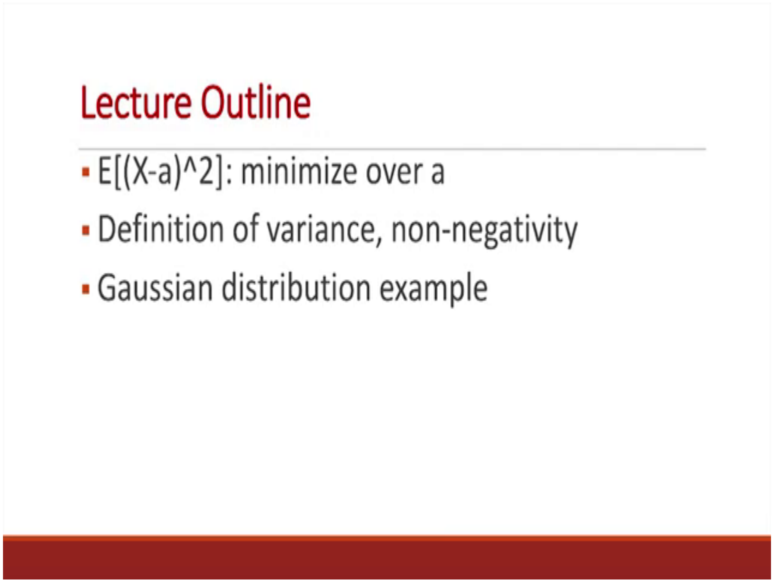
Probability Foundations for Electrical Engineers
Prof. Aravind R
Department of Electrical Engineering
Indian Institute of Technology, Madras

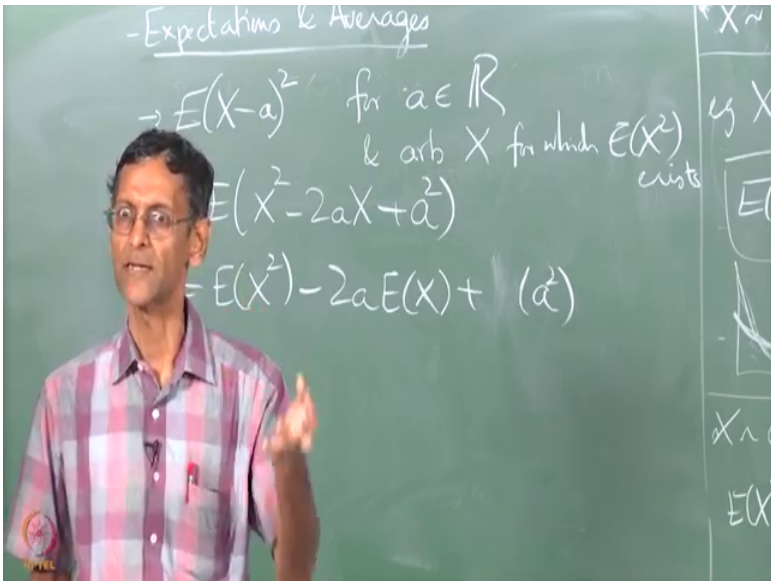
Lecture - 66
Variance

(Refer Slide Time: 00:14)



First of all.

(Refer Slide Time: 00:18)



What is the value of this E of X minus a whole square for some arbitrary a; some real a; if somebody asks you this note that if I already have a parenthesis I do I can skip the outer parenthesis right this means it is E of X minus a whole squared this is my g of x. Assume, that right it is finite it exists and. So, on no let us not right now scratch our head.

So, for some a and some and write any arbitrary and arbitrary X for whom right for which; let us say E X square exists this is the only condition I want. If E of X squared exists this almost surely will exist if a is finite right there is no issue there ok.

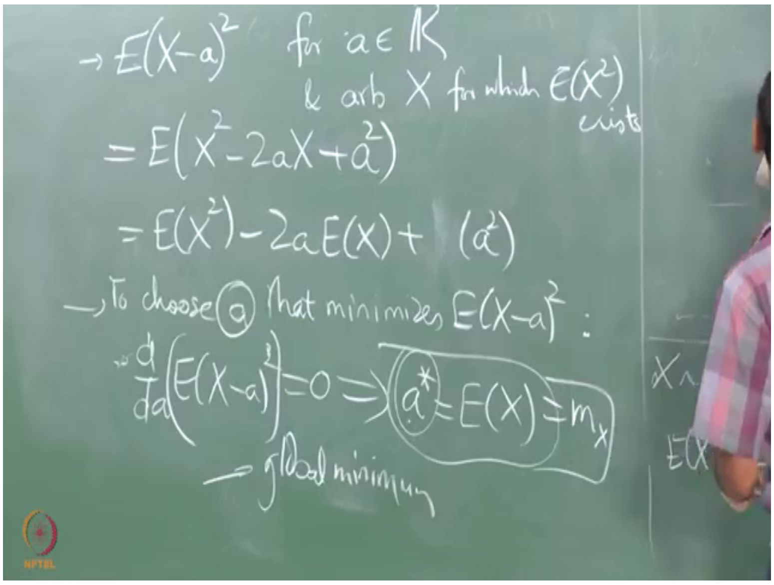
So, what is? So, if I want to simplify this why am I simplify is right why am I justified in expanding that function there? Because, X is after all always a real number it may be a random variable, but it is a real random variable right. So, it is always a real number. So, this is identical to E of Y X squared minus two a X plus a square you can use ordinary algebra to diff to simplify expressions involving random variables. And we are going to do that heavily from now on; because for every real value taken by X assigned it is valid. So, it is there is no right no way that this cannot hold.

Now, we use a linearity property of the E operator. So, what happens to this which is this nothing, but a square? So, I can drop this E a square is some constant as far as the expectation is concerned. So, it is expected values a squared itself. So, as well as we saw yesterday right so, this is a special case of E of a X plus b which is a E X plus b 1 of the what yesterday today if I did not say it you can write it down also E of X plus b is a EX plus b same thing applied here. Remember v of X squared exists E of X will also exist that is how another property which we have not formally proved, but if a higher power exists the lower power has to exist, because in general X squared is a is a sitting on top of X for large values right

You have what is called bounded convergence and all that right. If this is convert, if this if any integral or sum involving this converges any integral or sum involving this also has to converge; this is dominated convergence theorem that is what it is right. If you want to know more of the math's you can go look at that right what we call dominated convergence right. So, I have this.

So, what do I want to do with this? I want to pick the a which minimizes this quantity. And I assume I can make incremental changes in a that is a is not a is completely under my control I i can I have the whole field of reals to choose the value of a, but I want to minimize I want to find that a which minimizes this a is in a sense an estimate of X, which minimizes the this mean squared value special mean squared value it is not just E of X squared, but E of X minus a the whole square. So, we want to choose a.

(Refer Slide Time: 04:04)



What do we do? You differentiate with respect to a very simple. So, implies what?. So, we get a star the meaning there is a unique minimizer which we write as a star and not as a right do you or do you not get this does not depend on a no matter how fancy it is it does not depend on a at all. So, it will go away the only thing that depends on a is this part remember you are not right this is a smear number as far as a is concerned by changing a you do not change this. So, this is the significance of the mean basically this is equal to m X.

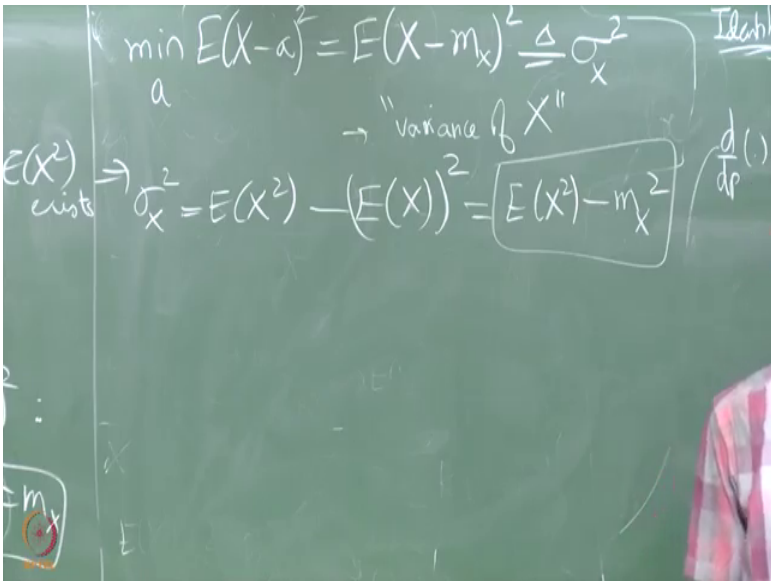
So, if somebody asks you an interview viva question; what is the significance of the mean? What will you say one among many other things it is that number which minimizes the; it is the best in the absence of any other information about X the mean value is the one that is plum in the middle in the sense of minimizing this quantity. It is right it is that value are from which the distance from the average distance to all other values taken by x is a minimum weighted by the probability of course.

So, this significance of the mean itself is missed, then what is that minimum value? This may be the minimizer, but if you substitute X equal to a ma s sorry a star equal to m X in here what do you get you get the what is called the variance the minimum of this.

The minimum over a of E of X minus a whole squared that is it is unique right it is a global minimum. So, this is a global minimum remember why is it a global minimum? Because what kind of a function is this with respect to a? It is a parabola every parabola have right; if he does not has a low I mean it is an upset it is a proper parabola in the sense that it is or what do you call it a convex concave.

I mean we do not use the word convex sorry concave it is a convex u parabola it is not the parabola which is the other way around right. So, it has a minimum not a maximum in other words right. So, this is this is a global minimum and this it is achieved at a star which is m and.

(Refer Slide Time: 07:20)



And the minimum over a of E of X minus a whole squared is basically E of X minus m X whole squared this is. So, important that it is written as sigma X squared called variance of X. It is by far more important than E of X squared itself; although E of X squared is needed to understand where it comes from and if I want to simplify that all I am doing is I am sub as of now I am just simply substituted a star or m X for a, but I can simplify whatever; what happens if we simplify this?

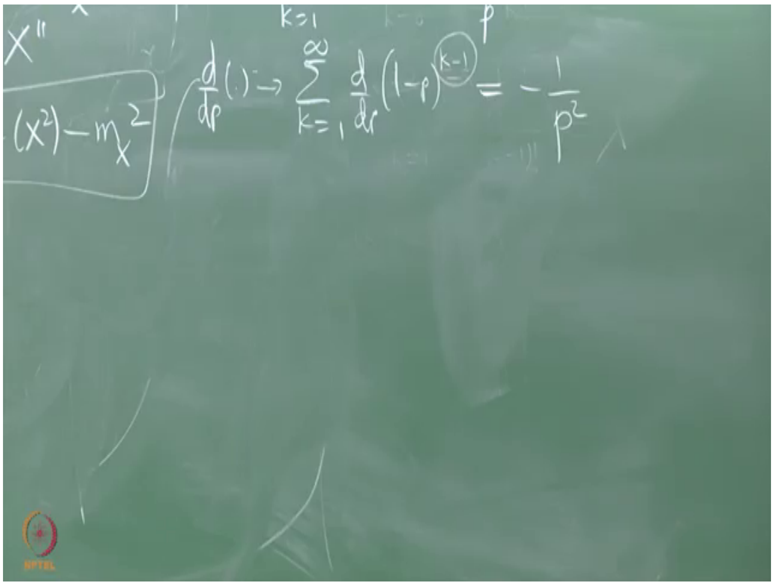
Basically E of X squared I am going to substitute in place of a what am I going to write I am going to write E of X itself what do I get if I substitute E of X here.

Student: (Refer Time: 08:33).

So, this is a nifty neat formula for the variance, in terms of just the mean squared value and the mean; you do not have to do a long calculation like this you can just simply remember this. And not only that the variance written with squared for a reason the variance can never be negative it always has to be pot non negative.

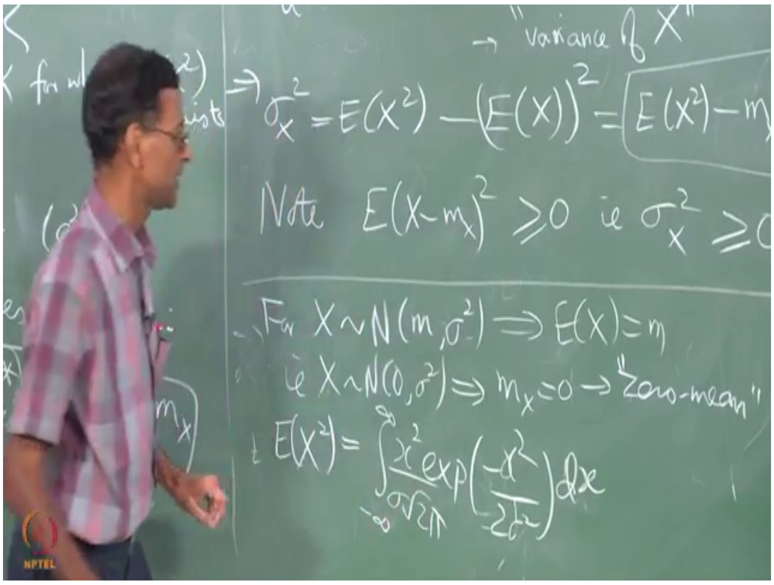
Why does it always have to be non negative? This quantity can is a g of x which is always non negative you know what we said yesterday the expected value of a non negative function is always non negative E of X squared; I pointed out yesterday has to be non negative, similarly the variance also has to be non negative.

(Refer Slide Time: 09:41)



Right.

(Refer Slide Time: 09:48)



And, why this squared because in statistics they also have a separate use for sigma X by itself this is positive square root of the variance; they call it the standard deviation we are not going to use that much here in this course it is enough if we stick to the variance, but in some equations you will you may see sigma X coming by itself like. For example, in the Gaussian case, we will do we are going to do that today and conclude right you see the sigma X by if you see it by itself you remember it is a square root of the variance you do not have to remember the other term associated with it right; just remember variance and variance is exactly this you do not know square root involved ok

That is why in the for the Gaussian case it turns out right. When the Gaussian case, as we said the other day a long back right the m parameter does not affect the spread. So, the it turns out that the variance a lot of things about the variance, but lend me. So, I claim that it is enough for me to look at the to calculate the for the for the Gaussian case is enough to look at 0 sigma squared, because it turns out that that we will ok. We will justify this in a little bit, but let me let me just at least do the 0 sigma squared and then later on I can I can justify what I just now said.

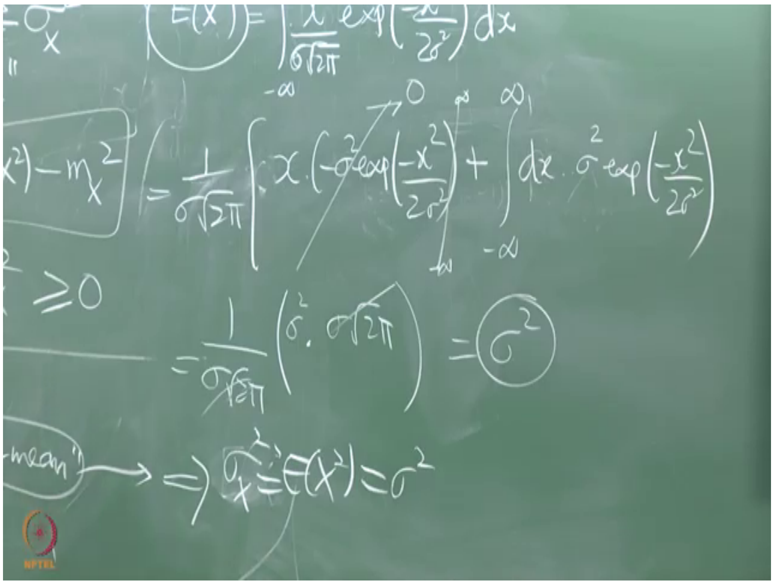
So, now note that I am just writing sigma squared here I am not putting sigma X square right as of; now I have I have no proof that right this is. In fact, the variance, but that is; what is going to come out to be right? So, we know that for this E of X is certainly m this we know that is m X is equal to m this parameter m, because of point of symmetry right.

And as I said there is no problem with the existence of area I mean the mean or the mean squared. Now for the variance I am going to take m equal to 0, because it turns out that. So, I E if I take X is 0 sigma squared this is implies that m X is 0 this is this is. In fact, this is called the example of 0 mean.

So, you can have a great many situations where the mean is 0. In this particular case you have a symmetric PDF about which it says PDF in the 0 sigma squared case is symmetric about the origin and therefore, the mean is 0 mean exists it has to be 0. In the uniform case that we talked about earlier that a plus b by 2; I can again have mean 0, if I take what if I take b equal to minus a? That once again it becomes symmetric about 0 instead of something else ok. So, I am just saying that for this case it is easier; now to deal with the; vary the E of X squared for this.

So, it turns out that if I take 0 sigma squared E of X squared is sum is an integral like this minus infinity to infinity X squared E power minus X squared by two sigma squared dx and of course, there is a square root of sigma 2 root by root of pi here the difference this is x square term here that in the. So, once again you integrate by parts integrate by parts by taking just x into x something. So, if you integrate this by parts maybe I have should do it here just one step. So, I am going to look at the integral for E of X squared this is not right this is not the variance or anything I am just like just simply finding E of X squared that is all ok.

(Refer Slide Time: 13:54)



So, I have E of X squared equal to integral minus infinity to infinity X squared X by sigma 2 root to pi x of minus x squared by 2 sigma squared. So, this I will write as x into x ok, and I will write this as x times the integral of the remaining part is it turns out of course, I will also do one thing let me pull out this 1 by sigma 2 to pi outside ok.

So, it is x times minus sigma squared this is the anti derivative of what this is the anti derivative of x into this there is a minus sign here, because there is a minus sign here and it becomes plus integral minus infinity of the derivative of this which is dx right u dv mi uv minus v d u, this is this is my v dx into sigma squared into x buff minus x square by 2 sigma squared right.

And this is evaluated between minus infinity and infinity. Actually, we are sweeping you know some details under the carpet here, because these are all infinite integrals and. So, on, but what do you think happens to this first term? It will it will be zero at both ends, because as I said earlier this E power minus X squared is very strong and x power k has no effect on if you take limit as x tend to infinity of x power k; E power minus X squared it is 0. So, and what is this is basically back to a Gaussian integral is it not. If you remove this a sigma squared pull it outside what is the rest of it?

This whole thing is 0 so this is 1 by sigma root 2 pi into that thing is sigma squared times it turns out it is sigma squared into there, there it will be root two pi into sigma or something be exactly this or rather this thing taken here the integral is one without the sigma square. So, the answer is sigma square. So, what have we got here we have got the fact that E of X squared is sigma squared and the mean is 0 right. So, what is the variance sigma square? So, it turns out.

For this case that E of X I will write it this way sigma X squared equals E of X squared equals sigma square. So, the parameter sigma square directly becomes the variance, because the mean is 0. So, why do I claim that for any m sigma squared sigma squared is a variance, because basically it turns out m the only thing that m does is shifted right and left and it turns out that the variance is directly the spread parameter of the PDF.

So, you can actually do I am not going to take it any further, because it is very boring to take it any further I mean this particular example, but for the general Gaussian case also right it turns out m is the mean and sigma squared is the is the variance we have proved it for the 0 sigma squared case m sigma squared case also exactly the same thing happens.

And, I have already talked about the effect of sigma squared on the shape of the PDF right that we talked about long back. Now you can understand where that shape is coming variance, basically tells you the spread of the PDF in general not just for Gaussian if you look at the exponential distribution for example, the variance will be one by lambda squared 2 it is 2 by lambda squared minus 1 by lambda squared from the mean.

So, the variance would be one by lambda squared it is half of E of X squared again lambda controls the spread the smaller the lambda the more the spread for the for the geometry case the variance will be 2 by p squared or wait a minute maybe not. So, no so, it is 1 minus p by p square sorry; what am I saying anyway ok. Let us not I am going end the lecture here, but just note that all these results are perfectly I mean are very standard big tables of this I am sure you can find all kinds of online resources to help you out.

So, the aim of this these lectures should not be to re replicate all the stuff that we have seen on the net right we can get elsewhere. So, instead we will look only at the conceptual meaning of what we are trying to calculate especially this kind of stuff