^2} = 5.02.$$

$(7.3 - 5.9) \pm (1.740)(5.02)$ (using df = 17)

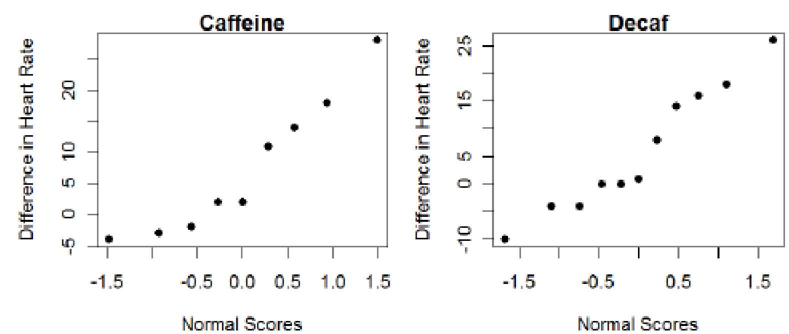
(-7.33,10.13) or $-7.33 < \mu_1 - \mu_2 < 10.13$.

(b) The interval contains zero, thus it is reasonable to believe that the mean difference in heart rates is zero, or that there is no difference between the caffeine and decaf groups.

(c) It is reasonable to believe that caffeine drinkers have a mean heart rate that is up to 7.33 beats/min slower than decaf drinkers. It is also reasonable to believe that caffeine drinkers have a mean heart rate that is up to 10.13 beats/min faster than decaf drinkers. This data is consistent with the belief that caffeine affects heart rate (yet, does not provide compelling evidence for this).

(d) Parts (b) and (c) appear to be contradictory, but they are not. Because the interval spans zero, no conclusive evidence for a difference (or lack of difference) is present. Note to students: Is the interval useless? Not entirely. We do have compelling evidence that if caffeine does affect heart rate, it doesn't lower it by more than 7.33 beats/min or raise it more than 10.13 beats/min with 95% confidence.

6.7.12 Normal probability plots support the normality condition:



6.7.13 Let 1 denote red and let 2 denote green.

$$SE_{(\bar{Y}_1-\bar{Y}_2)} = \sqrt{0.36^2 + 0.36^2} = 0.509.$$

$(8.36 - 8.94) \pm (2.021)(0.509)$ (using df = 40)

(-1.61,0.45) or $-1.61 < \mu_1 - \mu_2 < 0.45$.

6.7.14 No; the sample sizes are moderately large, so the sampling distribution of the difference in means is approximately normal, despite the skewness in the underlying data.

6.7.15 (a) The SE for the difference in means is $\sqrt{4.4^2 + 3.3^2} = 5.5$. Using 2 as the (approximate) t multiplier, the confidence interval is thus (151.7 – 125.6) ± 2*5.5 or 26.1 ± 11 or (15.1, 37.1).

(b) We are 95% confident that the average female Sumatran elephant foot circumference exceeds that of the average male by between 15.1 cm and 37.1 cm.

(c) Yes. The 95% confidence interval is well above zero so we have compelling evidence that the population means are not the same.

6.S.1 (a) $\bar{y} = 51.0$; $s = 3.195$; $SE = 3.195/\sqrt{4} = 1.597 \approx 1.6$.

(b) $51.0 \pm (3.182)(1.6)$

$(45.9, 56.1)$ or $45.9 < \mu < 56.1\%$.

(c) Wider. The higher the confidence level, the wider the interval must be.

• **6.S.2 (a)** $\bar{y} = 2.275$; $s = 0.238$. The standard error of the mean is

$$SE_{\bar{Y}} = \frac{s}{\sqrt{n}} = \frac{0.238}{\sqrt{8}} = 0.084 \text{ mm.}$$

(b) From Table 4 with $df = n - 1 = 7$, we find that $t_{0.025} = 2.365$.
The 95% confidence interval for μ is

$$\bar{y} \pm t_{0.025}\frac{s}{\sqrt{n}}$$

$2.275 \pm (2.365)(.084)$

$(2.08, 2.47)$ or $2.08 < \mu < 2.47$ mm.

(c) μ is the population mean stem diameter of plants of Tetrastichon wheat three weeks after flowering.

6.S.3 (a) The confidence interval formula is valid if (1) the data can be regarded as a random sample from a large population, (2) the observations are independent of one another, and (3) the population is normal.

(b) We have no information with which to check conditions (1) and (2). A normal probability plot, or a histogram, shows that the data look normal.

(c) The most important condition is (1).

• **6.S.4** We use the inequality

$$\frac{\text{Guessed SD}}{\sqrt{n}} \le \text{Desired SE.}$$

In this case the desired SE is 0.03 mm and the guessed SD (from Exercise 6.46) is 0.238 mm. Thus, the inequality is

$$\frac{0.238}{\sqrt{n}} \le 0.03 \text{ or } \frac{0.238}{0.03} \le \sqrt{n}\text{, so } 7.933^2 \le n\text{, which means that } n \ge 62.9.$$

Thus, the experiment should include 63 plants.

6.S.5 (a) $4.3 \pm (2.093)(2.03/\sqrt{20})$

$(3.35, 5.25)$ or $3.35 < \mu < 5.25$ puffs

(b) We are 95% confident that the average number of puffs for all fruit fly larva incubated at 37° C for 30 minutes is between 3.35 and 5.25.

(c) The horizontal banding in the normal probability plot is not surprising here because the data are discrete. The number of puffs can only take on non-negative integer values.

6.S.6 (a) $28.86 \pm (2.576)(4.24/\sqrt{1353})$

$(28.56, 29.16)$ or $28.56 < \mu < 29.16$ days.

(b) The confidence interval is not consistent with the hypothesis because 29.5 is not in the interval.

6.S.7 (a) The mean of all reported cycles is smaller because the women with shorter cycles had more cycles during the fixed time period, and therefore contributed more observations to the data.

(b) It would not be valid because the 5412 observations are not independent -- there is a hierarchical structure in the data.

6.S.8 (a) $5.1679 \pm (2.052)(0.1237)$

$(4.91, 5.42)$ or $4.91 < \mu < 5.42$ kg.

(b) $5.1679 \pm (2.771)(0.1237)$

$(4.83, 5.51)$ or $4.83 < \mu < 5.51$ kg.

(c) We are 95% confident that the average birthweight of all lambs that were born in April, were all the same breed, and were all single births, is between 4.91 kg and 5.42 kg.

(d) Since the stated research goal is to compare means of two populations, the SE should be reported. (If the researchers were interested in communicating or comparing the inherent variability of birthweight for the breeds, then the SD might be more appropriate).

• **6.S.9 (a)** We must be able to view the data as a random sample of observations from a large population, the observations in the sample must be independent of each other, and the population distribution must be approximately normal. (Note, however, that because the sample size (n = 28) is not very small, some non-normality of the population distribution would be acceptable.)

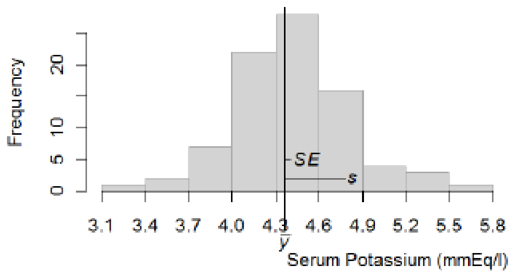
(b) The shape of the histogram is an estimate of the shape of the population distribution. Thus, the histogram can be used to check the normality condition of the population

(c) If twin births were included, the independence of the observations would be questionable, because birthweights of the members of a twin pair are likely to be dependent.

6.S.10 The confidence interval should be (6.2,7.4). The confidence interval is an interval estimate of the population mean. The data only take on integer values, but the mean of the population need not be an integer (and probably is not).

6.S.11 (a) $0.42 / \sqrt{84} = 0.04583$

(b)



(c) $4.36 \pm (1.984)(0.04583)$

$(4.269, 4.451)$ or $4.269 < \mu < 4.451$ mEq/l.

(d) We are 95% confident that the average serum potassium concentration in the blood of all healthy women is between 4.269 mEq/l and 4.451 mEq/l.

6.S.12 No. The confidence interval would be much too narrow; only a minority of healthy women would fall within the confidence interval. Instead, the interval $\bar{y} \pm 2\text{SD}$ would be a reasonable choice for reference limits.

6.S.13 (a) We would predict the SD of the new measurements to be about 0.42 mEq/l, because this is our best estimate (based on Exercise 6.54) of the population SD.

(b) $0.42 / \sqrt{200} = .030$ mEq/l.

6.S.14 (a) $\bar{y} = 62.767$; $s = 1.01127$; $SE = 1.01127 / \sqrt{6} = 0.4128 \approx 0.41$.

(b) $62.77 \pm (2.015)(0.41)$

$(61.94, 63.60)$ or $61.94 < \mu < 63.60\%$.

6.S.15 (a) $\bar{y} = 145.3$; $s = 12.87$; $SE = 12.87 / \sqrt{1139} = 0.381$.
The confidence interval is $145.3 \pm (1.96)(.381)$ or $(144.55, 146.05)$ or

$144.55 < \mu < 146.05$ g/l.

(b) No. The obtained 95% confidence interval is a confidence interval for the population mean hemoglobin level. It does not give limits for individual data points.

(c) No. See the answer to part (b).

6.S.16 (a) Assuming we had public safety in mind, we would want to be confident that the mean count was below 225 MPN/100ml, thus we would want an upper bound for the interval.

(b) For these data, $\bar{y} = 242.5$; $s = 42.5$; $SE = 42.5 / \sqrt{16} = 10.6$.
The one-sided (upper bound) 95% confidence interval is $242.5 + (1.753)(10.6)$ or $\mu < 261.1$ MPN/100ml.

(c) Although we are confident that the mean fecal count is below 261.1 MPN/100ml, there is not enough evidence to be confident that the count is below 225 MPN/100ml, the safe level. We do not have compelling evidence that the water is safe.

6.S.17 We can estimate the SD as being approximately 8. (Note that $(110 - 80)/4 = 7.5 \approx 8$.) It follows that $SE \approx 8 / \sqrt{38} \approx 1.3$. The t multiplier is roughly 2. Thus, (iii) is an approximate 95% confidence interval for the population mean blood pressure.

6.S.18 (a) Using the graph, we visually estimate $s \approx \frac{110-80}{4} \approx 8$. Our desired margin of error for 95% confidence is 2.0 mmHg, thus our desired SE is approximately $2.0/2 = 1.0$ mmHg. To find the sample size, we solve the following equation for n:

$$\frac{8}{\sqrt{n}} \le 1.0$$

which yields $n \ge 64$.

(b) Because 64 is more than 10% of 500 (the population size), the finite population correction factor should be applied, resulting in an SE of $\frac{s}{\sqrt{n}} \times \sqrt{\frac{500-64}{500-1}}$ rather than $\frac{s}{\sqrt{n}}$. That is, the usual SE is multiplied by 0.93, making the true SE somewhat smaller than the usual SE. Thus, an interval calculated from the usual SE will be somewhat too wide.

6.S.19 The 85 pups have come from only 10 mothers. The pups are not independent observations.

6.S.20 (a) $\bar{y} = 306$; $s = 51$; $SE = 51 / \sqrt{13} = 14.1$.

The confidence interval is $306 \pm (2.179)(14.1)$ or $(275, 337)$ or

$275 < \mu < 337$ calories.

(b) We are 95% confident that the mean calorie content is at least 275 calories, which exceeds the 252 calories listed on the package. There is compelling evidence that the value listed on the package is wrong.

(c) If we assume that the number of calories is normally distributed with mean equal to 252 calories, then ½ of the packages would have fewer calories than advertised (since the mean and median are equal for normally distributed data). Thus it makes sense that the manufacturer would desire the mean number of calories to exceed the number reported on the label.

6.S.21 **(a)** False. The confidence interval includes zero, so we are not confident that μ_1 and μ_2 are different.

(b) True. This is what a confidence interval tells us.

(c) False. The confidence interval is used to make an inference about the difference between μ_1 and μ_2; it does not tell us about individual data points (such as the length of hospitalization for a nitric oxide infant).

6.S.22 **(a)** False. We know that $\bar{y}_1 - \bar{y}_2$ is 6.9. The CI is for $\mu_1 - \mu_2$.

(b) False. This would only be true if we knew that $\mu_1 - \mu_2$ were exactly 6.9; but in that case we wouldn't be constructing a confidence interval to estimate $\mu_1 - \mu_2$.

(c) True. The confidence interval includes zero, so we would retain the null hypothesis of "no effect" if we conducted a hypothesis test.

6.S.23 In order to be 100% confident that our interval includes the parameter we need to include all real numbers in the interval.