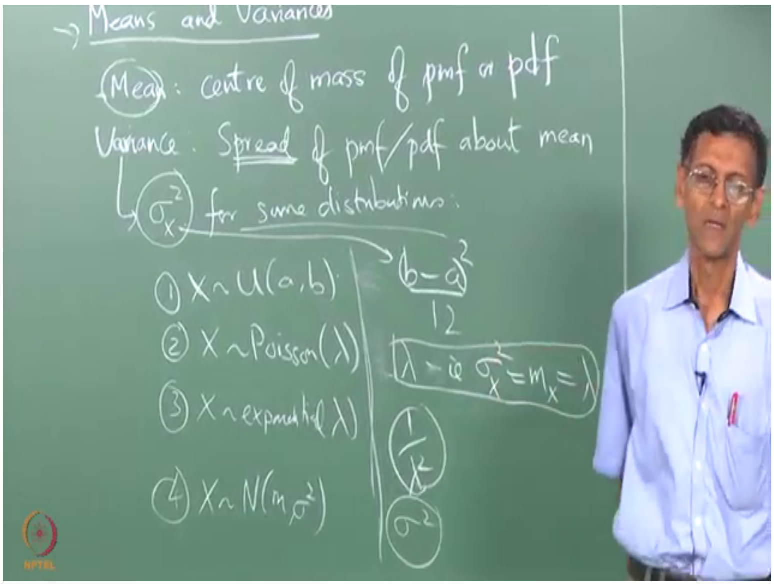
Probability Foundations for Electrical Engineers
Prof. Aravind R
Department of Electrical Engineering
Indian Institute of Technology, Madras

Lecture – 67
Variance

(Refer Slide Time: 00:20)



Last week we started we defined the means and variances of probability distributions both discrete and continuous right that turns out that the meaning is the same right whether it is a discrete distribution or a continuous distribution.

So, what is the mean? In words the mean as a center of mass and the variance is what? Spread these are the two things that you have to always remember about mean variance. And quite these are very likely questions on a on an interview. So, if you look at some I think what we did the other day a day was to give some examples, but let me take this variance a little more detail and couple of minutes we spent and spend a couple of minutes on it this sigma X squared right the notation and this is for the random variable X of course, and then here we have m X.

So the variance for various for some distributions right. So, supposing for somebody let me say let me write it like say first let us say they the simplest ones like the uniform distribution for example, a b that is anyone know the expression for the variance since nobody is answering immediately let me say b minus a divided whole square divided by

12. So, what does this tell you? What is it what information you get from this expression? The variance depends only on;.

Student: (Refer Time: 02:45)

Only on the.

Student: (Refer Time: 02:47).

On the width of the distribution does not depend on absolute values of a and b.

So, you can shift the distribution up and down; the variance will not change there is going to be a more general result that we will see right. The mean of course, will always remain in the midpoints; obviously, if you shift the distribution the mean will change, but. So, for the mean actually the value that is the origin of the distribution that X equal to 0 value is a is a great significance that that depends that determines the influences the mean, but the variance is not dependent on where you put the origin of the x axis that is what right to want to say. Then I am looking mostly the continuous well look at it mix up a continuous and the discrete.

The idea is that you should erase these distinctions and look at holistic view of mean and variation. Supposing I say X is Poisson ; so, for the Poisson case and exponential geometric and so on, I have its they are all single parameter distributions they are uniform it turns out this depends on two parameters a, b right. Now for the Poisson distribution; what did we say about the mean? We said the mean was lambda right; m X was lambda we already talked about that. Variance is what? Is also lambda; so, listing values of sigma X square here not this not equals here. So, this is the values of sigma X square in this column right; the variance in this column right we should have to be very careful not to confuse what gets recorded.

So, the variance is also lambda; so, sigma X squared equals m X equals lambda. This may look a little funny, but this unfortunately the way or fortunately it is the way it comes. So, if you increase lambda you end up increasing both the mean and variance; you push the distribution to the right and you are also in end up spreading it right.

And remember right Poisson distribution works for an enormously large range of values of lambda in practice; you can get absolutely small values you can also get very large values right ok. So, then if I look at the exponential for the exponential we said the mean was 1 by lambda; the variance is 1 by lambda squared.

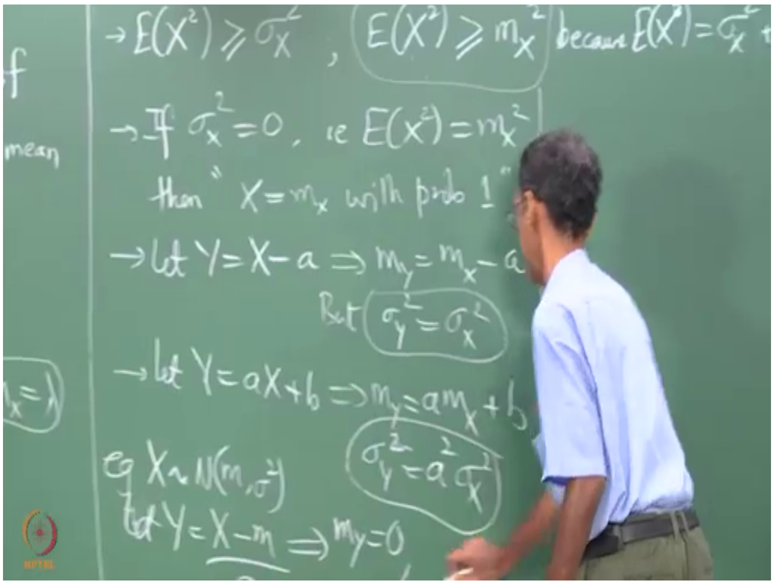
So, note that now you get an lambda inverse of square. So, as you make the lambda smaller and smaller you get larger and larger variance up by the square right. So, ; so, if you make so; that means, so, what happens is of course, the distribution starts to spread quite rapidly as lambda becomes smaller and smaller ok; then the Gaussian which beyond I am going to stop with Gaussian.

So, these two are single parameter distributions right well again I want to point out geometric also I can write down here, but I am not going to write geometric and exponential are very similar in behavior in many ways. The geometric variance is 1 minus p by p square again you get that 1 by p squared 1 by lambda squared kind of behavior; lambda and p are kind of are very similar parameters in the true distributions I for the geometric and exponential; p becomes small variance of the geometric will become large; mean also becomes large, lambda becomes small same thing for the exponential ok.

So, let me as a last example and just let me say X is a m sigma squared of course, we have sigma squared is the variance here. So, here again right there you have the two parameters; you have a perfect decoupling between the mean and variance right. There parameter m affects only that is there only the mean and the parameter sigma squared is only the variance. So, you can set the mean and variance independently in a Gaussian distribution without a choose one without affecting the other ok. So, let me ; so, with this thing I hope my intuition is changed.

So, we will move on just instead of taking up more computational examples which is not my main point of what I am trying to do here. So, anyway given it I suppose by now all the; you have written down all the formulas that apply right. So, I do not I am not going to write them out here again instead let us move on to some general properties right.

(Refer Slide Time: 07:57)



In the sense of this E X squared is always greater than or equal to sigma X squared E X squared is always bigger than equal to m X square. Why? Because what is E X square? Exactly equal to sigma X square plus m X squared and this is a relation between non negative positive quantities in general right.

So the difference here is sigma X square and the difference here is the mean square. So, minus E X squared equal to sigma X squared for any 0 mean random variable E X; variance is exactly the same as this mean squared value. So, I can have m X equal to 0 which means this would be equal to that right.

So, so remember right this particular thing may be is a clearly the you know in a more in it may be a little even more counter intuitive of the two, but it is true right. And it even for a simple, but now redistribution with parameter p you can appreciate why e of X square has to be bigger than m X square; what is there it turns out both E of X and E of X squared are both equal to p, but what is E of X the whole square; rows m X square? It is p squared. So, what happens to a number between 0 and (Refer Time: 09:43) you square it; it becomes smaller than value right.

So, you can see why that a of X squared is always bigger than or at least no smaller than m X square. Supposing they are equal; in other words if E of X squared, you get this with equality what does this mean, what does this mean? When is this right or supposing you take a limiting case of sigma X squared becoming very very small.

Student: (Refer Time: 10:27).

All the probability is concentrated at one point. In fact, the randomness keeps reducing there is no spread in the PDF in the limit as sigma X squared goes to 0.

So, so you can say that X this you this is a probabilistic statement then X will be equal to m X with probability 1; that is a probability that X is equal to m X is unity. And this is also write a statement in the limit; so, as you consider a sequence of PDF's that is something which we may talk about towards the end of this course right or pmf. Sequence of distributions index by some parameter says say the variance keeps decreasing as you go further and further into the sequence. Then there is no you have the inescapable conclusion is that the random variable is taking values closer and closer to the mean and with very high probability.

So, this is what in fact, is at the at the bottom of things like the law of large numbers; that is hopefully we will talk a little bit about it before we sign off this course ok. Then what else to be, but do I want to say right now in terms of summarizing some properties of mean and variance ; let us Y be equal to X minus a ; supposing I just construct a do a very simple transformation like I just do X minus a.

So, it turns out if I do this; I do not change the variance of X, in other words this implies that mean of course, will shift, but the variance will remain unchanged. This you can please check this if you are not right, it is very easy to plug in the formula; go here and plug it in right m sigma Y square and in terms of sigma X squared you will see they are exactly identical; exactly the same.

So, this is this kind of manipulation is very useful in the continuous case right, where X takes a continuous of values is also applicable in the discrete case, but in the discrete case usually you have integer what you call domains right for the omega X that is X taking values in it is, but a can be some non integer number as well.

So, if you remember right; for example, if you take the binomial and you subtract some you subtract let us say whatever some three point number 3.5 like that from each value of the binomial that is what we are doing, when you do X minus a. You do not get an integer valued random variable at the output of that. So, Y will not in general be integer valued even if X is integer value when a is some arbitrary number, but the result is still perfectly valid is no problem with the result ok.

Now, let us look at a more useful version of this a slightly expanded version. Supposing I say Y is a X plus b now remember this a and a are different, but it is better to write it as a X plus b rather than b X plus a right they were more interested I think we did the same thing earlier.

So, a let us write look at a X plus b how does this affect this operation how does it affect the mean and variance? Mean is very easy all you have to do is apply e on both sides, use lotus and what do you what does it end up with?

Student: (Refer Time: 14:18).

a I am assuming you know m X N sigma X squared; so, in terms of that.

Student: a square

Not a square

Student: amx plus b

That is right sigma Y squared.

Student: a square

Is a square.

Student: Sigma X square

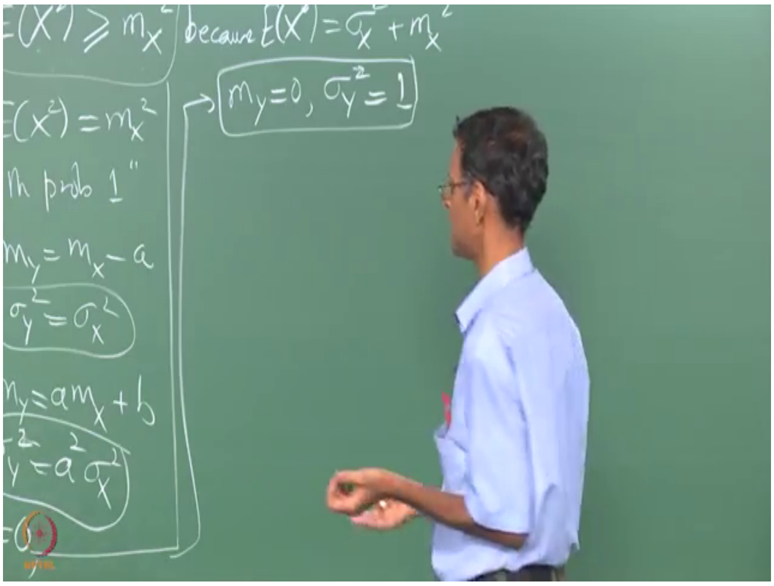
Sigma X square this is not depend on b again because I have to go a little fast I am not putting up putting all these derivations, but it is actually quite simple; all you have to do is first get your Y squared by just simply squaring both sides of this. And if you square both sides you get a squared X squared plus b to a b X plus then you take expected value on both sides and then subtract off m Y squared, which is square this then you will find that b squared will cancel and you will be left with exactly this.

So, this combination of scaling and shifting only changes the variance and does not change they mean sorry; it changes only the mean and leaves a variance intact. So, why is this useful? It means that for example, in the Gaussian distribution I can go very easily back and forth between arbitrary mean and variance and 0 1 mean variance. Supposing X is Gaussian with mean m sigma squared variance. Now I already pointed out this kind of thing which is a generalization of this is the shape preserving transformation.

We mentioned this very clearly when I talked about transformations right. So, the shape of the PDF is the same. So, in particular the Gaussian will remain Gaussian; Y will remain Gaussian. So, what is, but supposing I choose a supposing I choose then I let; let Y be equal to X minus m divided by sigma; not sigma square if I do if I define Y as X minus m by sigma.

So, X X minus m by sigma is a special case of this isn't it? So, what happens? What is it what are the mean and variance of Y? Mean is 0, variance is one remember you have to divide by the square root the variance here we it is not calling it by the other name, but this implies that m Y 0.

(Refer Slide Time: 17:27)



Maybe I should go back here and write and; obviously, this result works in the opposite direction as well. So, starting from 0 1; you can always inject some mean and variance by going the by doing this in other words.