Time Examined

Day 11

Today

I How good is a clock?

II Inventing measures of quality

I Recall our tripartite model of a clock:

(1) A physical oscillator

(2) An energy source to keep our oscillator moving

and

(3) A counter

We have begun to understand many features of the first two components.

Going forward it will also be helpful to characterize the third component. In particular we would like to have clear answers to the following questions: What makes one clock better than another? How do we measure how good a clock is? Is there a quantitative way to characterize the quality of a clock? Can we combine the measurements of more than one clock to get a more accurate measure of time?

There are several themes that run through the scientific characterization of how well we measure a quantity or build an instrument. We turn to those themes now.

II Using the pitching machine data sets invent a method for computing a single number that characterizes how good each machine is. You should use the same method for each machine.

Suppose I now relax the constraint

Center of rectangle to x via pythagoras
broke ties favoring area smaller
Average Euclidean distance
Area of circumscribing rectangle

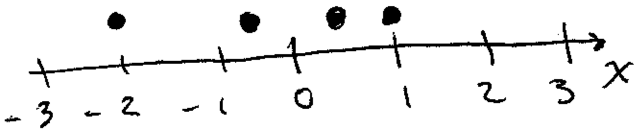
Machine	Grp 1	Grp 2	Grp 3	Grp 4
Ranco	2	3	3	
Big Bruiser	3	4	2	
Fireball	1	2	1	
Smyth's	4	1	4	

and allow you two numbers to characterize each machine. What would they be? How would you use them to rank the pitching machines?

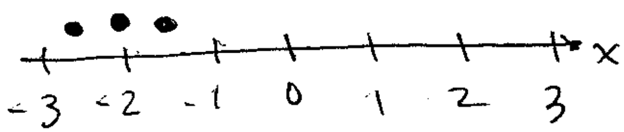
In our group discussion we touched on:
Averaged path length to x
path length to central pitch

Suppose that instead of pitching machines we had a tennis ball server and we fixed its launch angle and speed so that the only variable was the horizontal angle at which it launched balls. Say that in 3 different settings we got the following results:

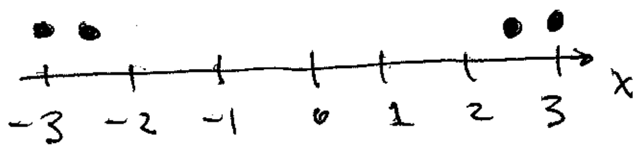
Setting 1



Setting 2



Setting 3



The average or mean can often be used as a representative of a sample of values. It is defined by

$$\text{ave}(a_1, a_2, \ldots, a_N) = \frac{a_1 + a_2 + \cdots + a_N}{N} \equiv \bar{a}$$

For example, for the tennis settings we have

Setting 1: $\text{ave}(-2, -0.5, 0.5, 1) = \dfrac{-2 - 0.5 + 0.5 + 1}{4}$

$= -\frac{1}{4} = -0.25 = \bar{x}_{s1}$

Can we adapt any of your inventions to the tennis server?

Scientists have found a set of measures that are quite flexible and useful in characterizing finite samples of data like the ones we have been exploring here.

Setting 2: $\text{ave}(-2.5, -2, -1.5)$

$= \dfrac{-2.5 - 2 - 1.5}{3} = -2 = \bar{x}_{s2}$

Setting 3: $\text{ave}(-3, -2.5, 2.5, 3)$

$= \dfrac{-3 - 2.5 + 2.5 + 3}{4} = 0 = \bar{x}_{s3}$

Notice that the average as a representative, isn't always very close to any members of the sample.

While these values do capture a central tendency of the data, it is unfortunate that they don't get at the spread of the data. To capture this spread we introduce the variance and the standard deviation

Variance $s^2 = \frac{1}{N-1} \sum_i (x_i - \bar{x})^2$

and

standard deviation $s = \sqrt{s^2} = \sqrt{\frac{1}{N-1} \sum_i (x_i - \bar{x})^2}$

or

$$S_{s2}^2 = \frac{1}{2}\left[(-0.5)^2 + (0)^2 + (0.5)^2\right]$$

$$= \frac{1}{4}$$

Then $S_{s2} = \sqrt{S_{s2}^2} = \frac{1}{2}$.

Meanwhile,

$$S_{s3}^2 = \frac{1}{4-1}\left[(-3-0)^2 + (-2.5-0)^2 + (2.5-0)^2 + (3-0)^2\right]$$

$$= \frac{1}{3}\left[18 + 12.5\right] = \frac{30.5}{3} = 10.1\overline{6}\overline{6}$$

or $S_{s3} = 3.189$

The standard deviation is just the square root of the variance.

Let's illustrate these calculations with settings 2 and 3:

$$\bar{x}_{s2} = -2$$

So

$$S_{s2}^2 = \frac{1}{3-1}\left[(-2.5+2)^2 + (-2+2)^2 + (-1.5+2)^2\right]$$

Indeed the variance and std. dev. capture the big difference in spread between the two settings.