

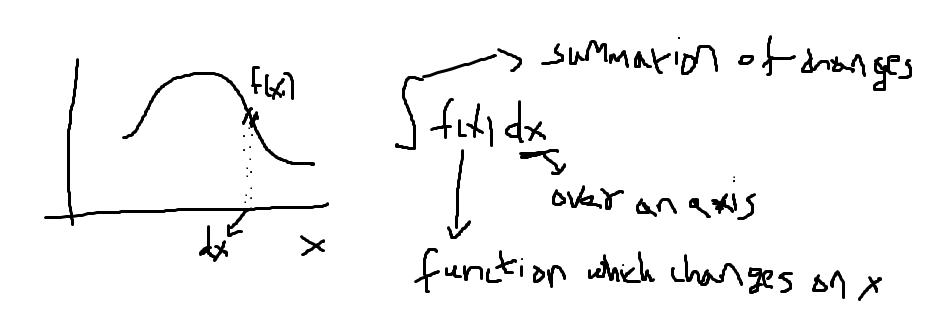
Two observations:

1. x
2. Densities with respective x.

As x changes the probable value it attains in the distribution changes. It senses for an xi that what contribution it takes for an overall x in the distribution. Distribution is over x but not on p(x)

So, the study of probability function took place to identify what probable values it takes for an x. For now, let’s assume its p(x).

**Integration**:

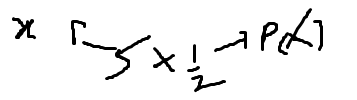


I hope you understand the intention of having dx. This is to entitle the changes happening to be on f(x) and sum up all the changes along the curve(f(x)).

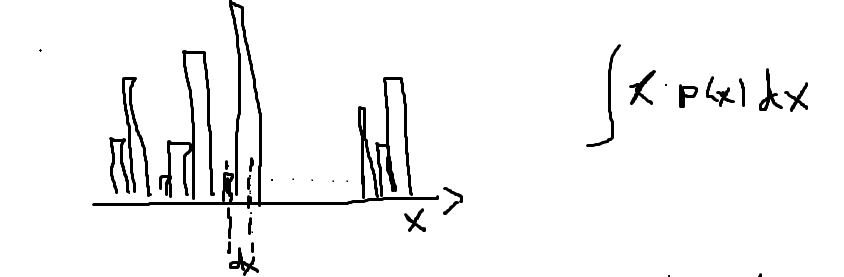
That is the reason it is very important to understand where is the intention moving->is it along x or is it along f(x)

Now let’s understand for our case study:

To recall we see the way we change the random x, we tend to see its change in its contribution for overall random distribution. The term x.p(x) is the probability contribution of your x on the distribution. It behaves like:



That is the reason x.p(x) is the term which moves along x. So mathematically let’s take the concept of integration as above shown



The small dx is taken to understand the change of how the property xi.p(xi) occurs on x and integrating it shows the average sum of these changes.

As we are clear that p(xi).xi is the probability contribution of xi over all the random variable x, integrating the changes is applied on the x.

This is the mean.

Just to pass down that Law of large numbers defines the study of these property on an infinite experimentation so that x lies over infinite samples and mean is always brought to a steady, stable value. So, to represent:

