Stat/For/Hort 571 — Fall 2003

Assignment 1 — Brief Solutions

1. (a) Evaluate the following:

   i. $\sum_{i=1}^{8} x_i = 43.63$

   ii. $\sum_{i=1}^{8} x_i^2 = 295.8189$

   iii. $(\sum_{i=1}^{8} x_i)^2 = 1903.577$

   iv. $\sum_{i=1}^{8} y_i = 617.51$

   v. $\sum_{i=1}^{8} x_i y_i = 3366.623$

   vi. $(\sum_{i=1}^{8} x_i)(\sum_{i=1}^{8} y_i) = 26941.96$

   vii. $\sum_{i=1}^{8} a x_i = 305.41$ with $a = 7$

   viii. $a \sum_{i=1}^{8} x_i = 305.41$ with $a = 7$

   ix. $\sum_{i=1}^{8} a = 32$ with $a = 4$

   x. the answers are not the same, and that the order in which we do the calculations makes a difference

   xi. the answers are not the same, and that the order in which we do the calculations makes a difference

   xii. the answers are the same, and that the order in which we do the calculations here does not make a difference

   (b)

   $$\bar{x} = 5.45375 \quad \bar{y} = 77.18875$$

   (c)

   $$s_x^2 = 8.2674 \quad s_y^2 = 0.1292$$

   (d) i.

   $$
   \begin{aligned}
   \sum_{i=1}^{8}(x_i - \bar{x}) &= (4.47 - 5.45375) + (9.75 - 5.45375) + (1.41 - 5.45375) \\
   &\quad + (8.26 - 5.45375) + (4.17 - 5.45375) + (7.63 - 5.45375) \\
   &\quad + (5.28 - 5.45375) + (2.66 - 5.45375) \\
   &= -0.98375 + 4.29625 - 4.04375 + 2.80625 - 1.28375 \\
   &\quad + 2.17625 - 0.17375 - 2.79375 \\
   &= 0
   \end{aligned}
   $$

ii.

$$\sum_{i=1}^{8}(x_i - \bar{x})^2 = (4.47 - 5.45375)^2 + (9.75 - 5.45375)^2 + (1.41 - 5.45375)^2$$

$$+ (8.26 - 5.45375)^2 + (4.17 - 5.45375)^2 + (7.63 - 5.45375)^2$$
$$+ (5.28 - 5.45375)^2 + (2.66 - 5.45375)^2$$
$$= 0.9678 + 18.4578 + 16.3519 + 7.8750 + 1.6480$$
$$+ 4.7361 + 0.0302 + 7.8050$$
$$= 57.8718$$

$$\sum_{i=1}^{8} x_i^2 - n(\bar{x})^2 = 295.8189 - 8 \times 5.45375^2 = 57.8718$$

$$\sum_{i=1}^{8} x_i^2 - \frac{\left(\sum_{i=1}^{8} x_i\right)^2}{n} = 295.8189 - \frac{(43.63)^2}{8} = 57.8718$$
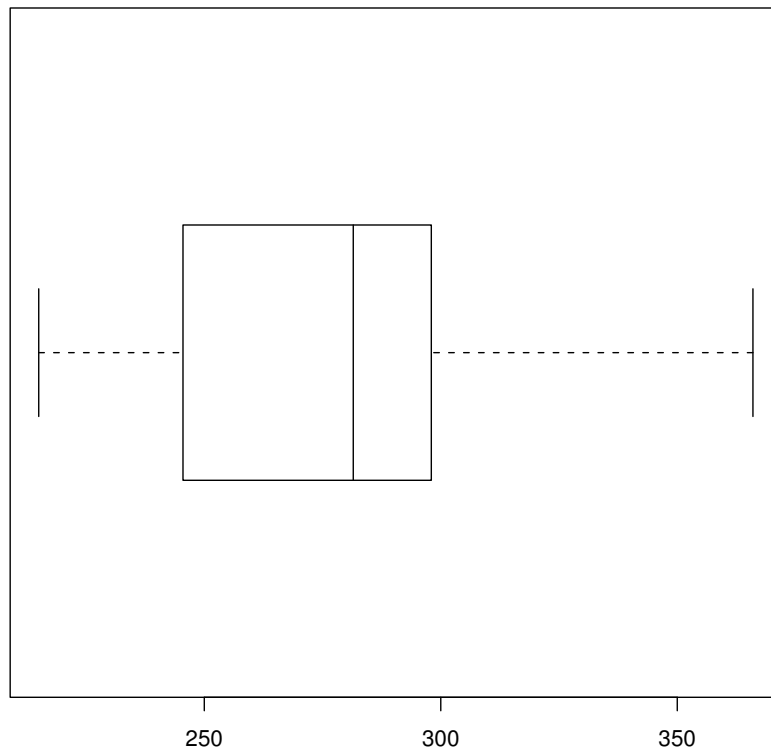
So, we have

$$\sum_{i=1}^{8}(x_i - \bar{x})^2 = \sum_{i=1}^{8} x_i^2 - n(\bar{x})^2 = \sum_{i=1}^{8} x_i^2 - \frac{\left(\sum_{i=1}^{8} x_i\right)^2}{n}$$

2. (a) The decimal point is 1 digit(s) to the right of the |

```
21 | 558
22 | 56
23 | 025
24 | 011456
25 | 0468
26 | 7
27 | 0459
28 | 0003455667
29 | 12334889
30 | 0789
31 | 145
32 |
33 | 09
34 | 6
35 |
36 | 6
```

Mostly bell-shaped.

(b) Median $= x_{[0.50]} = \frac{x_{(26)} + x_{(27)}}{2} = 281.5$

Range $= 366\text{-}215 = 151$

IQR $= x_{[0.75]} - x_{[0.25]} = \frac{x_{(39)} + x_{(40)}}{2} - \frac{x_{(13)} + x_{(14)}}{2} = 298 - 245.5 = 52.5$

$x_{[0.40]} = x_{(21)} = 274$

Box plot for 2(c)

3. The stem-and-leaf display is quite informative:

```
The decimal point is 2 digit(s) to the right of the |

leaf(Healthy)| stem | leaf(infected)
-----------------------------------------
           |   0  | 2
           |   1  | 78
           |   2  | 0
           |   3  | 4
           |   4  | 14
           |   5  |
           |   6  |
           |   7  | 1344
           |   8  | 9
           |   9  | 9
        7  |  10  |
           |  11  | 2
           |  12  |
      236  |  13  |
        5  |  14  |
        2  |  15  |
        5  |  16  |
        6  |  17  |
        7  |  18  |
```

Summary statistics:

```
           Healthy          Infected
Median      1446             577.5
Mean        1480             549.4
sd          248.9849         343.4586
cv          0.1682331        0.6251196
```

Tentatively, we can draw a conclusion that the virus can reduce the stem volume.

4. In this case,the Population should be the students in that university. And we can extrapolate the inference to students in all the universities because age and whether the people of the same age level work or not might be important factors for determing whether the pain killer had any side effects of nausea. If my goal were to make inference to the public at large, I might make my sample comprised of the persons from many age level and from different field.