Chapter 1

# A Review of Basic Concepts (Optional)

1.1 a. High school GPA is a number usually between 0.0 and 4.0. Therefore, it is quantitative.

b. Country of citizenship: USA, Japan, etc is qualitative

c. The scores on the SAT's are numbers between 200 and 800. Therefore, it is quantitative.

d. Gender is either male or female. Therefore, it is qualitative.

e. Parent's income is a number: $25,000, $45,000, etc. Therefore, it is quantitative.

f. Age is a number: 17, 18, etc. Therefore, it is quantitative.

1.3 a. The variable of interest is earthquakes.

b. Type of ground motion is qualitative since the three motions are not on a numerical scale. Earthquake magnitude and peak ground acceleration are quantitative. Each of these variables are measured on a numerical scale.

1.5 a. Town where sample collected is qualitative since this variable is not measured on a numerical scale.

b. Type of water supply is qualitative since this variable is not measured on a numerical scale.

c. Acidic level is quantitative since this variable is measured on a numerical scale (pH level 1 to 14).

d. Turbidity level is quantitative since this variable is measured on a numerical scale

e. Temperature quantitative since this variable is measured on a numerical scale.

f. Number of fecal coliforms per 100 millimeters is quantitative since this variable is measured on a numerical scale.

g. Free chlorine-residual(milligrams per liter) is quantitative since this variable is measured on a numerical scale.

h. Presence of hydrogen sulphide (yes or no) is qualitative since this variable is not measured on a numerical scale.

1.7 a. The population of interest is all decision makers. The sample set is 155 volunteer students. Variables measured were the emotional state and whether to repair a very old car (yes or no).

b. Subjects in the guilty-state group are less likely to repair an old car.

1.9 a. The experimental units are the amateur boxers.

b. Massage or rest group are both qualitative; heart rate and blood lactate level are both quantitative.

c. There is no difference in the mean heart rates between the two groups of boxers (those receiving massage and those not receiving massage). Thus, massage did not affect the recovery rate of the boxers.

d. No. Only amateur boxers were used in the experiment. Thus, all inferences relate only to boxers.

1.11 a. The population is all adults in Tennessee. The sample is 575 study participants.

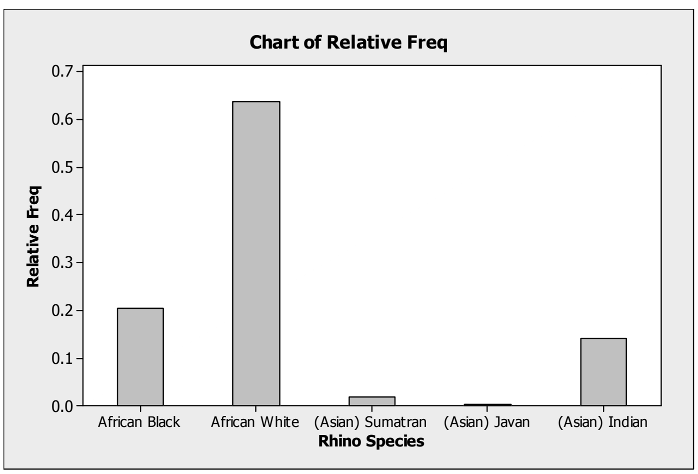
b. The number of years of education is quantitative since it can be measured on a numerical scale. The insomnia status (normal sleeper or chronic insomnia) is qualitative since it can not be measured on a numerical scale.

c. Less educated adults are more likely to have chronic insomnia.

1.13 a.

| **Rhino Species** | **Population** | **Relative Freq** |
|---|---|---|
| African Black | 3610 | 0.203 |
| African White | 11330 | 0.637 |
| (Asian) Sumatran | 300 | 0.017 |
| (Asian) Javan | 60 | 0.003 |
| (Asian) Indian | 2500 | 0.140 |

b.




c. African rhinos make up approximately 84% of all rhinos whereas Asian rhinos make up the remaining 16% of all rhinos.

1.15 a. Pie chart

b. The type of firearms owned is the qualitative variable.

c. Rifle (33%), shotgun (21%), and revolver (20%) are the most common types of firearm.

d.

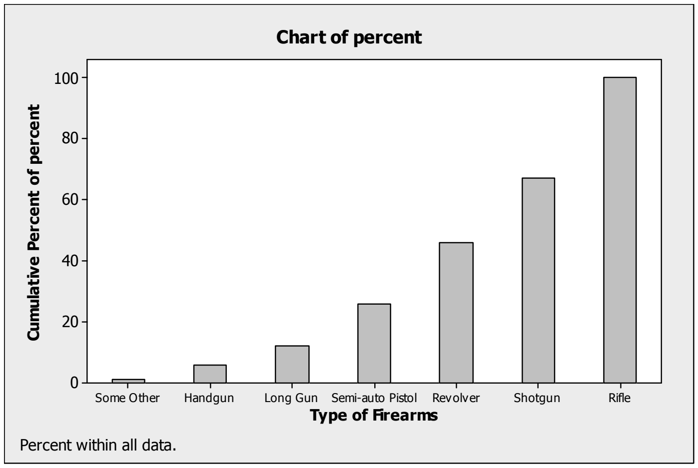
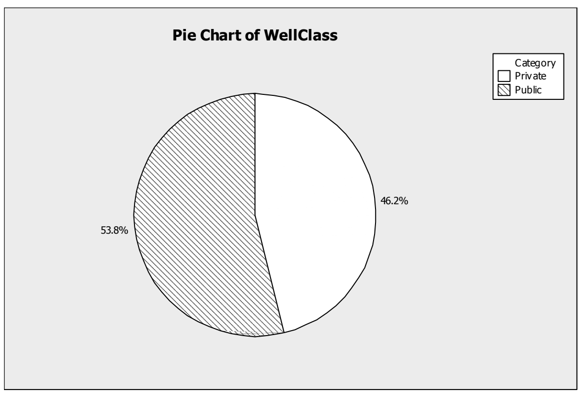
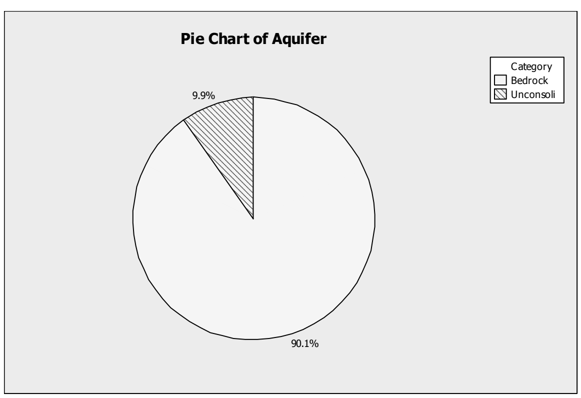


Chart of percent
Cumulative Percent of percent
100
80
60
40
20
0
Some Other
Handgun
Long Gun
Semi-auto Pistol
Revolver
Shotgun
Rifle
Type of Firearms
Percent within all data.

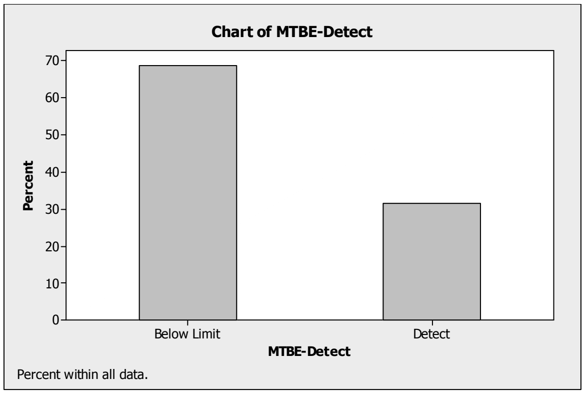
1.17 a.


Pie Chart of WellClass
Category
Private
Public
46.2%
53.8%

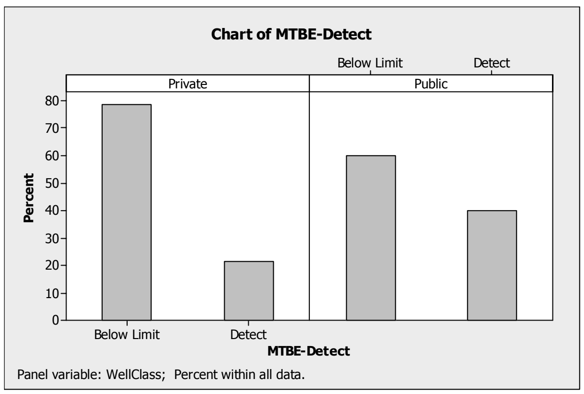
b.


Pie Chart of Aquifer
Category
Bedrock
Unconsoli
9.9%
90.1%

c.




d.




Public wells (40%); Private wells (21%).

1.19 a. A stem-and-leaf display of the data using MINITAB is:

```
Stem-and-leaf of FNE       N  = 25
Leaf Unit = 1.0

 2     0 67
 3     0 8
 6     1 001
10     1 3333
12     1 45
(2)    1 66
11     1 8999
 7     2 0011
 3     2 3
 2     2 45
```

b. The numbers in bold in the stem-and-leaf display represent the bulimic students. Those numbers tend to be the larger numbers. The larger numbers indicate a greater fear of negative evaluation. Yes, the bulimic students tend to have a greater fear of negative evaluation.

c. A measure of reliability indicates how certain one is that the conclusion drawn is correct. Without a measure of reliability, anyone could just guess at a conclusion.

1.21 Yes.

1.23 a.

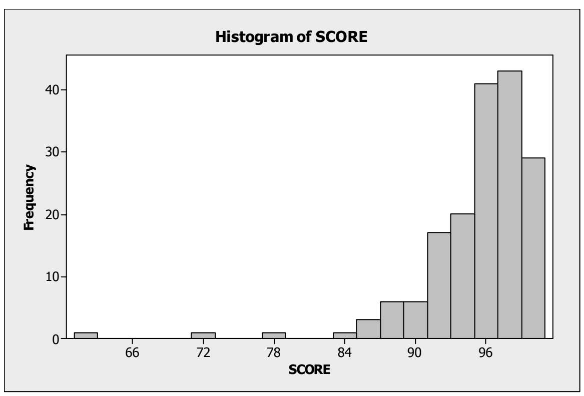
```
Stem-and-leaf of SCORE  N  = 169
Leaf Unit = 1.0

 1     6   2
 1     6
 2     7   2
 3     7   8
 4     8   4
 15    8   66677888899
 56    9   00001111111222222222233333333344444444444
(100)  9   555555555555555555555666666666666666666666777777777777777778888888+
 13    10  0000000000000
```

b. .98 or 98 out of every 100 ships have a sanitation score that is at least 86.

c.

```
Stem-and-leaf of SCORE  N  = 169
Leaf Unit = 1.0

 1     6   2
 1     6
 2     7   2
 3     7   8
 4     8   4
 15    8   66677888899
 56    9   00001111111222222222233333333344444444444
(100)  9   555555555555555555555666666666666666666666777777777777777778888888+
 13    10  0000000000000
```

d.




e. Approximately 95% of the ships have an acceptable sanitation standard.

1.25 a. 2.12; average magnitude for the aftershocks is 2.12.

b. 6.7; difference between the largest and smallest magnitude is 6.7.

c. .66; about 95% of the magnitudes fall in the interval mean $\pm$ 2(std. dev.) $= (.8, 3.4)$

d. $\mu =$ mean; $\sigma =$ Standard deviation

1.27 a. $\overline{y} = 94.91$, s = 4.83

b. $\overline{y} \pm 2s = 94.91 \pm 2 * 4.83 \Rightarrow (85.25, 104.57)$.

c. .976; yes

1.29 a. The average daily ammonia concentration $\bar{y}$ =

$$\frac{\sum y_i}{n} = \frac{1.53+1.50+1.37+1.51+1.55+1.42+1.41+1.48}{8}$$
$$= \frac{11.77}{8} = 1.47 \text{ ppm}$$

b. $$s^2 = \frac{\sum y_i^2 - n\bar{y}^2}{n-1} = \frac{\sum y_i^2 - \frac{\left(\sum y_i\right)^2}{n}}{n-1}$$

$$= \frac{\left(1.53^2+1.50^2+1.37^2+1.51^2+1.55^2+1.42^2+1.41^2+1.48^2\right) - \frac{(11.77)^2}{8}}{8-1}$$

$$s^2 = \frac{17.3453 - \frac{(11.77)^2}{8}}{8-1} = \frac{.0287}{7} = .0041$$
$$s = \sqrt{s^2} = \sqrt{.0041} = .0640$$

We would expect about most of the daily ammonia levels to fall with
$\hat{y} \pm 2s \Rightarrow 1.47 \pm 2(.0640) \Rightarrow (1.34, 1.60)$ ppm.

c. The morning drive-time has more variable ammonia levels as it has the larger standard deviation.

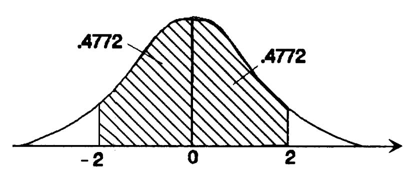
1.31 a. $(-111, 149)$

b. $(-91, 105)$

c. A student is more likely to get a 140-point increase on the SAT-Math test.

1.33 a. The $z$-score for $\mu - 2\sigma$ is $z = \frac{(\mu - 2\sigma) - \mu}{\sigma} = -2$

The $z$-score for $\mu + 2\sigma$ is $z = \frac{(\mu + 2\sigma) - \mu}{\sigma} = 2$

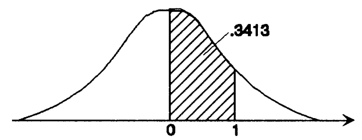



$P(\mu - 2\sigma \le y \le \mu + 2\sigma) = P(-2 \le z \le 2)$
$= P(-2 \le z \le 0) + P(0 \le z \le 2)$

Using Table 1 in Appendix D, $P(-2 \le z \le 0) = .4772$ and $P(0 \le z \le 2) = .4772$.
So $P(\mu - 2\sigma \le y \le \mu + 2\sigma) = .4772 + .4772 = .9544$

b. The $z$-score for $y = 108$ is $z = \dfrac{y-\mu}{\sigma} = \dfrac{108-100}{8} = 1$

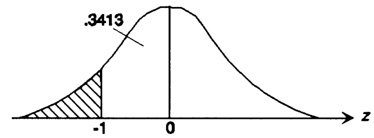
$P(y \geq 108) = P(z \geq 1)$




Using Table 1 of Appendix D, we find $P(0 \leq z \leq 1) = .3413$, so $P(z \geq 1) = .5 - .3413 = .1587$

c. The $z$-score for $y = 92$ is $z = \dfrac{y-\mu}{\sigma} = \dfrac{92-100}{8} = -1$

$P(y \leq 92) = P(z \leq -1)$

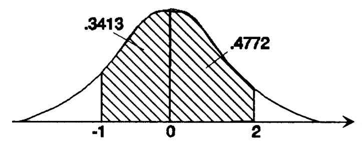



Using Table 1 of Appendix D, we find $P(-1 \leq z \leq 0) = .3413$, so $P(z \leq -1) = .5 - .3413 = .1587$

d. The $z$-score for $y = 92$ is $z = \dfrac{y-\mu}{\sigma} = \dfrac{92-100}{8} = -1$

The $z$-score for $y = 116$ is $z = \dfrac{y-\mu}{\sigma} = \dfrac{116-100}{8} = 2$

$P(92 \leq y \leq 116) = P(-1 \leq z \leq 2)$

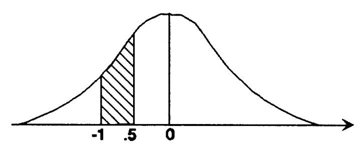



Using Table 1 of Appendix D, $P(-1 \leq z \leq 0) = .3413$ and $P(0 \leq z \leq 2) = .4772$. So $P(92 \leq y \leq 116) = P(-1 \leq z \leq 2)$ $= .3413 + .4772 = .8185$.

e. The $z$-score for $y = 92$ is $z = \dfrac{y-\mu}{\sigma} = \dfrac{92-100}{8} = -1$

The $z$-score for $y = 96$ is $z = \dfrac{y-\mu}{\sigma} = \dfrac{96-100}{8} = -.5$

$P(92 \leq y \leq 96) = P(-1 \leq z \leq -.5)$

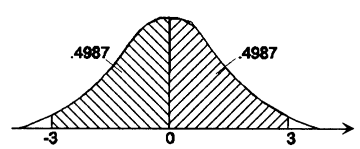



Using Table 1 of Appendix D, $P(-1 \leq z \leq 0) = .3413$ and $P(-.5 \leq z \leq 0) = .1915$. So $P(92 \leq y \leq 96)$ $= P(-1 \leq z \leq -.5) = .3413 - .1915 = .1498$.

f. The $z$-score for $y = 76$ is $z = \dfrac{y-\mu}{\sigma} = \dfrac{76-100}{8} = -3$

The $z$-score for $y = 124$ is $z = \dfrac{y-\mu}{\sigma} = \dfrac{124-100}{8} = 3$

$P(76 \leq y \leq 124) = P(-3 \leq z \leq 3)$




Using Table 1 of Appendix D, $P(-3 \leq z \leq 0) = .4987$ and $P(0 \leq z \leq 3) = .4987$. So $P(76 \leq y \leq 124) =$ $P(-3 \leq z \leq 3) = .4987 + .4987 = .9974$.

1.35 a. Let $x$ = alkalinity level of water specimens collected from the Han River.

Using Table 1, Appendix D,

$$P(y>45)=P\left(z>\frac{45-50}{3.2}\right)=P(z>-1.56)=.5+.4406=.9406.$$

b. Using Table 1, Appendix D,

$$P(y<55)=P\left(z<\frac{55-50}{3.2}\right)=P(z<1.56)=.5+.4406=.9406.$$

c. Using Table 1, Appendix D,

$$P(51<y<52)=P\left(\frac{51-50}{3.2}<z<\frac{52-50}{3.2}\right)=P(.31<z<.63)=.2357-.1217=.1140.$$

1.37 a. Using Table 1, Appendix D,

$$P(40<y<50)=P\left(\frac{40-37.9}{12.4}<z<\frac{50-37.9}{12.4}\right)=P(.17<z<.98)$$
$$=.3365-.0675=.2690.$$

b. Using Table 1, Appendix D,

$$P(y<30)=P\left(z<\frac{30-37.9}{12.4}\right)=P(z<-.64)=.5-.2389=.2611.$$

c. We know that if $P(z_L<z<z_U)=.95,$ then $P(z_L<z<\ 0)+P(0<z<z_U)=.95$ and

$$P(z_L<z<\ 0)=P(0<z<z_U)=.95/2=.4750.$$

Using Table 1, Appendix D, $z_U=1.96$ and $z_L=-1.96.$

$$P(y_L<y<y_U)=.95\Rightarrow P\left(\frac{y_L-37.9}{12.4}<z<\frac{y_U-37.9}{12.4}\right)=.95$$

$$\Rightarrow\frac{y_L-37.9}{12.4}=-1.96 \qquad \text{and} \quad \frac{y_U-37.9}{12.4}=1.96$$

$$\Rightarrow y_L-37.9=-24.3 \qquad \text{and} \quad y_U-37.9=24.3\Rightarrow y_L=13.6 \text{ and } y_U=62.2$$

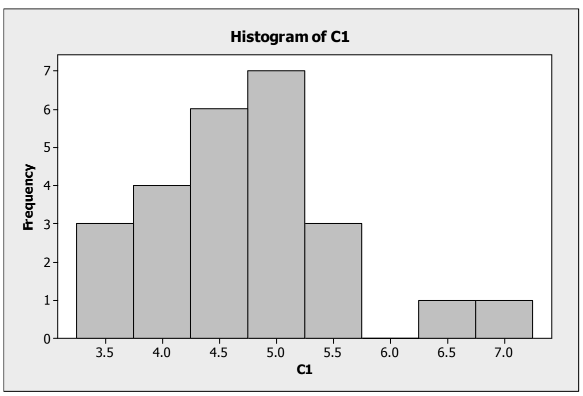
1.39 Using Table 1, Appendix D, $P(-1.5<Z<1.5)=2*0.4332=0.8664$. Approximately 87% of the time *Six Sigma* will met their goal.

1.41 a. The twenty-five means are:

| | | |
|---|---|---|
| 4.75 | 4.58 | 4.00 |
| 4.83 | 4.58 | 4.92 |
| 5.33 | 3.50 | 3.92 |
| 6.58 | 5.33 | 4.33 |
| 4.75 | 3.33 | 6.83 |
| 5.00 | 4.08 | 4.83 |
| 4.00 | 4.58 | 4.25 |
| 3.67 | 5.08 | 5.58 |
| 4.33 | | |

| Class | Frequency | Relative Frequency |
|---|---|---|
| 3.20 - 3.70 | 3 | 3/25 = .12 |
| 3.70 - 4.20 | 4 | 4/25 = .16 |
| 4.20 - 4.70 | 6 | 6/25 = .24 |
| 4.70 - 5.20 | 7 | 7/25 = .28 |
| 5.20 - 5.70 | 3 | 3/25 = .12 |
| 5.70 - 6.20 | 0 | 0/25 = .00 |
| 6.20 - 6.70 | 1 | 1/25 = .04 |
| 6.70 - 7.20 | 1 | 1/25 = .04 |

We can see that the histogram is less spread out than in the previous problem.




The mean of the sampling distribution is 4.680 and the standard deviation is .838. As expected, the standard deviation is smaller.

b. $\bar{\bar{y}} = \dfrac{\sum_{i=1}^{n} \bar{y}_i}{n} = \dfrac{117}{25} = 4.68$ $\qquad S_{\bar{y}} = \dfrac{\sum \left(\bar{y}_i - \bar{\bar{y}}\right)^2}{n-1} = \dfrac{20.112}{25-1} = .838$

This standard deviation is smaller than the one in the previous problem. Since the sample size is larger in this problem, we expect the standard deviation of $\bar{y}_i$'s to be smaller.

1.43 a. $E(\bar{y}) = \mu_{\bar{y}} = \mu = 0.10 \qquad Var(\bar{y}) = \frac{\sigma^2}{n} = \frac{(0.10)^2}{50} \cong 0.0002 \qquad \sigma_{\bar{y}} = \frac{s}{\sqrt{n}} = \frac{0.10}{\sqrt{50}} \cong 0.0141$

b. Since the sample size is greater than 30, the sample distribution of $\bar{y}$ is approximately normal by The Central Limit Theorem.

c. $$P(\bar{y} > 0.13) = P\left(Z > \frac{0.13 - 0.10}{\frac{0.10}{\sqrt{50}}}\right) = P(Z > 2.12) = 0.50 - 0.4830 = 0.0170$$

1.45 a. For confidence coefficient .99, $\alpha = .01$ and $\alpha/2 = .01/2 = .005$. From Table 1, Appendix D, $z_{.005} = 2.58$. The confidence interval is:

$$\bar{y} \pm z_{\alpha/2} \frac{s}{\sqrt{n}} \Rightarrow 1.13 \pm 2.58\left(\frac{2.21}{\sqrt{72}}\right) \Rightarrow 1.13 \pm .67 \Rightarrow (.46,\ 1.80)$$

We are 99% confident that the true mean number of pecks made by chickens pecking at blue string is between .458 and 1.802.

b. Yes, there is evidence that chickens are more apt to peck at white string. The mean number of pecks at white string is 7.5. Since 7.5 is not in the 99% confidence interval for the mean number of pecks at blue string, it is not a likely value for the true mean for blue string.

1.47 a. $E(y) = \mu_{\bar{y}} = \mu = 99.6$

b. From Table 1 of Appendix D, $Z = 1.96$

$$\bar{y} \pm z(s_{\bar{y}}) = \bar{y} \pm z\left(\frac{s}{\sqrt{n}}\right) = 99.6 \pm 1.96\left(\frac{12.6}{\sqrt{122}}\right) = 99.6 \pm 2.2 \Rightarrow (97.4, 101.8)$$

c. We are 95% confident that the true mean Mach rating score is between 97.4 and 101.8.

d. Yes, since the value of 85 is not contained in the confidence interval it is unlikely that the true mean Mach rating score could be 85.

1.49 Using Table 2, Appendix D,

$$\bar{y} \pm t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right) = \bar{y} \pm t_{0.005}\left(\frac{s}{\sqrt{n}}\right) = 19 \pm 3.055\left(\frac{2.2}{\sqrt{13}}\right) = 19 \pm 1.9 \Rightarrow (17.1, 20.9)$$

We are 99% confident that the true mean quality of the methodology of the Wong scale is between 17.1 and 20.9.

1.51 a. Null Hypothesis = $H_0$

b. Alternative Hypothesis = $H_a$

c. Type I error is when we reject the null hypothesis when the null hypothesis is in fact true.

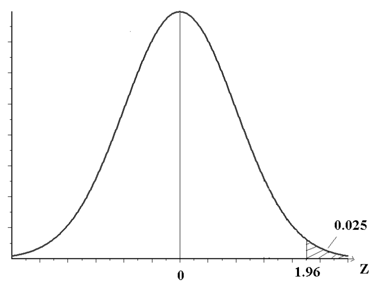
d. Type II error is when we do not reject the null hypothesis when the null hypothesis is in fact not true.

e. Probability of Type I error is $\alpha$

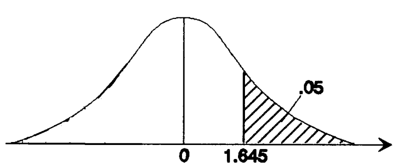
f. Probability of Type II error is $\beta$

g. $p$-value is the observed significance level, which is the probability of observing a value of the test statistics at least as contradictory to the null hypothesis as the observed test statistic value, assuming the null hypothesis is true.

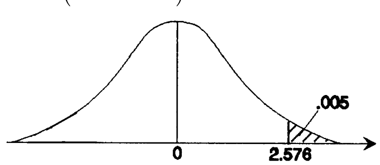
1.53 a. $\alpha = P$ (reject $H_0$ when $H_0$ is in fact true)
$= P(z > 1.96) = .025$



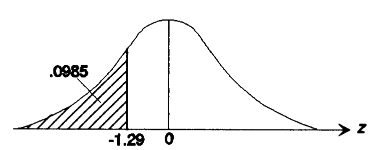

b. $\alpha = P(z > 1.645) = .05.$



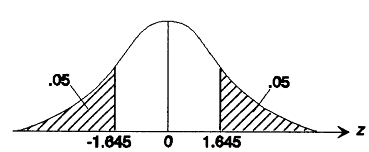

c. $\alpha = P(z > 2.576) = .005.$



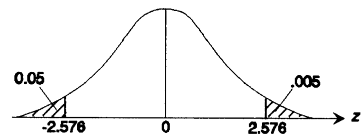

d. $\alpha = P(z < -1.29) = .0985.$




e. $\alpha = P(z < -1.645 \text{ or } z > 1.645)$
$= .05 + .05 = .10$



f. $\alpha = P(z < -2.576 \text{ or } z > 2.576)$
$= .005 + .005 = .01$




1.55 a. To determine if the average level of mercury uptake in wading birds in the Everglades in 2000 is less than 15 parts per million, we test:

$H_0 : \mu = 15$
$H_a : \mu < 15$

b. A Type I error is rejecting $H_0$ when $H_0$ is true. In terms of this problem, we would be concluding that the average level of mercury uptake in wading birds in the Everglades in 2000 is less than 15 parts per million, when in fact, the average level of mercury uptake in wading birds in the Everglades in 2000 is equal to 15 parts per million.

c. A Type II error is accepting $H_0$ when $H_0$ is false. In terms of this problem, we would be concluding that the average level of mercury uptake in wading birds in the Everglades in 2000 is equal to 15 parts per million, when in fact, the average level of mercury uptake in wading birds in the Everglades in 2000 is less than 15 parts per million.

1.57 a. To determine if the mean social interaction score of all Connecticut mental health patients differs from 3, we test:

$H_0 : \mu = 3$
$H_a : \mu \neq 3$

The test statistic is $z = \dfrac{\overline{y} - \mu_0}{\sigma_{\overline{y}}} = \dfrac{2.95 - 3}{1.10 / \sqrt{6,681}} = -3.72$

The rejection region requires $\alpha / 2 = .01 / 2 = .005$ in each tail of the $z$ distribution. From Table 1, Appendix D, $z_{.005} = 2.58$. The rejection region is $z < -2.58$ or $z > 2.58$.

Since the observed value of the test statistic falls in the rejection region $(z = -3.72 < -2.58), H_0$ is rejected. There is sufficient evidence to indicate that the mean social interaction score of all Connecticut mental health patients differs from 3 at $\alpha = .01$.

b. From the test in part a, we found that the mean social interaction score was statistically different from 3. However, the sample mean score was 2.95. Practically speaking, 2.95 is very similar to 3.0. The very large sample size, $n = 6681$, makes it very easy to find statistical significance, even when no practical significance exists.

c. Because the variable of interest is measured on a 5-point scale, it is very unlikely that the population of the ratings will be normal. However, because the sample size was extremely large, $(n = 6681)$, the Central Limit Theorem will apply. Thus, the distribution of $\overline{y}$ will be normal, regardless of the distribution of $y$. Thus, the analysis used above is appropriate.

.

1.59 a. The test statistic is $z = \frac{\bar{y} - \mu_0}{\sigma_{\bar{y}}} = \frac{10.2 - 0}{31.3 / \sqrt{50}} = 2.3$

b. To determine if the mean level of feminization differs from 0%, we test:

$$H_0 : \mu = 0$$
$$H_a : \mu \neq 0$$

Since the alternative hypothesis contains $\neq$, this is a two-tailed test. The $p$-value is $p = P(z \le -2.3) + P(z \ge 2.3) = .5 - P(-2.3 < z < 0) + .5 - P(0 < z < 2.3) =$ $.5 - .4893 + .5 - .4893 = .0214$. Reject $H_0$.

c. The test statistic is $z = \frac{\bar{y} - \mu_0}{\sigma_{\bar{y}}} = \frac{15 - 0}{25.1 / \sqrt{50}} = 4.23$.

Since the alternative hypothesis contains $\neq$, this is a two-tailed test. The $p$-value is $p = P(z \le -4.23) + P(z \ge 4.23) = .5 - P(-4.23 < z < 0) + .5 - P(0 < z < 4.23) \approx$ $.5 - .5 + .5 - .5 = 0$. Reject $H_0$.

1.61 Some preliminary calculations are:

$$\bar{y} = \frac{\sum y}{n} = \frac{110}{5} = 22$$

$$s^2 = \frac{\sum y^2 - \frac{\left(\sum y\right)^2}{n}}{n-1} = \frac{2,436 - \frac{110^2}{5}}{5-1} = 4$$

$$s = \sqrt{s^2} = \sqrt{4} = 2$$

To determine if the data collected were fabricated, we test:

$$H_0 : \mu = 15$$
$$H_a : \mu \neq 15$$

The test statistic is $t = \frac{\bar{y} - \mu_0}{s / \sqrt{n}} = \frac{22 - 15}{2 / \sqrt{5}} = 7.83$

If we want to choose a level of significance to benefit the students, we would choose a small value for $\alpha$. Suppose we use $\alpha = .01$. The rejection region requires $\alpha / 2 = .01 / 2 = .005$ in each tail of the t distribution with df $= n - 1 = 5 - 1 = 4$. From Table 2, Appendix D, $t_{.005} = 4.604$. The rejection region is $t < -4.604$ or $t > 4.604$.

Since the observed value of the test statistic falls in the rejection region $(t = 7.83 > 4.604), H_0$ is rejected. There is sufficient evidence to indicate the mean data collected were fabricated at $\alpha = .01$.

1.63 There are three things to describe:

1) Mean: $\mu_{\bar{y}_1-\bar{y}_2} = \mu_1 - \mu_2$

2) Std Deviation: $\sigma_{\bar{y}_1-\bar{y}_2} = \sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$

3) Shape: For sufficiently large samples, the shape of the sampling distribution is approximately normal.

1.65 For this experiment let $\mu_1$ and $\mu_2$ represent the mean ratings for Group 1 (support favored position) and Group 2 (weaken opposing position), respectively. Then we want to test:
$H_0 : \mu_1 - \mu_2 = 0$ (i.e. no difference in mean ratings)
$H_a : \mu_1 - \mu_2 \neq 0$ (i.e. $\mu_1 \neq \mu_2$)

Calculate the pooled estimate of variance:
$s_p^2 = \dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = \dfrac{(26-1)(12.5)^2 + (26-1)(12.2)^2}{26+26-2} = 152.545$ where $s_p^2$ is based on $(n_1+n_2-2) = (26+26-2) = 50$ degrees of freedom.

Now we compute the test statistic:
$$t = \frac{(\bar{y}_1 - \bar{y}_2) - D_0}{\sqrt{s_p^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} = \frac{(28.6-24.9)-0}{\sqrt{152.545\left(\dfrac{1}{26} + \dfrac{1}{26}\right)}} = 1.08$$
The rejection region for this two-tailed test at $\alpha = 0.05$, based on 50 degrees of freedom is $|t| > t_{0.025} = 2.009$. Since the computed t value does not fall in the rejection region, we fail to reject the null hypothesis. There is insufficient evidence of a difference between the true mean rating scores for the two groups.

Assumptions: This procedure requires the assumption that the samples of rating scores are randomly and independently selected from normal populations with equal variances.

1.67 For this experiment let $\mu_1$ and $\mu_2$ represent the performance level of students in the control group and the rudeness condition group, respectively. Then we want to test:

$H_0 : \mu_1 - \mu_2 = 0$ (i.e. no difference in mean ratings)
$H_a : \mu_1 - \mu_2 > 0$ (i.e. $\mu_1 > \mu_2$)

Calculate the pooled estimate of variance:
$s_p^2 = \dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = \dfrac{(53-1)(7.38)^2 + (45-1)(3.992)^2}{53+45-2} = 36.806$ where $s_p^2$ is based on $(n_1+n_2-2) = (53+45-2) = 96$ degrees of freedom.

Now we compute the test statistic:

$$t = \frac{(\bar{y}_1 - \bar{y}_2) - D_0}{\sqrt{s_p^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} = \frac{(11.81-8.51)-0}{\sqrt{36.806\left(\dfrac{1}{53} + \dfrac{1}{45}\right)}} = 2.68$$

Since the $p-\text{value} = .0043 < \alpha = 0.01$, we reject the null hypothesis and conclude that there is significant evidence that the true mean performance level for students in the rudeness condition is lower than the true mean performance level for students in the control group.

Assumptions: This procedure requires the assumption that the samples are randomly and independently selected from normal populations with equal variances.

1.69 a. Using MINITAB, the output for comparing the mean level of family involvement in science homework assignments of TIPS and ATIPS students is:

```
Two sample T for GSHWS

COND          N      Mean     StDev   SE Mean
ATIPS        98      1.43      1.06      0.11
TIPS        128      2.55      1.27      0.11

95% CI for mu (ATIPS) - mu (TIPS ): ( -1.43,  -0.82)
T-Test mu (ATIPS) = mu (TIPS ) (vs not =): T = -7.24 P = 0.0000 DF = 222
```

Let $\mu_1 =$ mean level of involvement in science homework assignments for TIPS students and $\mu_2 =$ mean level of involvement in science homework assignments for ATIPS students.
To compare the mean level of family involvement in science homework assignments of TIPS and ATIPS students, we test:

$$H_0: \mu_1 = \mu_2$$
$$H_a: \mu_1 \neq \mu_2$$

From the printout, the test statistic is $t = -7.24$ and the *p*-value is $p = .0000$. Since the *p*-value is less than $\alpha\left(p = .0000 < .05\right), H_0$ is rejected. There is sufficient evidence to indicate a difference in the mean level of family involvement in science homework assignments between TIPS and ATIPS students at $\alpha = .05$.

b. Using MINITAB, the output for comparing the mean level of family involvement in mathematics homework assignments of TIPS and ATIPS students is:

```
Two sample T for MTHHWS

COND          N      Mean     StDev   SE Mean
ATIPS        98      1.48      1.22      0.12
TIPS        128      1.56      1.27      0.11

95% CI for mu (ATIPS) - mu (TIPS ): ( -0.41,  0.25)
T-Test mu (ATIPS) = mu (TIPS ) (vs not =): T = -0.50  P = 0.62  DF = 212
```

Let $\mu_1 =$ mean level of involvement in mathematics homework assignments for TIPS students and $\mu_2 =$ mean level of involvement in mathematics homework assignments for ATIPS students.

To compare the mean level of family involvement in mathematics homework assignments of TIPS and ATIPS students, we test:

$$H_0: \mu_1 = \mu_2$$
$$H_a: \mu_1 \neq \mu_2$$

From the printout, the test statistic is $t = -0.50$ and the $p$-value is $p = .62$. Since the $p$-value is not less than $\alpha\left(p = .62 > .05\right), H_0$ is not rejected. There is insufficient evidence to indicate a difference in the mean level of family involvement in mathematics homework assignments between TIPS and ATIPS students at $\alpha = .05$.

c. Using MINITAB, the output for comparing the mean level of family involvement in language arts homework assignments of TIPS and ATIPS students is:

```
Two sample T for LAHWS

COND          N      Mean     StDev   SE Mean
ATIPS        98      1.01      1.09     0.11
TIPS        128      1.20      1.12     0.099

95% CI for mu (ATIPS) - mu (TIPS ): ( -0.48,  0.106)
T-Test mu (ATIPS) = mu (TIPS ) (vs not =): T = -1.25  P = 0.21  DF = 211
```

Let $\mu_1 =$ mean level of involvement in language arts homework assignments for TIPS students and $\mu_2 =$ mean level of involvement in language arts homework assignments for ATIPS students. To compare the mean level of family involvement in language arts homework assignments of TIPS and ATIPS students, we test:

$H_0: \mu_1 = \mu_2$
$H_\text{a}: \mu_1 \neq \mu_2$

From the printout, the test statistic is $t = -1.25$ and the $p$-value is $p = .21$. Since the $p$-value is not less than $\alpha\left(p = .21 > .05\right), H_0$ is not rejected. There is insufficient evidence to indicate a difference in the mean level of family involvement in language arts homework assignments between TIPS and ATIPS students at $\alpha = .05$.

d. Since both sample sizes are greater than 30, the only assumption necessary is:

1. The samples are random and independent.

From the information given, there is no reason to dispute this assumption.

1.71 a. Each participant acted as a speaker and an audience member

b. Let $\mu_d = \mu_{\text{speaker}} - \mu_{\text{audience}} =$ true mean number of laugh episodes for speakers minus the true mean number of laugh episodes as an audience member.

c. No, you need sample statistics for differences.

d. When testing the hypothesis: $H_0: \mu_d = 0$ vs $H_a: \mu_d \neq 0$, the t-test revealed that the $p-$value $< .01$. Thus, we can reject $H_0$ and conclude that there is a significant difference in the true mean number of laugh episodes for speakers and audience members.

1.73 a. Let $\mu_1 - \mu_2 =$ FULLDARK-TRLIGHT . Then we want to test the hypothesis:
$H_0: \mu_1 - \mu_2 = 0$
$H_a: \mu_1 - \mu_2 \neq 0$ no difference
Assumption: The population differences between the mean standardized growth of genes in the full-dark condition and the genes in the transient light condition are approximately normally distributed.

```
Two-sample T for fulldark vs trlight

               N    Mean   StDev  SE Mean
fulldark      10  -0.285  0.435     0.14
trlight       10   0.93   1.21      0.38

Difference = mu (fulldark) - mu (trlight)
Estimate for difference:  -1.212
99% CI for difference:  (-2.478, -0.055)
T-Test of difference = 0 (vs not =): T-Value = -2.97  P-Value = 0.016  DF = 11
```

Since the $p$-value is greater than 0.01, we can not conclude that the true means are significantly different.

b. $H_0 : \mu_1 - \mu_2 = 0$

$H_a : \mu_1 - \mu_2 \neq 0$ no difference

Assumption: The population differences between the mean standardized growth of genes in the full-dark condition and the genes in the transient light condition are approximately normally distributed.

```
Two-sample T for FULL-DARK vs TR-LIGHT

             N    Mean  StDev  SE Mean
FULL-DARK  103   -0.318  0.597    0.059
TR-LIGHT   103    0.10   1.14     0.11


Difference = mu (FULL-DARK) - mu (TR-LIGHT)
Estimate for difference:  -0.420
99% CI for difference:  (0.090, 0.750)
T-Test of difference = 0 (vs not =): T-Value = -3.32  P-Value = 0.001  DF = 154
```

No, the sample of the first ten observations did not show any difference, but the entire sample of 103 has a $p$-value of .001 which indicates that there is a significant difference in the means.

c. Let $\mu_1 - \mu_2 = \text{FULLDARK-TRDARK}$. Then we want to test the hypothesis:

$H_0 : \mu_1 - \mu_2 = 0$

$H_a : \mu_1 - \mu_2 \neq 0$ no difference

Assumption: The population differences between the mean standardized growth of genes in the full-dark condition and the genes in the transient light condition are approximately normally distributed.

```
Two-sample T for FULL-DARK vs TR-DARK

            N    Mean  StDev  SE Mean
FULL-DARK  10  -0.285  0.435     0.14
TR-DARK    10  -0.425  0.567     0.18


Difference = mu (FULL-DARK) - mu (TR-DARK)
Estimate for difference:  0.140
99% CI for difference:  (-0.520, 0.801)
T-Test of difference = 0 (vs not =): T-Value = 0.57  P-Value = 0.58  DF = 16
```

No difference in the means with the first ten data points since the *p*-value is only .543. When all 103 data points are used, there is still not a significant difference in the full-dark and transient-dark at the .01 significance level since the *p*-value is .017 as seen below.

```
     Two-sample T for FULL-DARK vs TR-DARK

                  N    Mean  StDev  SE Mean
FULL-DARK       103  -0.318  0.597    0.059
TR-DARK         103  -0.091  0.746    0.073

Difference = mu (FULL-DARK) - mu (TR-DARK)
Estimate for difference:  -0.2274
99% CI for difference:  (-0.472, 0.0174)
T-Test of difference = 0 (vs not =): T-Value = -2.42  P-Value = 0.017 DF = 194
```

d. Let $\mu_1 - \mu_2 = \text{TRLIGHT-TRDARK}$. Then we want to test the hypothesis:

$H_0 : \mu_1 - \mu_2 = 0$

$H_a : \mu_1 - \mu_2 \neq 0$ no difference

Assumption: The population differences between the mean standardized growth of genes in the full-dark condition and the genes in the transient light condition are approximately normally distributed.

```
        Two-sample T for TR-LIGHT vs TR-DARK

                N    Mean  StDev  SE Mean
TR-LIGHT  10    0.93   1.21     0.38
TR-DARK   10  -0.425  0.567     0.18

Difference = mu (TR-LIGHT) - mu (TR-DARK)
Estimate for difference:  1.352
99% CI for difference:  (-0.058, 2.646)
T-Test of difference = 0 (vs not =): T-Value = 3.23  P-Value = 0.01  DF = 12
```

Yes, there is a significant difference between the transient dark and transient light means of the first ten data points since the *p*-value is .008 as presented above. However, when all 103 data points are used in the analysis which is presented below we see that the *p*-value is .153 which is indicates that there is not significant difference in the means.

```
Two-sample T for TR-LIGHT vs TR-DARK

             N    Mean  StDev  SE Mean
TR-LIGHT  103    0.10   1.14     0.11
TR-DARK   103  -0.091  0.746    0.073


Difference = mu (TR-LIGHT) - mu (TR-DARK)
Estimate for difference:  0.192
99% CI for difference:  (-0.157, 0.541)
T-Test of difference = 0 (vs not =): T-Value = 1.44  P-Value = 0.153  DF = 176
```

1.75 We will test to see if the ratio of the variances differ from 1 or not:

$$H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1 \Rightarrow \sigma_1^2 = \sigma_2^2$$

$$H_a : \frac{\sigma_1^2}{\sigma_2^2} \neq 1 \Rightarrow \sigma_1^2 \neq \sigma_2^2$$

Two assumptions are required for the F test are as follows:

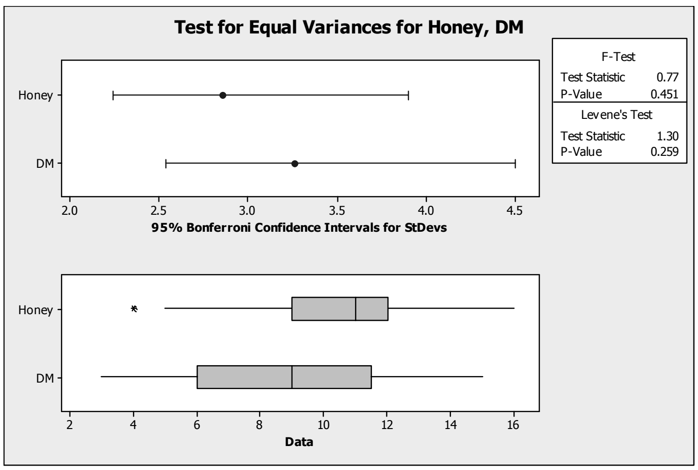
1. The two sample populations are normally distributed.
2. The samples are randomly and independently selected from their respective populations

We will use the test statistic: $F = \frac{s_1^2}{s_2^2} = \frac{10.604}{8.151} = 1.30$

$df = (n_1 - 1), (n_2 - 1) = (33 - 1 = 32), (35 - 1 = 34)$ respectively

Referring to Table D.4 of Appendix D: $\alpha = 0.10$, we see that we will reject the null hypothesis if the calculated value of F exceeds the tabulated value: $F_{0.05} = 1.84$. Since the calculated F value = 1.30 does not fall in the rejection region, the data does not provide sufficient evidence to show that the variances of the two groups are unequal; $\frac{\sigma_{\text{DM}}^2}{\sigma_{\text{H}}^2} = 1$.

Note: we will always place the larger sample variance in the numerator of the F test.




1.77 Let $\sigma_{\text{B}}^2$ = variance of the FNE scores for bulimic students and $\sigma_{\text{N}}^2$ = variance of the FNE scores for normal students.

From Exercise 1.59, $s_{\text{B}}^2 = 24.1636$ and $s_{\text{N}}^2 = 27.9780$

To determine if the variances are equal, we test:

$H_0 : \sigma_{\text{B}}^2 = \sigma_{\text{N}}^2$

$H_a : \sigma_{\text{B}}^2 \neq \sigma_{\text{N}}^2$

The test statistic is $F = \frac{\text{Larger sample variance}}{\text{Smaller sample variance}} = \frac{s_{\text{N}}^2}{s_{\text{B}}^2} = \frac{27.9780}{24.1636} = 1.16$

The rejection region requires $\alpha / 2 = .05 / 2 = .025$ in the upper tail of the F distribution with numerator df $\nu_2 = n_2 - 1 = 14 - 1 = 13$ and denominator df $\nu_1 = n_1 - 1 = 11 - 1 = 10$. From Table 5, Appendix D, $F_{.025} \approx 3.62$. The rejection region is $F > 3.62$.

Since the observed value of the test statistic does not fall in the rejection region $(F = 1.16 < 3.62), H_0$ is not rejected; $\frac{\sigma_{\text{B}}^2}{\sigma_{\text{N}}^2} = 1$. It appears that the assumption of equal variances is valid.

1.79 $H_0: \sigma_1^2 / \sigma_2^2 = 1 \quad \left(\sigma_1^2 = \sigma_2^2\right)$

$H_a: \sigma_1^2 / \sigma_2^2 \neq 1 \quad \left(\sigma_1^2 \neq \sigma_2^2\right)$

where $\sigma_1^2$ = variance of the one wet sampler readings and $\sigma_2^2$ = variance of the three wet sampler readings.

Test statistic: $F = \dfrac{\text{larger sample variance}}{\text{smaller sample variance}} = \dfrac{s_1^2}{s_2^2} = \dfrac{6.3^2}{2.6^2} = 5.87$

Rejection region: Using $\alpha = .05, n_1\text{-}1 = 364$ numerator df, $n_2\text{-}1 = 364$ denominator df, $F_{.025} \approx 1.00$.

Reject $H_0$ if $F > 1.00$

Conclusion: Reject $H_0; \sigma_1^2 / \sigma_2^2 \neq 1$ There is sufficient evidence (at $\alpha = .05$ ) to indicate the variations in hydrogen readings for the two sampling schemes differ.

Assumptions: The distributions of the hydrogen readings for the one wet sampler and the three wet samplers are both approximately normal. The samplers of the hydrogen readings are randomly and independently selected from their populations.

1.81 a. $\bar{y} = \dfrac{\sum_{i=1}^{n} y_i}{n} = \dfrac{11+2+2+1+9}{5} = \dfrac{25}{5} = 5$

$$s^2 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1} = \frac{\sum_{i=1}^{n} y_i^2 - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n}}{n-1} = \frac{(11^2+2^2+2^2+1^2+9^2) - \frac{(25)^2}{5}}{5-1}$$

$$= \frac{211-125}{4} = \frac{86}{4} = 21.5$$

$s = \sqrt{s^2} = \sqrt{21.5} \approx 4.637$

b. $\bar{y} = \dfrac{\sum_{i=1}^{n} y_i}{n} = \dfrac{22+9+21+15}{4} = \dfrac{67}{4} = 16.75$

$$s^2 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1} = \frac{\sum_{i=1}^{n} y_i^2 - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n}}{n-1} = \frac{(22^2+9^2+21^2+15^2) - \frac{(67)^2}{4}}{4-1}$$

$$= \frac{1231-1122.25}{3} = \frac{108.75}{3} = 36.25 \qquad s = \sqrt{s^2} = \sqrt{36.25} \approx 6.021$$

c. $\overline{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{34}{7} = 4.857$

$$s^2 = \frac{\sum_{i=1}^{n} y_i^2 - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n}}{n-1} = \frac{344 - \frac{(34)^2}{7}}{7-1} = \frac{178.857}{6} \approx 29.81$$

$s = \sqrt{s^2} = \sqrt{29.81} \approx 5.460$

d. $\overline{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{16}{4} = 4$

$$s^2 = \frac{\sum_{i=1}^{n} y_i^2 - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n}}{n-1} = \frac{64 - \frac{(16)^2}{4}}{4-1} = \frac{0}{3} \approx 0$$

$s = \sqrt{s^2} = \sqrt{0} \approx 0$

1.83 a. $z = \frac{y-\mu}{\sigma} = \frac{10-30}{5} = \frac{-20}{5} = -4$

The sign and magnitude of the $z$-value indicate that the $y$-value is 4 standard deviations below the mean.

b. $z = \frac{y-\mu}{\sigma} = \frac{32.5-30}{5} = \frac{2.5}{5} = .5$

The $y$-value is .5 standard deviation above the mean.

c. $z = \frac{y-\mu}{\sigma} = \frac{30-30}{5} = \frac{0}{5} = 0$

The $y$-value is equal to the mean of the random variable $y$.

d. $z = \frac{y-\mu}{\sigma} = \frac{60-30}{5} = \frac{30}{5} = 6$

The $y$-value is 6 standard deviations above the mean.

1.85 a. The population of interest is all men and women.

b. The sample of interest is 300 people from Gainesville, Florida

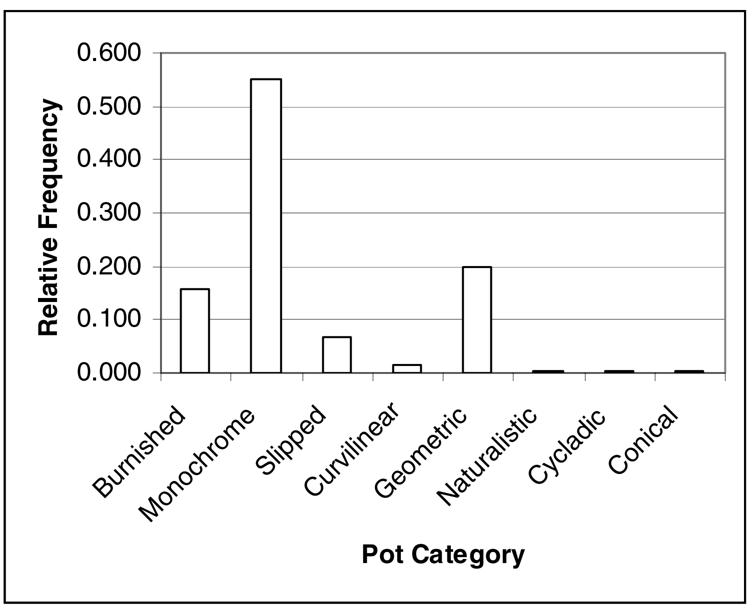
c. The study involves inferential statistics.

d. One variable is measured for each of the 20 objects placed. For each variable, the 2 possible outcomes were "yes" (place of objects was recalled) and "no" (place of object was not recalled). Since the outcomes "yes" and "no" are not measured on a numerical scale, the variables are qualitative.

1.87 Suppose we construct a relative frequency bar chart for this data. This will allow the archaeologists to compare the different categories easier. First, we must compute the relative frequencies for the categories. These are found by dividing the frequencies in each category by the total 837. For the burnished category, the relative frequency is 133 / 837 = .159. The rest of the relative frequencies are found in a similar fashion and are listed in the table.

| Pot Category | Number Found | Computation | Relative Frequency |
|---|---|---|---|
| Burnished | 133 | 133 / 837 | .159 |
| Monochrome | 460 | 460 / 837 | .550 |
| Slipped | 55 | 55 / 837 | .066 |
| Curvilinear Decoration | 14 | 14 / 837 | .017 |
| Geometric Decoration | 165 | 165 / 837 | .197 |
| Naturalistic Decoration | 4 | 4 / 837 | .005 |
| Cycladic White clay | 4 | 4 / 837 | .005 |
| Cononical cup clay | 2 | 2 / 837 | .002 |
| Total | 837 | | 1.001 |

A relative frequency bar chart is:




The most frequently found type of pot was the Monochrome. Of all the pots found, 55% were Monochrome. The next most frequently found type of pot was the Painted in Geometric Decoration. Of all the pots found, 19.7% were of this type. Very few pots of the types Painted in Naturalistic Decoration, Cycladic White clay, and Conical cup clay were found.

1.89 a. Using Minitab, the stem-and-leaf display for the data is:

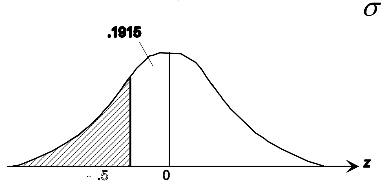
```
Stem-and-Leaf of LOSS      N = 19
Leaf Unit = 1.0

 8      0 11234559
(3)     1 123
 8      2 00
 6      3 9
 5      4 6
 4      5 6
 3      6 15
 1      7
 1      8
 1      9
 1     10 0
```

b. The numbers circled on the display in part a are associated with the eclipses of Saturnian satellites.

c. Since the five largest numbers are associated with eclipses of Saturnian satellites, it is much more likely that the greater light loss is associated with eclipses rather than occults.

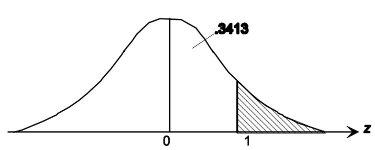
1.91 For each of these questions, we will use Table 1 in Appendix D.

a. The $z$-score for $y = 75$ is $z = \dfrac{y-\mu}{\sigma} = \dfrac{75-80}{10} = -.50$




$$\begin{aligned} P(y \le 75) &= P(z \le -.5) \\ &= .5 - P(-.5 \le z \le 0) \\ &= .5 - .1915 \\ &= .3085 \end{aligned}$$

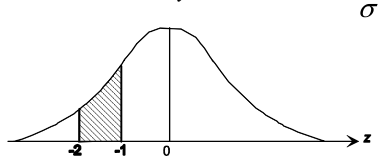
b. The $z$-score for $y = 90$ is $z = \dfrac{y-\mu}{\sigma} = \dfrac{90-80}{10} = 1.00$




$$\begin{aligned} P(y \ge 90) &= P(z \ge 1.00) \\ &= .5 - P(0 \le z \le 1.00) \\ &= .5 - .3413 \\ &= .1587 \end{aligned}$$

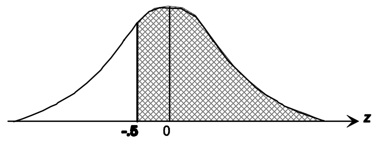
c. The $z$-score for $y = 60$ is $z = \dfrac{y-\mu}{\sigma} = \dfrac{60-80}{10} = -2$

The $z$-score for $y = 70$ is $z = \dfrac{y-\mu}{\sigma} = \dfrac{70-80}{10} = -1$




$$\begin{aligned} P(60 \le y \le 70) &= P(-2 \le z \le -1) \\ &= P(-2 \le z \le 0) \\ &\quad - P(-1 \le z \le 0) \\ &= .4772 - .3413 \\ &= .1359 \end{aligned}$$

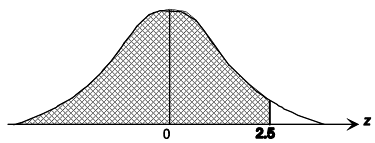
d.




$$P(y \geq 75) = P(z \geq -.5)$$
$$= .5 + P(-.5 \leq z \leq 0)$$
$$= .5 + .1915$$
$$= .6915$$

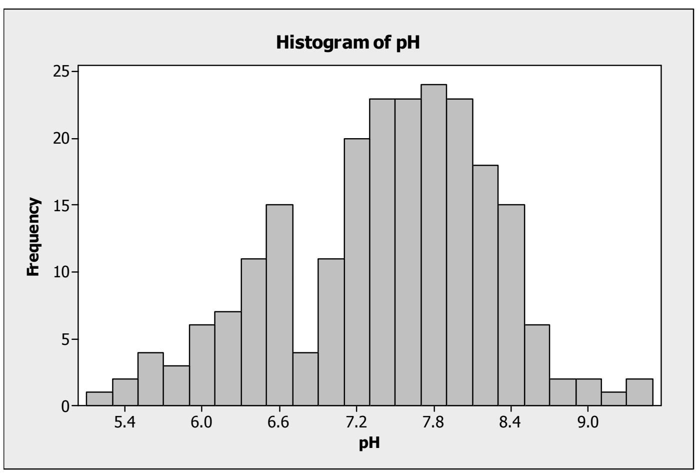
e. $P(y = 75) = P(z = -.5) = 0$

f. The $z$-score for $y = 105$ is $z = \dfrac{y-\mu}{\sigma} = \dfrac{105-80}{10} = 2.5$

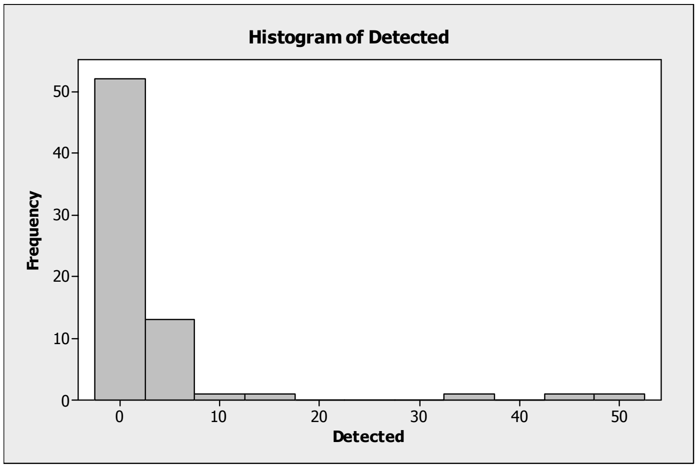



$$P(y \leq 105) = P(z \leq 2.5)$$
$$= .5 + P(0 \leq z \leq 2.5)$$
$$= .5 + .4938$$
$$= .9938$$

1.93 a. It appears that about 58 out of the 223 or .26 of New Hampshire wells have a pH level less than 7.0.



b. It appears that about 6 out of 70 or .086 have a value greater than 5 micrograms per liter.




c. $\bar{y} = 7.43$, $s = 0.82$, $\bar{y} \pm 2 * s \Rightarrow 7.43 \pm 2 * 0.82 \Rightarrow (5.79, 9.06); 95\%$ (Empirical Rule).

d. $\bar{y} = 1.22$, $s = 5.11$, $\bar{y} \pm 2 * s \Rightarrow 1.22 \pm 2 * 5.11 \Rightarrow (-9.00, 11.44); 75\%$ (Empirical Rule).

1.95 a. Let $y =$ change in SAT-MATH score. Using Table 1, Appendix D,

$$P(y \geq 50) = P\left(z \geq \frac{50-19}{65}\right) = P(z \geq .48) = .5 - .1844 = .3156.$$

b. Let $y =$ change in SAT-VERBAL score. Using Table 1, Appendix D,

$$P(y \geq 50) = P\left(z \geq \frac{50-7}{49}\right) = P(z \geq .88) = .5 - .3106 = .1894.$$

1.97 For this problem, $\mu_{\bar{y}} = \mu = 4.59$ and $\sigma_{\bar{y}} = \frac{\sigma}{\sqrt{n}} = \frac{2.95}{\sqrt{50}} = .4172$

a. $P(\bar{y} \geq 6) = P\left(z \geq \frac{6-4.59}{.4172}\right) = P(z \geq 3.38) \approx .5 - .5 = 0$

(using Table 1, Appendix D)

Since the probability of observing a sample mean CAHS score of 6 or higher is so small ($p$ is essentially 0), we would not expect to see a sample mean of 6 or higher.

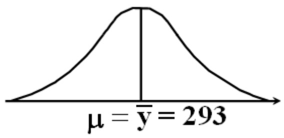
b. $\mu$ and/or $\sigma$ differ from stated values.

1.99 We must assume that the weights of dry seed in the crop of the pigeons are normally distributed. From the printout, the 99% confidence interval is: $(0.61,\ 2.13)$

We are 99% confident that the mean weight of dry seeds in the crop of all spinifex pigeons is between 0.61 and 2.13 grams.

1.101 a. $\mu_{\bar{y}} = \mu = 293$ $\sigma_{\bar{y}} = \frac{\sigma}{\sqrt{n}} = \frac{847}{\sqrt{50}} = 119.78$

b.




c. $P(\bar{y} > 550) = P\left(z > \frac{\bar{y} - \mu_{\bar{y}}}{\sigma_{\bar{y}}}\right) = P\left(z > \frac{550 - 293}{119.78}\right)$

$= P(z > 2.15) = .5 - P(0 < z < 2.15)$

$= .5 - .4842 = .0158$

1.103 Statistix was used to conduct the test desired. The output is shown below:

```
ONE-SAMPLE T TEST FOR TEMPERATURE

NULL HYPOTHESIS: MU =  2550
ALTERNATIVE HYP: MU <> 2550

MEAN           2558.7
STD ERROR      7.1927
MEAN - H0      8.7000
LO 95% CI     -7.5709
UP 95% CI      24.971
T                1.21
DF                  9
P              0.2573

CASES INCLUDED 10    MISSING CASES 0
```

To determine if the mean pouring temperature differs from the target setting, we test:

$H_o : \mu = 2550$
$H_a : \mu \neq 2550$

The test statistic is: $t = 1.21$

The *p*-value is: $p = .2573$

Since $\alpha = .01 < p = .2573, H_0$ cannot be rejected. There is insufficient evidence to indicate that the mean pouring temperature differs from the target setting.

1.105 Let $\mu_{\text{Repeated}}$ = mean height of Australian boys who repeated a grade and $\mu_{\text{Never}}$ = mean height of Australian boys who never repeated a grade.

a. To determine if the average height of Australian boys who repeated a grade is less than the average height of boys who never repeated, we test:

$H_0 : \mu_{\text{Repeated}} = \mu_{\text{Never}}$
$H_a : \mu_{\text{Repeated}} < \mu_{\text{Never}}$

The test statistic is $z = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{-.04 - .30}{\sqrt{\frac{1.17^2}{86} + \frac{.97^2}{1349}}} = -2.64$

The rejection region requires $\alpha = .05$ in the lower tail of the $z$ distribution. From Table 1, Appendix D, $z_{.05} = 1.645$. The rejection region is $z < -1.645$.

Since the observed value of the test statistic falls in the rejection region $(z = -2.64 < -1.645), H_0$ is rejected. There is sufficient evidence to indicate that the average height of Australian boys who repeated a grade is less than the average height of boys who never repeated at $\alpha = .05$.

b. Let $\mu_1 =$ mean height of Australian girls who repeated a grade and $\mu_2 =$ mean height of Australian girls who never repeated a grade.

To determine if the average height of Australian girls who repeated a grade is less than the average height of girls who never repeated, we test:

$H_0 : \mu_1 = \mu_2$
$H_a : \mu_1 < \mu_2$

The test statistic is $z = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{.26 - .22}{\sqrt{\frac{.94^2}{43} + \frac{1.04^2}{1366}}} = .27$

The rejection region requires $\alpha = .05$ in the lower tail of the $z$ distribution. From Table 1, Appendix C, $z_{.05} = 1.645$. The rejection region is $z < -1.645$.
Since $(z = .27 \not< -1.645)$, fail to reject $H_0 : \mu_{Never} - \mu_{\mathrm{Re}\,peat} = 0$. There is insufficient evidence to indicate that the average height of Australian girls who repeated a grade is less than the average height of girls who never repeated at $\alpha = .05$.

c. From the data, there is evidence to indicate that the average height of Australian boys who repeated a grade is less than the average height of boys who never repeated a grade. However, there is no evidence that the average height of Australian girls who repeated a grade is less than the average height of girls who never repeated.

1.107 a. Let $\mu_1 =$ mean mathematics test score for males and $\mu_2 =$ mean mathematics test score for females.

To determine if there is a difference between the true mean mathematics test scores of male and female 8th-graders, we test:

$H_0 : \mu_1 - \mu_2 = 0$
$H_a : \mu_1 - \mu_2 \neq 0$

The test statistic is $z = \frac{(\bar{y}_1 - \bar{y}_2) - 0}{\sqrt{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)}} = \frac{(48.9 - 48.4) - 0}{\sqrt{\left(\frac{12.96^2}{1764} + \frac{11.85^2}{1739}\right)}} = 1.19$

Since no $\alpha$ was given, we will use $\alpha = .05$. The rejection region requires $\alpha / 2 = .05 / 2 = .025$ in each tail of the $z$ distribution. From Table 1, Appendix D, $z_{.025} = 1.96$. The rejection region is $z < -1.96$ or $z > 1.96$.

Since the observed value of the test statistic does not fall in the rejection region $(z = 1.19 < 1.96), H_0$ is not rejected. There is insufficient evidence to indicate there is a difference between the true mean mathematics test scores of male and female 8th-graders at $\alpha = .05$.

b. For confidence coefficient $.90, \alpha = .10$ and $\alpha / 2 = .10 / 2 = .05$. From Table 1, Appendix D, $z_{.05} = 1.645$. The 90% confidence interval is:

$$(\bar{y}_1 - \bar{y}_2) \pm z_{.05}\sqrt{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)}$$

$$\Rightarrow (48.9 - 48.4) \pm 1.645\sqrt{\left(\frac{12.96^2}{1764} + \frac{11.85^2}{1739}\right)}$$

$$\Rightarrow .5 \pm .690 \Rightarrow (-.190,\ 1.190)$$

We are 90% confident that the true differences in mean mathematics test scores between males and females is between $-.190$ and 1.190.

This agrees with the results in part a, even though different $\alpha -$ levels were used. We are 90% confident that the interval contains the true difference in mean scores. Since the interval contains 0, 0 is a likely candidate for the true value of the difference. Thus, we would not be able to reject $H_0$.

c. Since both sample sizes are large, the only assumption necessary is that the samples are independent.

d. The observed significance level of the test in part a is the same as the *p*-value. The *p*-value is the probability of observing your test statistic or anything more unusual, given $H_0$ is true. Thus, the $p - \text{value} = P(z \le -1.19) + P(z \ge 1.19) = .5 - .3830 + .5 - .3830 = .2340$.

e. To determine if the males' test scores are more variable than the female test scores we test:
$H_0 : \sigma_1 = \sigma_2$
$H_a : \sigma_1 > \sigma_2$

The test statistic is $F = \frac{s_1^2}{s_2^2} = \frac{(12.96)^2}{(11.85)^2} = 1.196$

The rejection region requires $\alpha = .05$ in the upper tail of the F distribution with $\nu_1 = n_1 - 1 = 1764 - 1 = 1763$ and $\nu_2 = n_2 - 1 = 1739 - 1 = 1738$. From Table 5, Appendix D, $F_{0.05} > 1.00$ since both degrees of freedom are too large, thus we can reject the null hypothesis and conclude that the males' variability is significantly larger than that of the females: $\frac{\sigma^2_{\text{Males}}}{\sigma^2_{\text{Females}}} = 1$.

1.109 a. When testing the difference in the two sample means of milk prices, we see from the MINITAB printout that the $t = -6.02; p - \text{value} \approx 0;$ which is smaller than a significance level of 1%. Therefore, we can conclude that the two mean milk prices are significantly different at the 1% level.

b. When testing the difference in the two sample means of milk prices that occurred in a sealed bid market, we see from the MINITAB printout that the *p*-value of Levene's Test is 0.264 which is significantly greater than a significance level of 1% or 5%. Therefore, we can conclude that the variances of the two mean milk prices are not significantly different.

Also, the *p*-value for the F test is .048, which means that (unlike with Levene's Test) we would reject the null hypothesis at the 5% level but not at the 1% level. $F = 1.41,$ fail to reject $H_0 : \frac{\sigma_{TRI}^2}{\sigma_S^2} = 1$ at the 1% significance level.