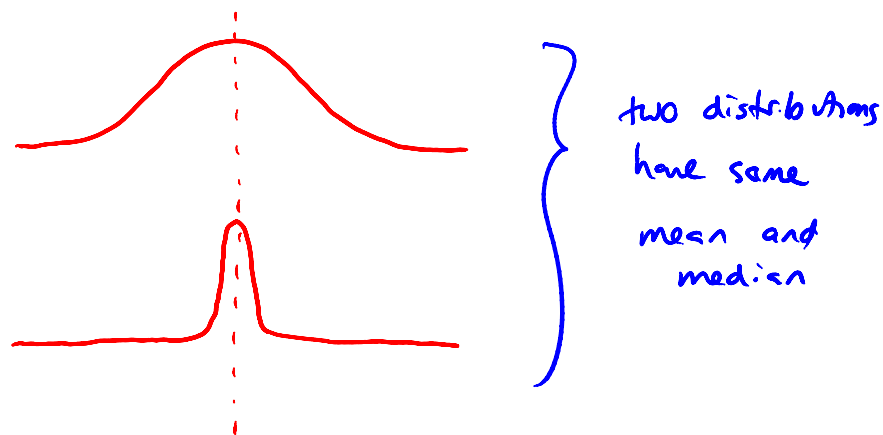


Section 3.2: Measures of Dispersion

Monday, May 06, 2013
1:31 PM

When you are looking at a distribution, you might want to quantify also the spread or variability of the distribution



for example, how "wide" the peak is if the distribution is unimodal

note: when do you want a small spread?
when trying to manufacture identical objects

when do you want a large spread?
when trying to make distinctions
- ranking objects

easiest measure to calculate:

range - difference between the largest and smallest measurements

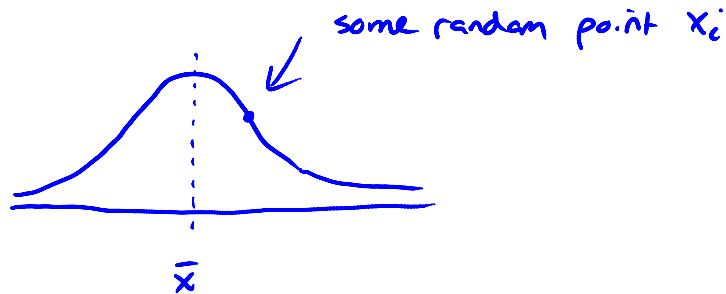
good part: easy to calculate

bad part: not very useful

strongly influenced by
outliers

annoying measure to calculate:

variance:



consider the data point x_i on the above distribution.

how far away from the mean is it? $x_i - \bar{x}$

(note: if you sum $\sum (x_i - \bar{x})$, you get zero)

but if you sum $\sum (x_i - \bar{x})^2$ (all positive)
then the result is a measure of
how far away from the mean the
points are

population variance:

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N}$$

N = size of population

sample variance

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

n = size of sample

standard deviation:

population:

$$\sigma = \sqrt{\sigma^2}$$

sample

$$s = \sqrt{s^2}$$

of what use is s ?

how do you determine placements in Olympic events such as the heptathlon / decathlon