STAT 270

I had intended to show you the workings of exercises 1.3.37 and 1.4.49, but ran out of time on Friday. I was slowed down by the need to verbalize everything instead of writing on the white board! Now I have the pens, so this is a one-timer. Here are the notes for 1.3.37 and 1.4.49.

1.3.37

> data1.3.37=c(6.5,12.0,14.9,10.0,10.7,7.9,21.9,12.5,14.5,9.2) > my.dotplot(data1.3.37)



> mean(data1.3.37)
[1] 12.01
> median(data1.3.37)
[1] 11.35

Not much difference but note how the one large value pulls the mean up past most of the points. For skewed data, median is better descriptive measure I would use it here -11.35 million km.

Incidentally, here is an example of a measurement exhibiting unexplained variation, and important for scientific study - presumably snow cover is important for affecting the proportion of sunlight that is reflected back to space.

Another possibility with this data is to use the 10% trimmed mean, which in this case is 11.46. However, the trimmed mean is so seldom used it is hardly worth studying. There is perhaps some theoretical value in it – note that the median is almost a 50% trimmed mean.

> trmeanx=mean(data1.3.37,.1)
> trmeanx
[1] 11.4625

1.4.49

data1.4.49=c(2.75,2.62,2.72,3.85,2.34,2.74,3.93,4.21,3.88,4.33,3.46,4.52,2.43,3.65,2.78, 3.56,3.01)

```
> sumx=sum(data1.4.49)
> sumxsq=sum(data1.4.49^2)
> sumx
[1] 56.78
> sumxsq
[1] 197.6948
> Sxx=sumxsq-((sumx)^2)/17
> Sxx
[1] 8.0496
> s.sq=Sxx/16
> s.sq
[1] 0.5031
> s=s.sq^.5
> s
[1] 0.7092954
```

But this formula pp39-40 does not really make s understandable. Instead I would do this:

```
> meanx=mean(data1.4.49)
> sq.devx=(data1.4.49-meanx)^2
> sdx.sq=sum(sq.devx)/16
> sdx.sq
[1] 0.5031
> sdx=sdx.sq^.5
> sdx
[1] 0.7092954
```

Note that the standard deviation of the data, 71, is quite close to the average absolute deviation, which in this case is .63. Think of the standard deviation as a "typical" deviation from the mean.

> madx=mean(abs(data1.4.49-meanx)) > madx [1] 0.6270588