Module 1

Linear Transformation of mean (Part 1)

Previously we looked at mean and variance to describe statical properties of datasets.

In this session we will discuss what happens to mean and variance when we linearly transform the data set, that mean we shift it around or stretch it.

Let's look at dataset A, which is given by blue dots over here. (see pic)

$D = \{-1, 2, 3\}$

Lets compute the mean value of this dataset to start with

$$E[D] = \frac{-1 + 2 + 3}{3} = \frac{4}{3} \quad \text{Ⓐ}$$

I am indicating the mean value of this data set by blue star (topic)

Now question is, what happens to dataset, when we shift it?

Let have a look at shift by 2 tonight (see pc)

We end up with the red dots now.
What happens to mean value, of the shifted dataset, original data set to right.

The answer is, the mean also shifts by 2.

Lets define D prime to be dataset: $1, 4, 5$, which is $D + 2$.

$$D' = \{1, 4, 5\} = D + 2$$

$$E[D'] = \frac{1+4+5}{3} = \frac{10}{3} = \frac{4}{3} + 2$$

$\frac{10}{3}$ becomes $\frac{4}{3}$ which comes from Ⓐ (page 2)

Plus 2 which is the shift of the rest of the original data set.

So we can generalize it now, to general shifts

$$E[D + a] = a + E[D]$$

(a is constant factor)

Up to this point we knew what happens to mean of data when we shift data set,

But whathappens when we stretch the dataset?

Lets say we also stretch by factor of 2.

That means we multiply every individual component in data set D by 2.

Then we end up with dataset indicated by ~~dots~~ red dots here (see pic)

And we define D doubleprime to be -2,4,6

$$D'' = \{-2, 4, 6\}$$

And if we now compute the mean of D'', then we end up with:

$$E[D''] = \frac{-2+4+6}{3} = \frac{8}{3} = \frac{4}{3} \cdot 2$$

and reunite $\frac{8}{3}$ as $\frac{4}{3}$

times 2 which is scaling factor (obviously)

In general we can write:

$$E[\alpha D] = \alpha E[D]$$

Putting everything together we can shift and scale our original dataset in new mean of the linear transformation

So if we compute the expected value of α times $D + a$ we get α times expected value of D plus a, where α is the scaling factor and a the shift

$$E[\alpha D + a] = \alpha E[D] + a$$

In this session, we looked at the effect of the linear transformations on mean value of Dataset

In next session we will look at exactly the same for variances.