Module 1

Linear Transformation of mean (Part 1)

Previously we looked at mean and variance to describe statistical properties of dataset.

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

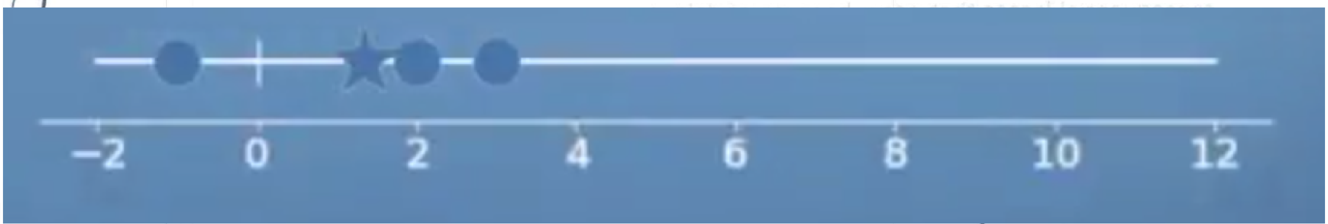
Lets look at dataset D, which is given by blue dots over here. (see pic)

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

Lets compute the mean value of this dataset to start with

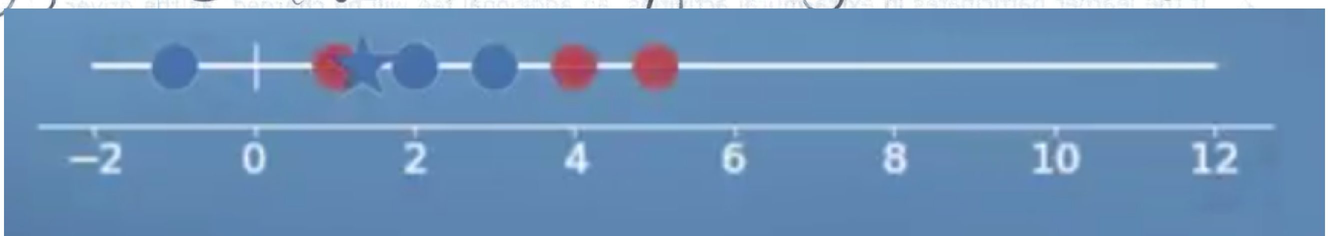
$E[D] = \frac{-1+2+3}{3} = \frac{4}{3}$ (A)

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



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

Lets have look at shift by 2 to right (see pic)



we end up with the red dots now.

What happens to mean value of the dataset, if we shift dataset, original dataset 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 came from (A) (page 2)

Plus 2 which is the shift of the offset 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 know what happens to mean of data when we shift data set, But what happens when we Stretch the dataset?

Lets say we also Stretch by factor of 2.

That means we multiply every individual Component in dataset D by 2.

Then we end up with dataset indicated by ~~[illegible]~~ 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 rewrite $8/3$ as $4/3$

time 2 which is Scaling factor (for Stretching)

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 is shift

$$\boxed{E[\alpha D+a] = \alpha E[D]+a}$$

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

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