(T)

STAT 213 L05

Solutions to Assignment #1

(2.38) 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. Thus, 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.

(2.52) Assume the data are a sample. The sample mean is:

$$\bar{x} = \frac{\sum x}{n} = \frac{3.2+2.5+2.1+3.7+2.8+2.0}{6} = \frac{16.3}{6} = 2.717$$

The median is the average of the middle two numbers when the data are arranged in order (since $n = 6$ is even). The data arranged in order are: 2.0, 2.1, 2.5, 2.8, 3.2, 3.7. The middle two numbers are 2.5 and 2.8. The median is:

$$\frac{2.5+2.8}{2} = \frac{5.3}{2} = 2.65$$

2.56 a. $\bar{x} = \frac{\sum x}{n} = \frac{7+\cdots+4}{6} = \frac{15}{6} = 2.5$

Median $= \frac{3+3}{2} = 3$ (mean of 3rd and 4th numbers, after ordering)

Mode = 3

b. $\bar{x} = \frac{\sum x}{n} = \frac{2+\cdots+4}{13} = \frac{40}{13} = 3.08$

Median = 3 (7th number, after ordering)

Mode = 3

c. $\bar{x} = \frac{\sum x}{n} = \frac{51+\cdots+37}{10} = \frac{496}{10} = 49.6$

Median $= \frac{48+50}{2} = 49$ (mean of 5th and 6th numbers, after ordering)

Mode = 50

2.66 a. The mean number of ant species discovered is:

$$\bar{x} = \frac{\sum x}{n} = \frac{3+3+...+4}{11} = \frac{141}{11} = 12.82$$

The median is the middle number once the data have been arranged in order: 3, 3, 4, 4, 4, 5, 5, 5, 7, 49, 52.

The median is 5.

The mode is the value with the highest frequency. Since both 4 and 5 occur 3 times, both 4 and 5 are modes.

b. For this case, we would recommend that the median is a better measure of central tendency than the mean. There are 2 very large numbers compared to the rest. The mean is greatly affected by these 2 numbers, while the median is not.

c. The mean total plant cover percentage for the Dry Steppe region is:

$$\bar{x} = \frac{\sum x}{n} = \frac{40+52+...+27}{5} = \frac{202}{5} = 40.4$$

The median is the middle number once the data have been arranged in order: 27, 40, 40, 43, 52.

The median is 40.

The mode is the value with the highest frequency. Since 40 occurs 2 times, 40 is the mode.

(continued)

2.66 (continued)

d. The mean total plant cover percentage for the Gobi Desert region is:

$$\bar{x} = \frac{\sum x}{n} = \frac{30+16+...+14}{6} = \frac{168}{6} = 28$$

The median is the mean of the middle 2 numbers once the data have been arranged in order: 14, 16, 22, 30, 30, 56.

The median is $\frac{22+30}{2} = \frac{52}{2} = 26$.

The mode is the value with the highest frequency. Since 30 occurs 2 times, 30 is the mode.

e. Yes, the total plant cover percentage distributions appear to be different for the 2 regions. The percentage of plant coverage in the Dry Steppe region is much greater than that in the Gobi Desert region.

2.68 a. The mean number of power plants is:

$$\bar{x} = \frac{\sum x}{n} = \frac{5+3+...+3}{20} = \frac{80}{20} = 4$$

The median is the mean of the middle 2 numbers once the data have been arranged in order: 1, 1, 1, 1, 1, 2, 2, 3, 3, 3, 4, 4, 4, 5, 5, 5, 6, 7, 9, 13

The median is $\frac{3+4}{2} = \frac{7}{2} = 3.5$.

The mode is the value with the highest frequency. Since 1 occurs 5 times, 1 is the mode.

b. Deleting the largest number, 13, the new mean is:

$$\bar{x} = \frac{\sum x}{n} = \frac{5+3+...+3}{19} = \frac{67}{19} = 3.526$$

The median is the middle number once the data have been arranged in order:
1, 1, 1, 1, 1, 2, 2, 3, 3, 3, 4, 4, 4, 5, 5, 5, 6, 7, 9

The median is 3.

The mode is the value with the highest frequency. Since 1 occurs 5 times, 1 is the mode.

By dropping the largest measurement from the data set, the mean drops from 4 to 3.526. The median drops from 3.5 to 3. There is no effect on the mode.

(continued)

(2.68) (continued)

c. Deleting the lowest 2 and highest 2 measurements leaves the following:

1, 1, 1, 2, 2, 3, 3, 3, 4, 4, 4, 5, 5, 5, 6, 7

d. The new mean is:

$$\overline{x} = \frac{\sum x}{n} = \frac{5+3+\ldots+3}{16} = \frac{56}{16} = 3.5$$

The trimmed mean has the advantage that any possible outliers have been eliminated.

(2.74) a. $s^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1} = \frac{84 - \frac{20^2}{10}}{10-1} = 4.8889$ $\quad s = \sqrt{4.8889} = 2.211$

b. $s^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1} = \frac{380 - \frac{100^2}{40}}{40-1} = 3.3333$ $\quad s = \sqrt{3.3333} = 1.826$

c. $s^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1} = \frac{18 - \frac{17^2}{20}}{20-1} = .1868$ $\quad s = \sqrt{.1868} = .432$

(2.76) a. Range = 4 – 0 = 4

$s^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1} = \frac{22 - \frac{8^2}{5}}{4-1} = 2.3$ $\quad s = \sqrt{2.3} = 1.52$

b. Range = 6 – 0 = 6

$s^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1} = \frac{63 - \frac{17^2}{7}}{7-1} = 3.619$ $\quad s = \sqrt{3.619} = 1.90$

c. Range = 8 – (–2) = 10

$s^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1} = \frac{154 - \frac{30^2}{18}}{10-1} = 7.111$ $\quad s = \sqrt{7.111} = 2.67$

d. Range = 2 – (–3) = 5

$s^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1} = \frac{29 - \frac{(-5)^2}{18}}{18-1} = 1.624$ $\quad s = \sqrt{1.624} = 1.274$

2.83 a. Range = 1.55 − 1.37 = .18

b. $$s^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1} = \frac{17.3453 - \frac{11.77^2}{8}}{8-1} = .0041$$

c. $s = \sqrt{.0041} = .064$

d. If the standard deviation of the daily ammonia levels during the morning drive-time is 1.45 ppm (compared to .064 ppm in the afternoon drive-time), then the morning drive-time has more variable ammonia levels.

2.90 a. $$\bar{x} = \frac{\sum x}{n} = \frac{206}{25} = 8.24$$

$$s^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1} = \frac{1778 - \frac{206^2}{25}}{25-1} = 3.357 \qquad s = \sqrt{s^2} = 1.83$$

b.

Interval	Number of Measurements in Interval	Percentage
$\bar{x} \pm s$, or (6.41, 10.07)	18	18/25 = .72 or 72%
$\bar{x} \pm 2s$, or (4.58, 11.90)	24	24/25 = .96 or 96%
$\bar{x} \pm 3s$, or (2.75, 13.73)	25	25/25 = 1 or 100%

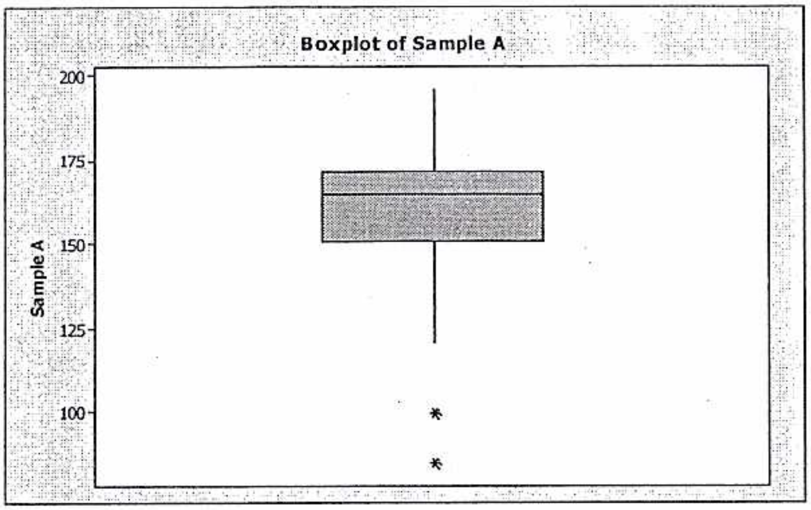
c. The percentages in part **b** are in agreement with Chebyshev's rule and agree fairly well with the percentages given by the Empirical Rule.

d. Range = 12 − 5 = 7

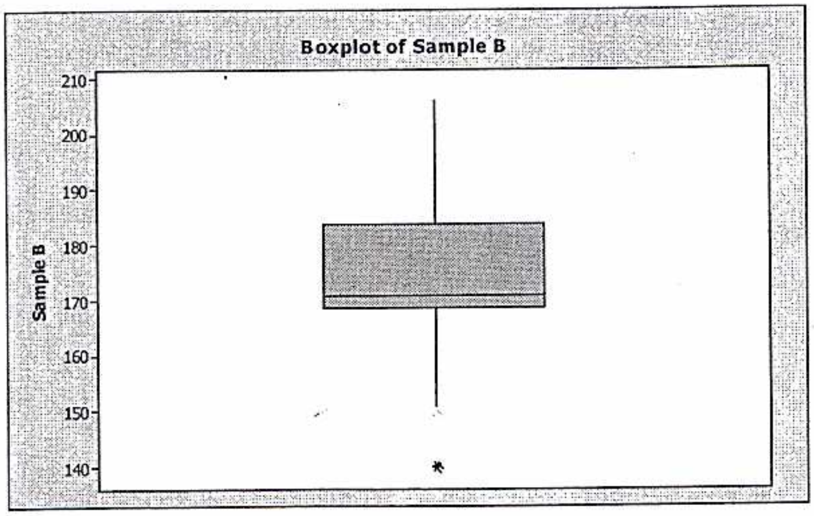
$s \approx$ range/4 = 7/4 = 1.75

The range approximation provides a satisfactory estimate of s.

2.126 a. Using Minitab, the box plot for sample A is given below.



Using Minitab, the box plot for sample B is given below.



b. In sample A, the measurements 84 and 100 are outliers. These measurements fall outside the outer fences.

$$\begin{aligned} \text{Lower outer fence} &= \text{Lower hinge} - 3(\text{IQR}) \\ &\approx 158 - 3(172 - 158) \\ &= 158 - 3(14) \\ &= 116 \end{aligned}$$

In addition, 122 and 196 may be outliers. They lie outside the inner fences. In sample B, 140.4 and 206.4 may be outliers. They lie outside the inner fences.

2.160 a.

$$\sum x = 13 + 1 + 10 + 3 + 3 = 30$$

$$\sum x^2 = 13^2 + 1^2 + 10^2 + 3^2 + 3^2 = 288$$

$$\bar{x} = \sum x = \frac{30}{5} = 6$$

$$s^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1} = \frac{288 - \frac{30^2}{5}}{5-1} = \frac{108}{4} = 27 \qquad s = \sqrt{27} = 5.20$$

b.

$$\sum x = 13 + 6 + 6 + 0 = 25$$

$$\sum x^2 = 13^2 + 6^2 + 6^2 + 0^2 = 241$$

$$\bar{x} = \frac{\sum x}{n} = \frac{25}{4} = 6.25$$

$$s^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1} = \frac{241 - \frac{25^2}{4}}{4-1} = \frac{84.75}{3} = 28.25 \qquad s = \sqrt{28.25} = 5.32$$

c.

$$\sum x = 1 + 0 + 1 + 10 + 11 + 11 + 15 = 49$$

$$\sum x^2 = 1^2 + 0^2 + 1^2 + 10^2 + 11^2 + 11^2 + 15^2 = 569$$

$$\bar{x} = \frac{\sum x}{n} = \frac{49}{7} = 1$$

$$s^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1} = \frac{569 - \frac{49^2}{7}}{7-1} = \frac{226}{6} = 37.67 \qquad s = \sqrt{37.67} = 6.14$$

d.

$$\sum x = 3 + 3 + 3 + 3 = 12$$

$$\sum x^2 = 3^2 + 3^2 + 3^2 + 3^2 = 36$$

$$\bar{x} = \frac{\sum x}{n} = \frac{12}{4} = 3$$

$$s^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1} = \frac{36 - \frac{12^2}{4}}{4-1} = \frac{0}{3} = 0 \qquad s = \sqrt{0} = 0$$

e.

a) $\bar{x} \pm 2s \Rightarrow 6 \pm 2(5.2) \Rightarrow 6 \pm 10.4 \Rightarrow (-4.4,\ 16.4)$

All or 100% of the observations are in this interval.

b) $\bar{x} \pm 2s \Rightarrow 6.25 \pm 2(5.32) \Rightarrow 6.25 \pm 10.64 \Rightarrow (-4.39,\ 16.89)$

All or 100% of the observations are in this interval.

c) $\bar{x} \pm 2s \Rightarrow 7 \pm 2(6.14) \Rightarrow 7 \pm 12.28 \Rightarrow (-5.28,\ 19.28)$

All or 100% of the observations are in this interval.

d) $\bar{x} \pm 2s \Rightarrow 3 \pm 2(0) \Rightarrow 3 \pm 0 \Rightarrow (3,\ 3)$

All or 100% of the observations are in this interval.