

Solutions to Homework #1

Stat 371-003

1. I installed R on my laptop, running Mac OS X.
I installed R version 2.7.2

I also went around to various Info Labs on campus.

At Steenbock, on Windows XP Pro, I could start R by doing:

Start: Programs: Courseware: R: R 2.6.1

This was R ver 2.6.1

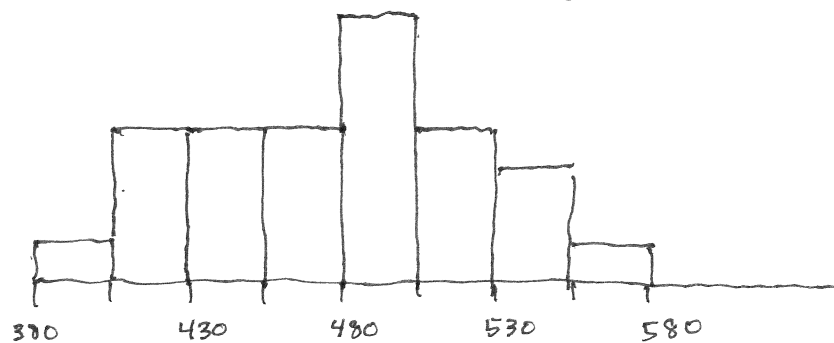
At CALS, I also found R on the Macs, which were running Mac OS X (ver 10.4.11). I invoked R by going to the Applications folder and clicking on "R".

This was R ver 2.3.0.

2.9 (pg 25)

A variety of solutions are allowable,
depending on your choice of cutpoints.

380 - 404		1
405 - 429		4
430 - 454		4
455 - 479		4
480 - 504		7
505 - 529		4
530 - 554		3
555 - 579		1
		28



2.20 (pg 31)

The mean is the sum divided by 9.

$$\frac{3.89 + 3.51 + \dots + 3.27}{9} = \frac{31.43}{9} \approx \boxed{3.49}$$

To determine the median, sort the numbers and find the middle one.

3.21 3.24 3.27 3.31 3.36 3.51 3.67 3.89 3.97

The median is 3.36.

2.28 (pg 32)

(a) The median is the point at which the area to the right is the same as the area to the left, say a bit less than 50.

(b) The distribution has a slightly longer left tail, so the mean would be a bit less than the median, say 45.

If you want to be more precise, you could measure the relative heights of each rectangle, and assume all data is at the midpoint of the respective bin.

<u>bin</u>	<u>height</u>	<u>bin</u>	<u>height</u>
2.5	1	37.5	8
7.5	2	42.5	9
12.5	2	47.5	12
17.5	3	52.5	11
22.5	4	57.5	17
27.5	6	62.5	8
32.5	8	67.5	9
		72.5	1

Then

Median ≈ 47.5

mean ≈ 45.5

2.33 (pg 39)

First sort the values

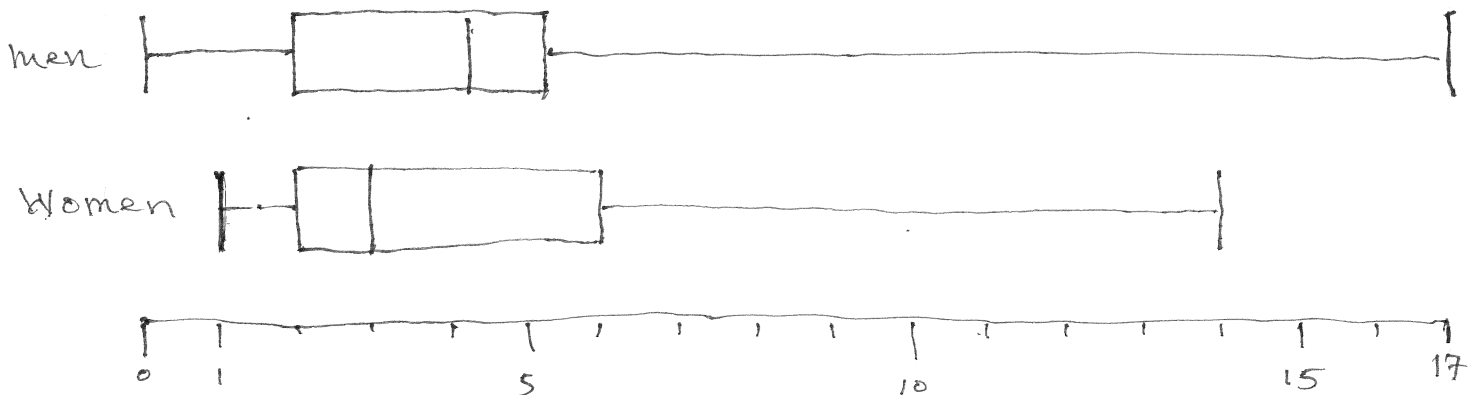
men: 0 1 2 2 3 4 4.5 4.5 5 6 8 17
 women: 1 1.5 1.5 2 3 3 3 4 5 6 8 13 14

Now find the min, max, median, 25%ile, 75%ile
 of course, the min, max & median are easy.

And the 25 and 75 %iles are easy for the women.

For the men, we'd place the 75th percentile a quarter of the way between 5 and 6.

	min	25%	median	75%	max
men	0	2	4.25	5.25	17
women	1	2	3	6	14



2.42 (pg 48)

(a) That particular sample $-3, -1, 0, 2, 2$ would work, since the mean is

$$\frac{-3 + (-1) + 0 + 2 + 2}{5} = \frac{0}{5} = 0.$$

(b) The SD is $\sqrt{\frac{(-3)^2 + (-1)^2 + (0)^2 + 2^2 + 2^2}{4}}$

$$= \sqrt{\frac{9 + 1 + 4 + 4}{4}} = \sqrt{18/4} \approx 2.12$$

(c) Everyone should get the same answer for part (b), since the SD depends just on these deviations.

2.54 (pg 50)

The bulk of the data are between 10 and 75, and the distribution is reasonably symmetric, so the mean is about $\frac{10+75}{2} = 42.5$, and the SD is about $\frac{75-10}{4} \approx 16$.

To be more precise, we could measure the heights of the bins and use the technique in the solution for problem 2.28.

This gives estimates mean ≈ 44 , SD ≈ 20 .