AMS 310 Homework 1 Solutions

Prof. F. Rispoli

1. (20pts) The monthly average New York City temperature in Fahrenheit are given below.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
High	38	42	50	61	71	79	84	83	75	64	54	43
Low	27	29	35	45	54	64	69	68	61	50	42	28

a) (5pts) Find the mean and median of the *low* temperatures.

b) (5pts) Draw a time plot of the average *high* temperatures.

c) (5pts) Compute the sample variance of the *high* temperatures using formula 1-1 AND 1-2. (Textbook Page 17). This is also known as the shortcut formula.

d) (5pts) Compute the sample variance of the *low* temperatures using formula 1-1 AND 1-2. (Textbook Page 17)

5 points for each part

a) $\bar{x}$ =47.67 and Median = 47.5

b)
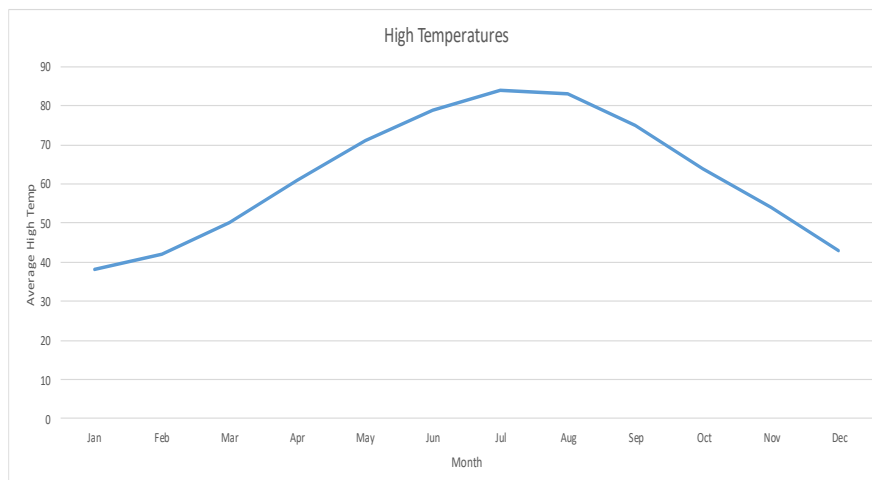
c) s^2 = 274 (show work doing it both ways)

e) s^2 = 247.33 (show work doing it both ways)

2. (20pts) In a study of a parasite in humans and animals. Researchers measured the lengths (in μm) of 90 individual parasites of certain species from the blood of a mouse. The measures are shown in the following table:

Length	19	20	21	22	23	24	25	26	27	28	29
Frequency	1	2	11	9	13	15	13	12	10	2	2

a) (5pts) Find the sample median and quartiles.
b) (5pts) Compute the sample mean and sample standard deviation.
c) (5pts) Compute the sample range and interquartile range.
d) (5pts) Find the 85th percentile.

(a) Q1: $np = 90 * 0.25 = 22.5$,round up to 23.The 23rd value in the sorted list is 22.

Q2(median): $np = 90 * 0.5 = 45$, $\frac{x_{(45)}+x_{(46)}}{2} = \frac{24+24}{2} = 24$.

Q3: $np = 90 * 0.75 = 67.5$, round up to 68.The 68th value in the sorted list is 26.

(1 pt for finding correct nps;

1 pt for "round up";

3 pts for final answers,1 pt each.)

(b) Sample mean: $\bar{x} = \frac{19*1+20*2+21*11+22*9+23*13+24*15+25*13+26*12+27*10+28*2+29*2}{90} = 24.09$ (2 pts)

Sample variance: $s^2 = \frac{((19-24.09)^2*1+(20-24.09)^2*2+\cdots+(28-24.09)^2*2+(29-24.09)^2*2}{89}$ (2 pts)

Sample standard deviation: $s = \sqrt{s^2}$ (1 pt)

(Students receive full credits of part b as long as they show the correct formula. No need to calculate the final answer for part b only)

(c) Sample range: $29 - 19 = 10$ (2 pts)
Interquartile range: $IQR = Q_3 - Q_1 = 26 - 22 = 4$ (3 pts)

(d) $np = 90 * 0.85 = 76.5$,round up to 77.
$x_{(77)} = 27$

(3 pts for find correct np

2 pts for the final answer).

3. (20pts) The following stem-and-leaf plots shows scores on a statistics final exam:

2	9
3	38
4	3688
5	116
6	14678
7	7889
8	234556699
9	001347
10	0

a) (4pts) Compute the sample median and quartiles.

b) (3pts) Compute the sample mean, sample variance, and sample standard deviation.

c) (4pts) Compute the 66th percentile.

d) (3pts) Compute the sample range and the interquartile range.

e) (3pts) Construct a boxplot.

f) (3pts) Are there any outliers? Give the boundaries for determining outliers

29,33,38,43,46,48,48,51,51,56,61,64,66,67,68,77,78,78,79,82,83,84,85,85,86,86,89,89,

90,90,91,93,94,97,100

(a) Q1: $np = 35 * 0.25 = 8.75$,round up to 9.

$Q_1 = x_{(9)} = 51$ (1 pt)

Q2(median): $np = 35 * 0.5 = 17.5$,round up to 18.

$Q_2 = x_{(18)} = 78$ (1 pt)

Q3: $np = 35 * 0.75 = 26.25$,round up to 27.

$Q_3 = x_{(27)} = 89$ (1 pt for "round up", 1pt for correct final answer)

(b) Sample Mean: $\bar{x} = \frac{29+33+38+43,...+94+97+100}{35}$

Sample Variance: $s^2 = \frac{(29-\bar{x})^2+(33-\bar{x})^2+,...+(100-\bar{x})^2}{35-1}$

Sample Standard Deviation: $s = \sqrt{s^2}$

1 pt for each, no need to give final answers.

(c) $np = 35 * 0.66 = 23.1$, round up to 24.
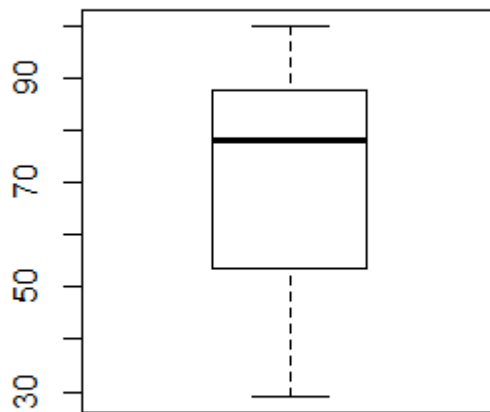
$x_{(24)} = 85$

(2 pts for correct np,

1 pt for rounding up,

1 pt for the final answer.)

(d) Sample Range: $100 - 29 = 71$ (1 pt)

Interquartile Range: $IQR = 89 - 51 = 38$ (2 pts)

(e)

(1 pt for correct min and max,

1 pt for correct Q_1 and Q_3,

1 pt for correct median.)

(f)

$Q_1 - 1.5(IQR) = 51 - 1.5 * 38 = -6$ (1 pt)

$Q_3 + 1.5(IQR) = 89 + 1.5 * 38 = 146$ (1 pt)

There're no outliers.(1 pt)

4. (20pts) The following are the amount of radiation received at a green house.

24 39 7 48 16 29 34 20 43 18 12 19 22 27 29 10 37 16 23 32

    a) (4pts) Construct a stem-and-leaf plot.

    b) (4pts) Compute the sample mean and standard deviation.

    c) (4pts) Compute the sample median and quartiles.

    d) (4pts) Construct a boxplot.

    e) (4pts) Are there any outliers? Give the boundaries for determining outliers.

(a)
```
0 | 7
1 | 026689
2 | 0234799
3 | 2479
4 | 38
```
(1 pt for each row. Total 5 pts)

(b)

Sample Mean: $\bar{x} = \frac{1}{20}\sum_{i=1}^{20}x_i = \frac{505}{20} = 25.25$

Sample Variance: $s^2 = \frac{\sum_{i=1}^{n}x_i^2 - \left(\sum_{i=1}^{n}x_i\right)^2/n}{n-1} = 124.30$

Sample Standard Deviation: $s = \sqrt{s^2} = \sqrt{124.3} = 11.15$

(Sample Mean (2 pts) :1 pt for correct formula

                    1 pt for correct final answer)

(Standard Deviation (3 pts) :2 pts for correct sample variance

                       1 pt for correct standard deviation

                       1 pt partial credit may be awarded if the student

                       shows a correct sample variance formula.)

(c) Recipe for calculating the sample 100p-th percentile:

      Step1: Order the n observations from the smallest to largest.

Step2: Find np.

Step3: If np is an integer, say k, calculate $\frac{X_{(k)}+x_{(k+1)}}{2}$.

If np is not an integer, round it up to the next integer and find the

Corresponding ordered value.

Ordered Data:

7, 10, 12, 16, 16, 18, 19, 20, 22, 23, 24, 27, 29, 29, 32, 34,

37, 39, 43, 48

First Quartile: Q1: $np = 20 * 0.25 = 5$,is an integer.

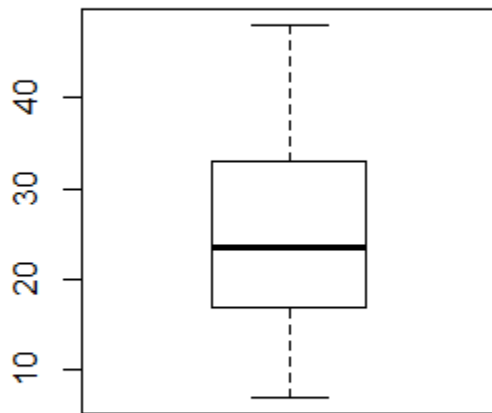$Q_1 = \frac{x_{(5)}+x_{(6)}}{2} = \frac{16+18}{2} = 17$ (1 pt)

Second Quartile Q2(median): $np = 20 * 0.5 = 10$,is an integer.

$Q_2 = \frac{x_{(10)}+x_{(11)}}{2} = \frac{23+24}{2} = 23.5$ (1 pt)

Third Quartile Q3: $np = 20 * 0.75 = 15$,is an integer.

$Q_3 = \frac{x_{(15)}+x_{(11)}}{2} = \frac{32+34}{2} = 33$ (1 pt)

(d)

(e)

$$IQR = Q_3 - Q_1 = 33 - 17 = 16 (1 \text{ pt})$$
$$Q_1 - 1.5(IQR) = 17 - 1.5 * 16 = -7 \ (1 \text{ pt})$$

$$Q_3 + 1.5(IQR) = 33 + 1.5 * 16 = 57 \ (1 \text{ pt})$$

There are no observations less than -7 or greater than 57. There're no

outliers.(1 pt)

5. (20pts) Consider the following data and answer the questions **using R**. Make sure your answer includes both R command and the final answer/diagram.

55 61 94 94 69 77 68 54 85 77 92 92 81 73 69 81 75 84 70 81 81 89 59 72 82 62

  a) (3pts) Draw a histogram.
  b) (3pts) Find the sample mean and sample variance.
  c) (4pts) Find the median, Q1, and Q3.
  d) (3pts) Find the 65th percentile.
  e) (4pts) Find the interquartile range.
  f) (3pts) Draw a boxplot.

Please note that NO partial credit will be awarded to this question. Student's answer must include BOTH R command AND the final answer/diagram.
(a) (3pts)

```
> q5 <- c(55, 61, 94, 94, 69, 77, 68, 54, 85, 77, 92, 92, 81,
73, 69, 81, 75, 84, 70, 81, 81, 89, 59, 72, 82, 62)
> hist(q5)
```

Histogram of q5
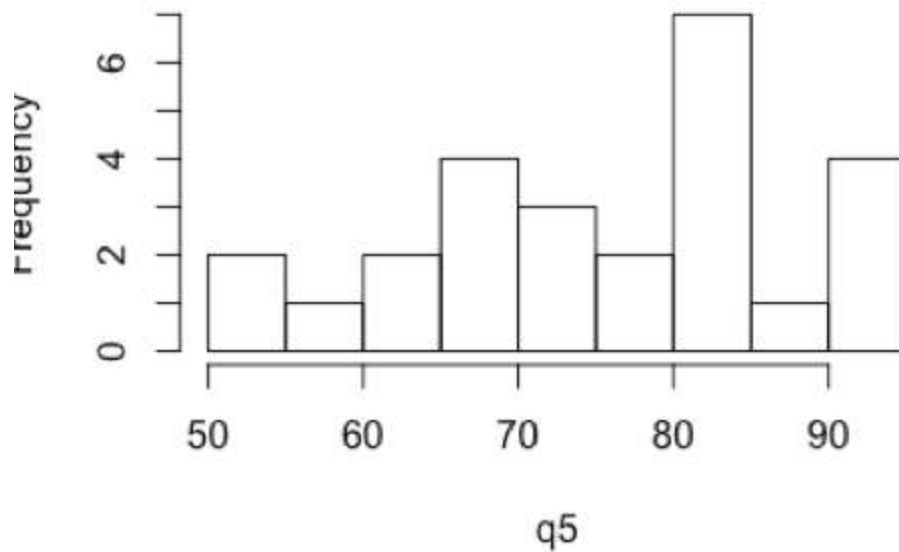
(b) (3pts)
```
> mean(q5)
[1] 76.03846
>
> var(q5)
[1] 138.8385
```

(c) (4pts)

```
> quantile(q5)
   0%   25%   50%   75%  100%
 54.0  69.0  77.0  83.5  94.0
```

(d) (3pts)

```
> quantile(q5, 0.65)
65%
 81
```

(e) (4pts)

```
> IQR(q5)
[1] 14.5
```

(f) (3pts)

```
> boxplot(q5)
```