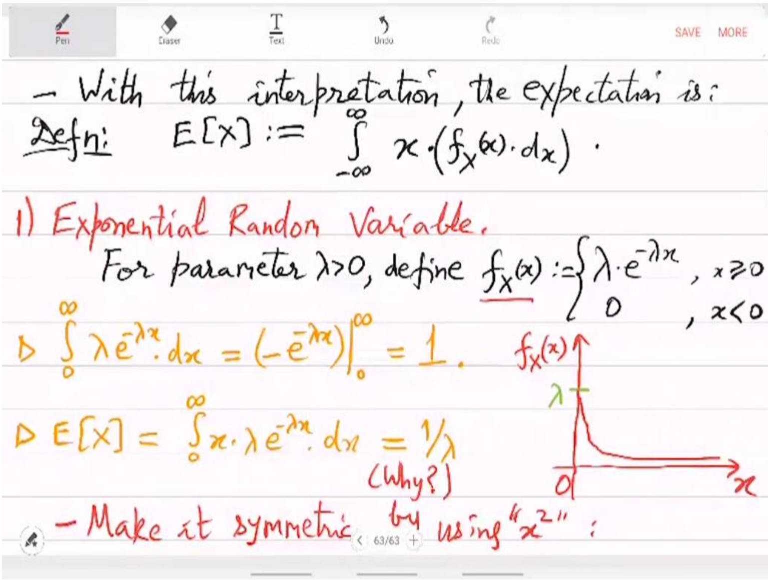
Probability for Computer Science
Prof. Nitin Saxena
Department of Computer Science and Engineering
Indian Institute of Technology - Kanpur

Module - 3
Lecture - 12
Continuous Random Variables

(Refer Slide Time: 00:13)



So, this is a general theory about how to deal with random variables which are actually taking values, let us say, on the whole real line or even on an interval like 0 to 1, because there are infinitely many values. So, we make sense of it through probability density function. And let us study now some named continuous random variables. So, first is exponential random variable.

So, for parameter lambda, define the probability density function to be lambda times e raised to minus lambda x for x non-negative, and 0 for x negative. That is the density function. So, what this is doing is, let us also assume for sanity that this is non-negative. So, what this is doing is, as x is increasing, the density is tending to 0. So, you can draw the following picture. So, x increases, and this is the value of f sub X.

And so, at x equal to 0, clearly the value is lambda, somewhere here. And then it falls. It keeps falling and it actually goes to 0 very slowly, as x tends to infinity. So, this is what the function is doing, it is exponentially falling. Why is it exponentially falling? We say exponential because it is 1 over e to the lambda x. It is falling like e raised to x, which is falling very fast. So, what you should check is its value over the whole real line.

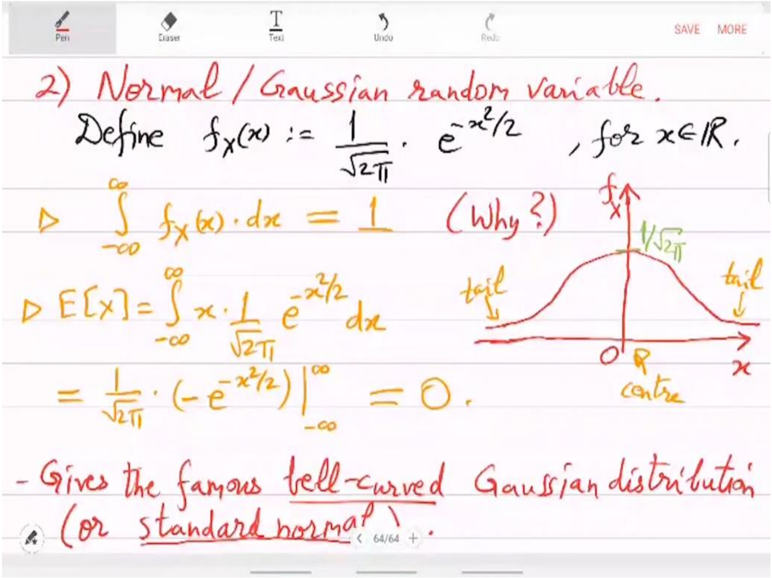
So, that is lambda e raised to minus lambda x times dx, this integral, which is just minus e to the minus lambda x, going from 0 to infinity; (()) (03:33) minus infinity to infinity. Yeah, let me correct this. It is not right. So, this f is actually 0 for negative x, right? So, this is only going from 0 to infinity. At infinity, this e raised to minus lambda x vanishes; and at 0 it is minus 1.

So, you get value 1, which is important, which shows that this is a probability density function, because the sum is, the whole integral is, integral over the real line is 1. Second thing is expectation of X. What is that? So, where do you expect the random variable big X to be around? If it follows this PDF, where is it around? Very clear; so, let us do that calculation or at least give the expression.

So, expression is again 0 to infinity x times lambda e raised to minus lambda x times dx. And I will not do this, but this is supposed to give you 1 over lambda. You just check this. So, that is the expected value. So, essentially, if you think of lambda to be 1, then this density function is e raised to minus x. It is really, I mean, truly exponentially falling. And the expected value of the random variable is to be around; you expect the random variable to be around 1.

The probability is kind of larger in the vicinity of 1. And when you go away from 1, then it becomes much smaller. That is the intuition. There is this problem of x being positive and negative in this definition, or, I mean, this definition is asymmetric; so, let us make it symmetric. So, let us make it symmetric by using x square. Using means, instead of e raise to minus where, so that it looks symmetry both ways, positive line and negative; positive half, negative half. That is a better picture.

(Refer Slide Time: 06:21)



And then you get something amazing, which is called normal or Gaussian random variable. So, now you define the density function to be 1 over square root of 2 pi e raised to minus x square by 2, over the whole, same description over the whole real line. The square root 2 pi and 2, we have put just to make it satisfy the integral equal to 1. You can check that. You have to check that this is equal to 1. What is the picture of this? So, let me draw that.

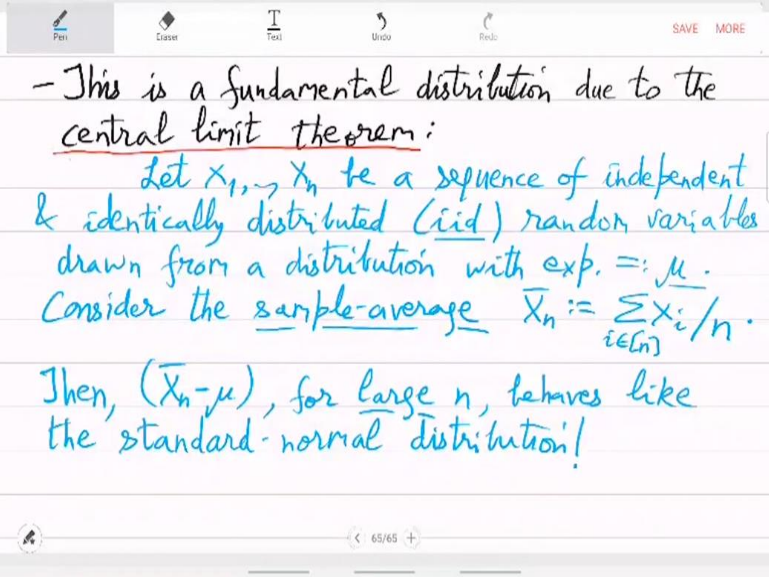
So, at small x equal to 0, this is 1 over square root 2 pi. And as x goes on the positive half, you can clearly see that the density falls. And when x is infinity, then it becomes 0. So, it falls something like this. And there is complete symmetry; so, the same thing on this side. So, this is what is called a bell curve. You must have seen this before. It is centred at x = 0, and the probability is very high around the centre; but then, towards the ends which are called tails, it is falling.

So, probability of being in the tail area, it is very low; probability of being around the centre zone is very high. That is a bell curve. And the expectation is the value x times x being in that vicinity. So, that is 1 over square root 2 pi e raised to minus x square by 2 dx, which is a calculation you can do. So, you will get 1 over square root 2 pi minus of e raise to minus x square by 2.

And the beauty is that, since x square is the same for in minus infinity and infinity, the difference is 0, which says, what you see in the picture. It says that random variable is supposed to be around in x = 0; it is centred there. That is where you expect practical experiments to be. And in fact, it is exactly the expectation. So, this is the famous bell curved Gaussian distribution, because Gauss defined this, in the eighteen hundreds.

This is a very fundamental object. I will talk more about this. It is also called standard normal distribution. So, this is the standard normal distribution centred at 0, and this factor being 1 over square root 2 pi; that is the maximum value it takes. That is kind of the probability density, in the band x = 0, in that vicinity is 1 over square root 2 pi. And then it is falling towards 0 as you move towards the tails. So, why is this important?

(Refer Slide Time: 11:49)



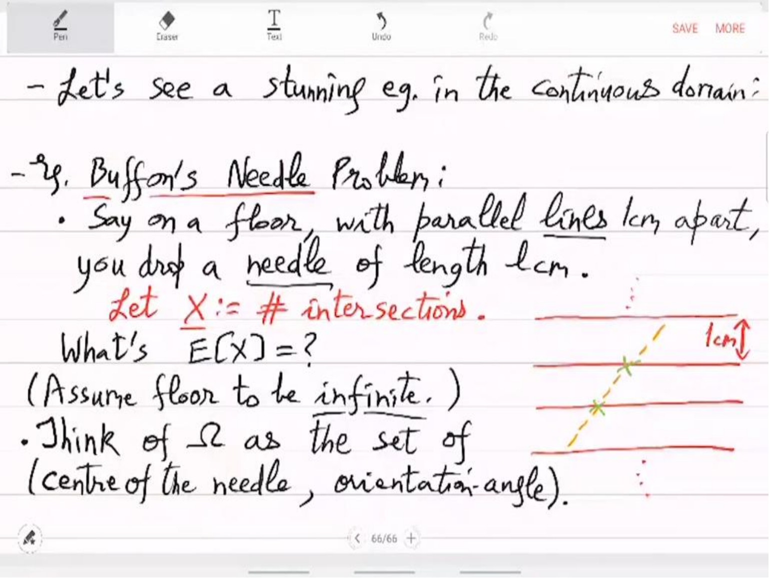
So, this is a fundamental distribution due to the central limit theorem. So, what is that? Central limit theorem essentially says that, no matter what random variable you start with, if you repeat that experiment too much, then ultimately you will get the Gaussian distribution. So, let me be a bit more specific. So, let X 1 to X n be a sequence of independent and identically distributed, that what is called IID, independent and identically distributed.

So, you can think of this Bernoulli trial being repeated, the binomial distribution; you can think of that for example; but this actually holds for any distribution, any random variables drawn from a distribution with expectation equal to mu. So, this random variable which you are repeating n times, each has expectation mu. Now, consider the sample average, X n bar sigma X i divided by n. So, that is the sample average.

So, you repeated this same experiment n times and look at the sample average. So, the question is, how far is this from the expectation? So, the central limit theorem says that, then, X n bar minus mu for large n behaves like the standard normal distribution. So, essentially, you can start with any distribution you want, and by averaging, you can get to the standard normal distribution. So, that makes it really fundamental and surprising.

This is not something you would have expected. Maybe we will do, we will sketch the proof later in the course. So, for now, this is just to emphasise the importance of the bell curve, this normal distribution.

(Refer Slide Time: 16:20)



Let us see now an example in the continuous domain, which is the example of Buffon's needle problem. So, obviously, there is a needle and you drop it on a floor. So, say, on a floor with parallel lines 1 centimetre apart, you drop a needle of length l centimetres. So, let me draw a picture for you. So, these are the lines, say they are infinitely many; so, infinite floor space. And this is all 1 centimetre. And you have a needle which is randomly dropped.

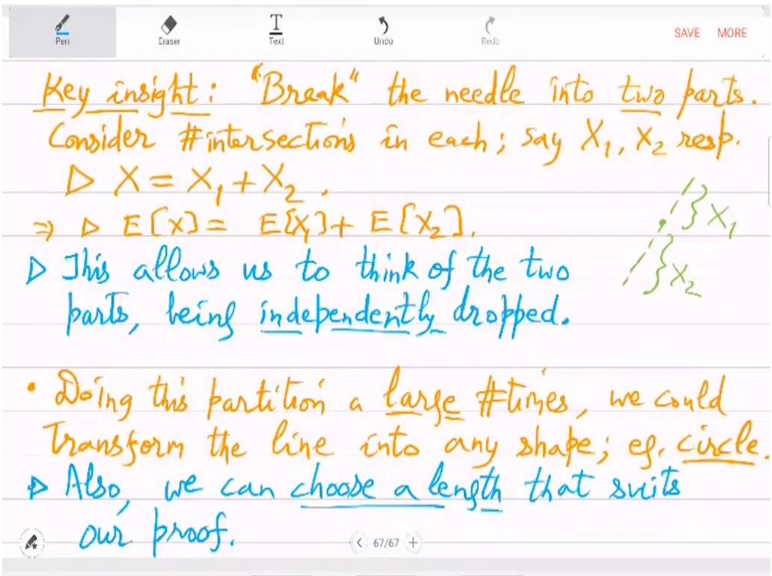
So, say it falls like this. So, the intersection points are these. So, this needle is intersecting with 2 lines and that is the random variable of interest. So, let X be the number of intersections. So, when you randomly drop a needle on this infinite floor space, how many intersections do you expect to happen? It is an interesting question; hard to imagine what the answer will be, because there are all the possibilities.

The needle may fall between 2 lines, and then there is 0 intersection; or it may fall perpendicular to the lines, when you get the maximum intersection; or anywhere in the middle. So, what is the answer? And you have to give an answer as a function of l. So, what is the expected value of this random variable? Assume the floor to be infinite. This is not very important; this is just for the theoretical calculation.

But in practice, as long as you are looking at a very large floor space, you will get this expectation which will be also true in an experiment. Think of infinite just as very large. So, what is the sample space that we are looking at? Sample space is, how can the needle fall, the ways. So, basically, the centre of the needle, what point is that and what is the angle? So, think of omega as the set of centre point of the needle comma orientation.

So, these 2 things when given, you can draw the needle, you know how has the needle fallen. So, already the centre points choices are infinitely many. And then obviously, same thing is true for the orientation. It goes from 0 degrees to 180 degrees. So, yes, the probability calculation is tricky, even the formalisation is tricky. So, instead of going into the very rigorous parts of it, let me just give you the key idea. What is the key idea for this?

(Refer Slide Time: 21:54)



What is the key insight to use? Think of breaking the needle into parts. So, break the needle, let us say, just 2 parts, and consider the number of intersections in each. So, let us say, in the first part it is X 1 and in the second it is X 2 respectively. So, here is your line; you break it up into 2 parts. So, intersections in the first part; intersections in the second part. Then the number of intersections is the sum of this.

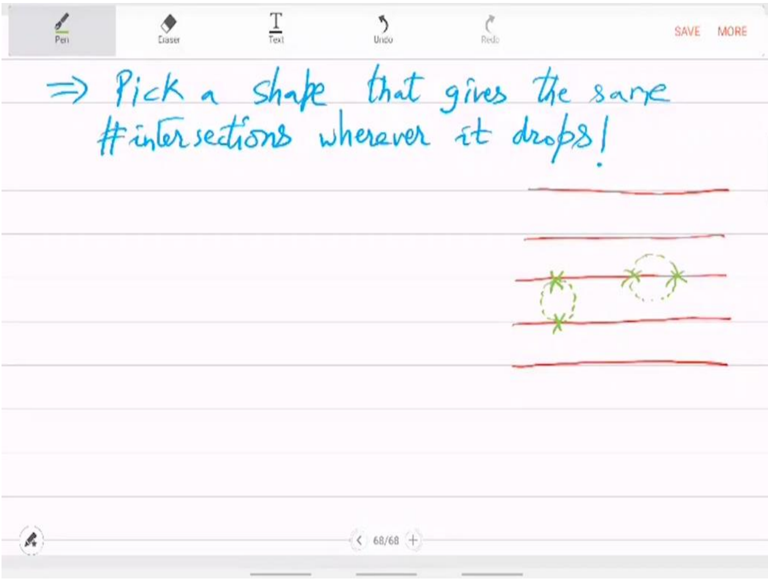
So, you are interested in this random variable which is a sum of random variables, which also means that the expectation is a sum of those. This is linearity of expectation, but what this formula has done, the insight you get out of this is that, now you can think of X 1 and X 2 completely independently. You do not have to think that they are joint. Now you have truly broken the needle, and then you can also just independently throw them on the floor.

So, this allows us to think of the 2 parts being independently dropped. So, independent is the keyword. So, you can independently drop these 2 parts. You do not have to think of them as being joint. And independently calculate the expectation, and then just sum it up. And while you are at it, you can keep breaking it into infinitesimal parts, and you can also make a circle out of it.

So, doing this partition large number of times, we could transform the line into other shapes. So, why not a circle? So, instead of this line, we can actually as well think of a circle, because, from the line, we can take these infinitesimal parts, join them to make a circle and then throw the circle on the floor and compute the probability. So, other thing is, also we can choose a length that suits us, because, for example, you can think of the line, instead you can think of a line which is 1 centimetres.

And from that, whatever you get, you can multiply it by l to get a result for l centimetres. Similarly, you can work with a circle of some circumference. And the answer you get, you can then multiply to get circle of or line of length l centimetres.

(Refer Slide Time: 26:56)



So, basically, pick a shape that gives the same number of intersections wherever it drops. So, shape, for example, if you take a circle which just fits between 2 lines, then you can see that wherever it falls, the number of intersections does not change. So, if the circle falls here or if the circle falls here, the intersections are always 2. So, using this, you can get the probability. So, we will finish this next time.