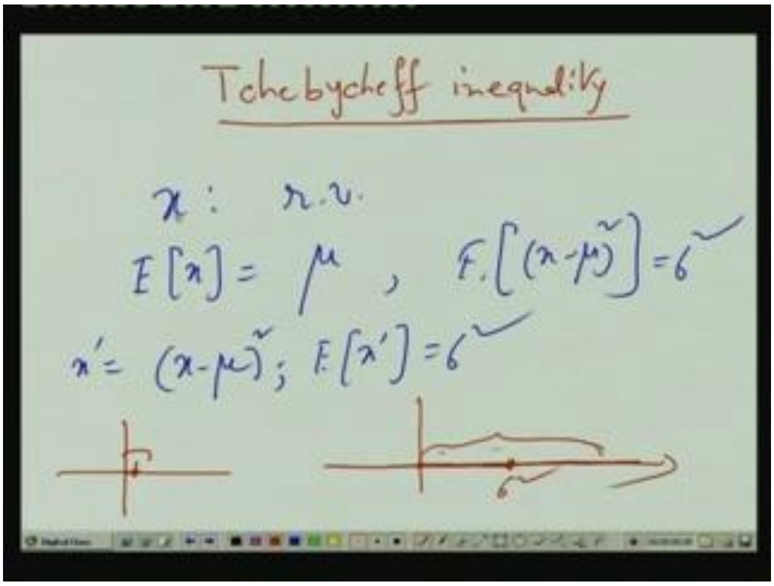
Probability & Random Variables
Prof. M. Chakraborty
Department of Electronics and Electrical Communication Engineering
Indian Institute of Technology, Kharagpur

Lecture - 27
Thebycheff Inequality and Estimation of an Unknown Parameter

So far I mean we have been discussing random sequences. A very important concept in that context is that of stochastic convergence, that is I mean classically we have seen in mathematics you know that if you are given a sequence x of n, then we obviously immediately talk about its convergence whether it converges to something or not. Similarly, if you are given random sequence, that is a sequence of random variables, obviously same question can come again that is whether it converges to some point or not.

So, what is the notation of convergence here, because it is not just a sequence of numbers, it is a sequence of random variables. We state various values depending on your experiment, right. So, what will be the notation of convergence here? This is very important. To get into this in fact, we will not be discussing that. Today just to develop the motivation for this how these notions are developed, just took it motivation for this topic. We will start at a topic an issue which maybe I should have considered earlier, but I mean I did not kept it for this lecture.

(Refer Slide Time: 02:20)



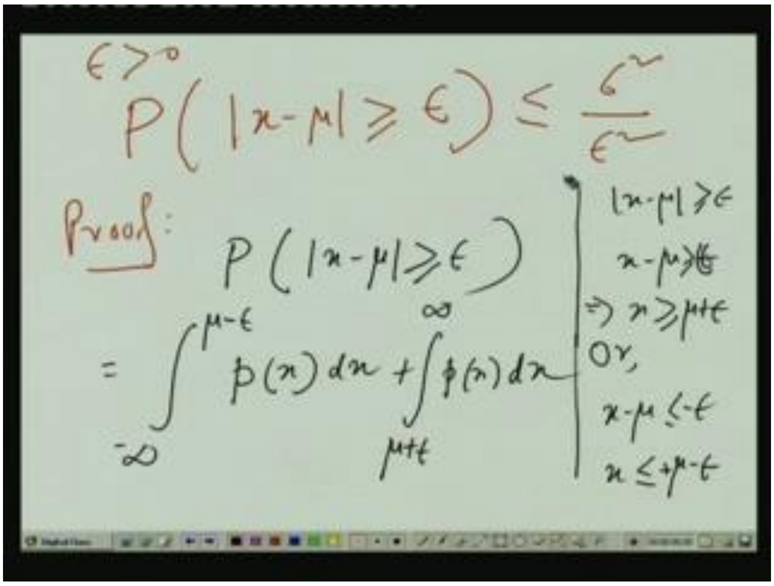
Now, suppose you are given a random variable x. X is a random variable expected value of say x is mu and expected value of x minus mu whole square which is the variance, it is a sigma square. Now, what does this mean? You first form a random variable x minus mu and square it up. You can call it x prime and then, e of x prime is sigma square, right. Now, you can see it means that on a line suppose this is origin. Now, x prime, its minimum value can be 0. X can be to the right of mu, to the left of mu, that is x minus mu can be positive or negative, but the square of x minus mu which is x prime is always positive or 0. It can never be negative. So, its minimum value is 0. That is why x equal to mu and as x deviates from mu either to the right or to the left x prime. The value of x prime increases.

So, it goes out like this as x takes further and further values, you know I mean goes further and further. Diaphragm mean mu its value increases. So, expected value of x prime could be somewhere here, and that is if I am able to take the square and if that is sigma square. Now, what does that mean that if I am observing x and then, this difference x minus mu square, it will take values around sigma square, right. So, in some experiments, x prime can value to the left of sigma square and in some experiment to the right of sigma square. So, its concentration is around sigma square to the left hand side. It can go only up to 0. That means, to the right hand side also, it will not go up to infinity because then the average cannot be sigma square finite number.

So, on either side it can go over a range like this. Then, my claim is I mean like this. So, then my statement is if sigma square or rather sigma both are positive is a small quantity. That means, if suppose a situation like this sigma on sigma square is only this much.
Then, I can roughly have a range like this around sigma square right for x prime, which means x minus mu this variable x prime. In fact, x prime sigma is x minus mu square will vary over a small range around sigma square.

So, physically this means that if I am now coming that we have random variable x and I am measuring x in the experiment. After experiment I am getting values close to mu because a deviation x minus mu, it will be over a small range. It can be positive or negative, but its range cannot be much because square of x minus mu which is x prime means it is much less which is sigma square. That is much less which means x minus mu that is the deviation of x around mu will be within a small range. That means if sigma square is really small, then if we measure x, we can say that I mean measurements for x, they are all coming across mu close to mu. So, loosely you can take even the particular measured value to an estimate of in this context an inequality comes to be a failure. This is Thebycheff inequality.

(Refer Slide Time: 07:02)

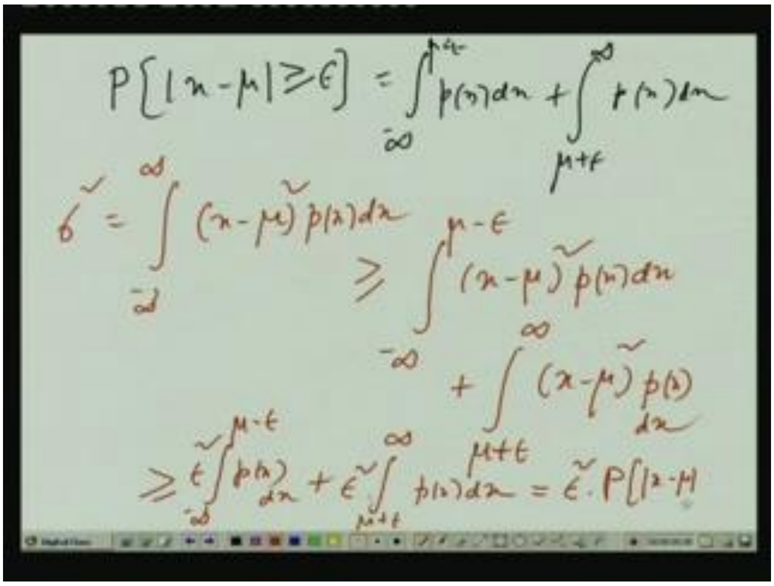


The probability of this deviation, its mod that is either to the right or to the left rather deviation across mu is positive or negative, but the mod value there is a magnitude of the deviation if there is greater than equal to some positive constant epsilon. Epsilon is greater than 0. So, probability that this deviation has magnitude greater than equal to epsilon is always less than equal to sigma square by your epsilon square. That means, if sigma is much less compared to epsilon, that is the variance of x is around x is less, and the epsilon there is the range of deviation that we are choosing is comparatively larger. Then the probability that the magnitude of deviation will be higher than epsilon that is x will indeed go beyond mu plus epsilon or mu minus epsilon will be very small because sigma square by epsilon square is small. I will again come to this point, but first let us prove it.

Now, what is this? This will be equal to 2 integral x minus mu greater than equal to epsilon means that is x minus mu mod greater than equal to epsilon means either x minus mu greater than equal to epsilon. That is either x minus mu greater than equal to epsilon. Epsilon is positive. So, I am considering positive deviation. There is deviation to the right of mean mu. So, x minus mu in that case is greater than equal to epsilon meaning either x is greater than equal to mu plus epsilon or if the deviation is negative, that is x is taking value to the i mean is very less than mu. So, deviation is to the left of mu. In that case, this is less than equal to minus epsilon. Either the deviation is greater than plus epsilon, greater than equal to or less than equal to minus epsilon.

In this case, you will get x less than equal to minus mu minus epsilon. So, what is the probability that x takes values either greater than equal to mu plus epsilon or value is less than equal to minus mu minus epsilon. Sorry, plus mu. This should be plus mu minus epsilon, either greater than equal to mu epsilon or less than equal to mu minus epsilon.
So, there should be two integrals that is one from minus infinity to mu minus epsilon. If you see Papoulis book, there is a mistake here. He puts a minus symbol with mu. That is wrong. It will be mu minus epsilon Px which is the probability density of x dx, and there is another integral mu plus epsilon up to infinity Px dx.

(Refer Slide Time: 11:53)



Now, we keep this result. This we keep somewhere that is I repeat here just for storing purpose, that is probability of x minus mu greater than equal to epsilon is nothing but minus infinity mu minus epsilon Px dx plus mu plus epsilon infinity Px dx. This result we keep separately. Now, we consider sigma square. What is sigma square? It is x minus mu whole square. It is expected value. So, this is multiplied by Px dx integrated. So, obviously since x minus mu square mu whole square, its positive P of x is non-negative number. I mean all are non-negative x minus mu whole square P of x. These are non-negative and dx is of course always positive.

This is this integral obviously is greater than equal to same integral, but over a shorter range, over a reduced range. It is from minus infinity you say mu minus epsilon x minus mu whole square Px dx plus earlier we are going from minus infinity. Now, we go from minus infinity only up to mu minus epsilon, and again jump over region and start again at mu plus epsilon and go to infinity. So, it is a reduced range. So, obviously the value since all the quantities being integrated are positive, the net value of the integral on the right hand side or the two integral summed added will be less than equal to sigma square.

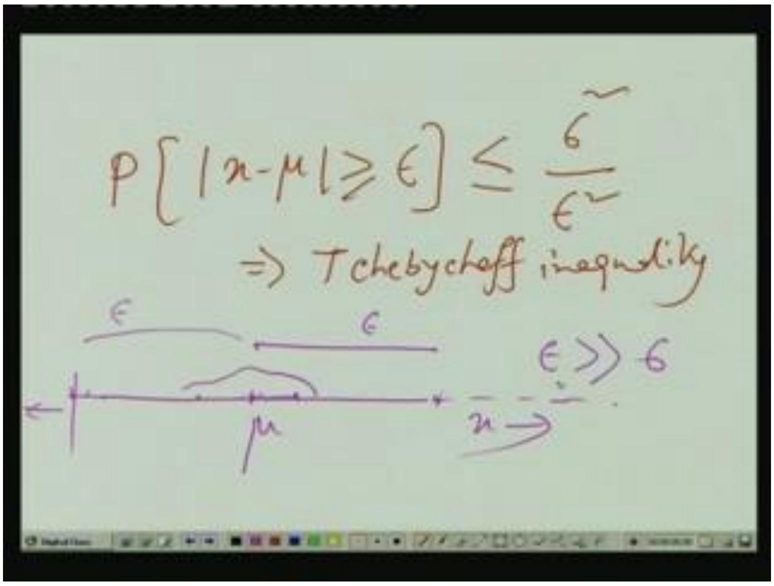
Now, in the first integral consider or may be in the second integral. To start with you consider x minus mu whole square this function as x takes value mu plus epsilon. I get only epsilon square. Mu mu cancels plus mu minus mu cancels. I get only epsilon square. Epsilon square Px dx. As x takes values still higher and higher, x minus mu whole square that also takes values larger and larger, that is higher and still higher than the original I mean starting value of epsilon square, right. I repeat again the starting value of x is the lower limit that is mu plus epsilon. If you put mu plus epsilon here plus mu and minus mu cancels, and you are left with epsilon square epsilon square, but as x increases further, so you put values here which is not just mu plus epsilon, but something more.

So, obviously may be delta. So, mu plus epsilon plus delta if we put that see you get epsilon plus delta whole square which is greater than epsilon square. That means this integral is less than equal to this second integral. This is less than and equal to mu plus epsilon up to infinity Px dx, because in this function x minus mu whole square takes the higher and higher value as x goes from mu plus epsilon up to infinity. I am holding this fixed at epsilon square. I am not allowing this quantity to go beyond epsilon square here.
So, obviously net value is that we get by this calculation will be less than what I get by this integral here. In this integral, this function x minus mu square is increasing function, but here I kept it fixed and epsilon square which is the minimum value and integrate the rest.

So, obviously net value net thing that we will get is less than what this integral would give. Same way you consider the first integral when takes the upper limit mu minus epsilon square epsilon mu mu cancels, minus epsilon whole square, which is epsilon square, but as x takes values less than that may be mu minus epsilon minus delta. If you put that here, you get minus epsilon minus delta whole square which is higher than the origin starting value of epsilon square, right. So, as x takes lesser and lesser values and goes up to the minus infinity, this function x minus mu square goes on increasing.

So, if I consider only epsilon square for it and integrate the rest, then obviously I will get a lesser value. That means this integral is less than epsilon square. So, here you take epsilon square common. So, what you are left with is simply what we wrote in the top. This is nothing but epsilon square into probability of x minus mu mod greater than equal to epsilon, which we wrote at the top. So, obviously probability of mod of x minus mu greater than equal to epsilon is nothing but less than equal to sigma square by epsilon square. This is the Thebycheff inequality.

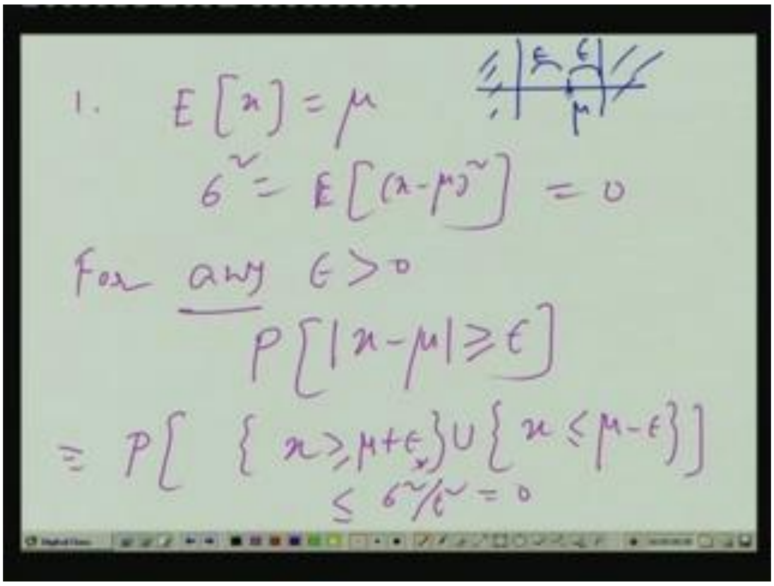
(Refer Slide Time: 18:22)



You remember this is what we wanted to prove. So, what is the significance of this? Suppose x is varying across mu, but sigma that is if you take really x minus mu and square it up, and take the average. That average value is sigma square. That means if sigma square is much less, if this corresponds to sigma square rather because sigma when we take sigma square, it is x minus will be whole square, but here x is for x. So, around this, if sigma square is much less, then we get values of x in most of trails around mu in a small zone. It does not go far to the right for two values and if you now take an epsilon like this, this much which is much greater than this is much greater than sigma.

Obviously, if this inequality means that the probability that x minus mu, it is either there is the deviation of x. X minus mu is either greater than equal to epsilon, that is x takes values to the right of this. X is either greater than equal to mu plus epsilon, that is takes values on this side or to the left. This side that is x is less than equal to mu minus epsilon. Probability of that will be much less almost close to 0, because sigma square is given to be sigma is given to be much less than epsilon. So, sigma square will be epsilon square almost equal to 0. So, obviously this implies the probability that x takes values far away from mu on either side that is x takes value either to the right of mu plus epsilon or to the left of mu minus epsilon. That will be very less almost 0 if epsilon chosen is much greater than sigma. So, we have been discussing physically that comes through this inequality also. Few things we can see.

(Refer Slide Time: 21:31)

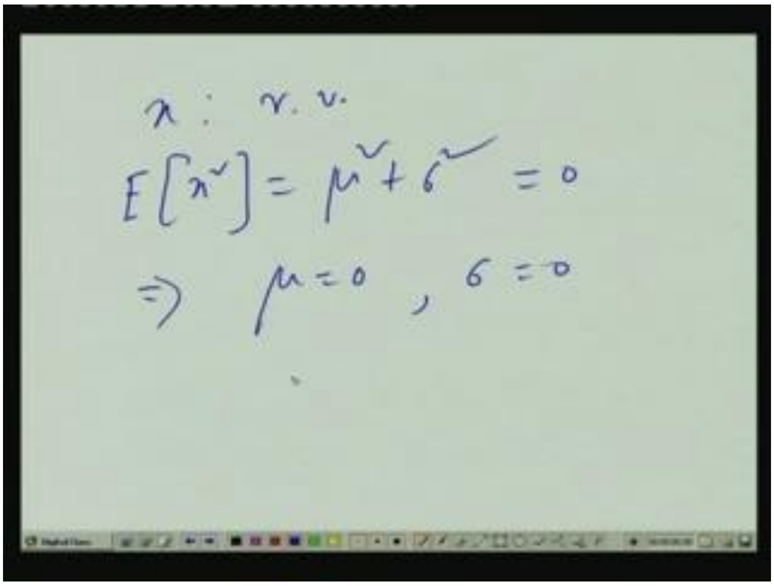


Suppose x is such random variable. It is given that E f x is mu, but sigma square which is E of suppose is given to be 0. There is x such random variable because mean is mu, but whose variants is given to be 0. Then, that means for any I underline for this for any epsilon greater than zero probability of I write it in x expanded x minus mu. Sorry just a minute which is also equivalent to probability of x taking values either greater than equal to that is why in x minus mu is positive. A deviation to the right of new is greater than equal to mu plus epsilon. This set union with x taking values to the left of mu that is mu minus epsilon. That is x taking values to the right off mu plus epsilon and left of mu minus epsilon. For any epsilon, this is by the Thebycheff inequality less than equal to sigma square by epsilon square, but sigma is given to be 0 and epsilon however small, it is not 0.

So, this is equal to 0. That means that x is such a random variable that it always takes the value mu with probability 1 because for any epsilon chosen, its probability of taking values higher than greater than equal to x minus mu and less than equal to x 1 as mu, sorry greater than equal to x minus epsilon. Greater than equal to mu minus epsilon and mu plus epsilon or less than equal to mu minus epsilon, that is graphically it means let us if your mu is here and epsilon for any choice of epsilon probability of x taking value to the right side or to the left, that is 0. You know this is true for any epsilon greater than 0 however small.

So, you now start getting epsilon 10 to 0. Obviously from right hand side and from left hand side will be approaching mu. So, as long as you are not at mu, but very close to mu on the left hand side and right side, probability of x taking those values is 0. You can let this limit tend to this epsilon tend to 0. This obviously means the probability of x taking value mu is 1. So, you say that x takes the value mu, x is such a random variable that it takes the value mu with probability 1, and variance is 0.

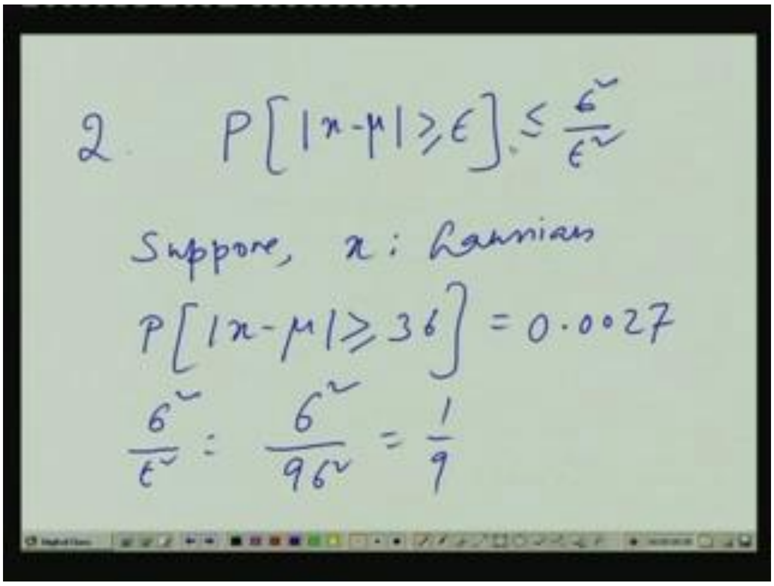
(Refer Slide Time: 25:36)



Further, x is a random variable and E of x square which is always this is always true mu square plus sigma square. That is expected value of x square which is also called the expected power. In electrical engineering language, it is expected power of x. Suppose this is given to be 0 and since, this is a square number and this is a square number and therefore, both are positive. This means mu is equal to 0. Sigma is equal to 0. Now, sigma equal to 0 means x is such random variable that takes the value mu with probability 1. That means, in all experiments we will continue to get the value of x equal to mu. What is mu here? It is 0.

So, that means, x is a random variable which will always take the value 0 with probability 1. You remember when you are discussing this vector space of random variable, then in that vector space we also considered zero random variable. We called it zero vector and that time, I made this statement that is such a vector there is such random variable which always takes the value 0 with probability 1. I mean this explains that further.

(Refer Slide Time: 26:59)

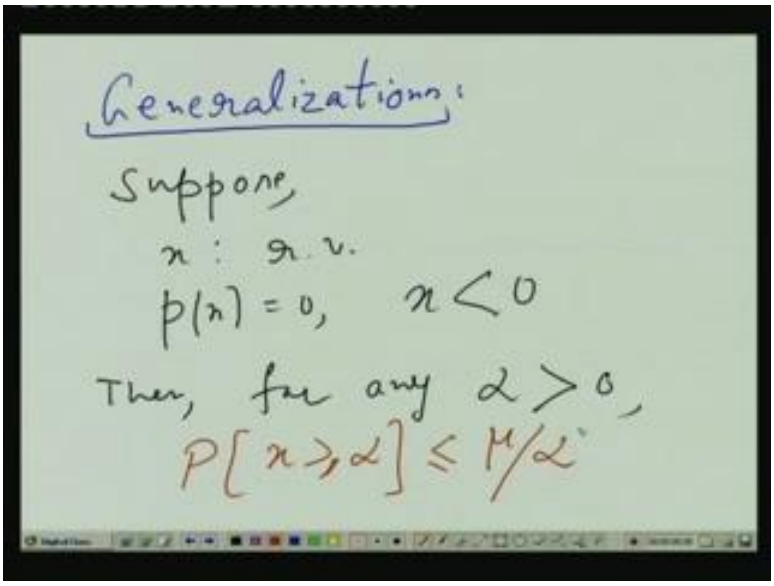


Secondly, if you consider this bound, now you know this provides an upper bound that is this probability should be less than equal to this, but sometimes this upper bound may be much higher than the exact probability. You took some epsilon, you are giving the probability density function of x and from that you can carry out calculation and find out the exact value for this probability. You could find that this actually is much less that sigma square a by epsilon x square. This you would keep in mind. This is just a bound. It shows this left hand side can never exceed the right hand side, but it in no way means that this will be close to sigma square by epsilon by epsilon square. In some cases yes, but in many cases no.

For example, if suppose x is given to be Gaussian or normal. It is Gaussian, in that case Gaussian mean mu and variant sigma square. You know what this probability density 1 by root 2 by sigma e to the power minus x minus will be whole square by twice sigma square. If this is given, then you can find out yourself that if you are chosen epsilon to be just 3 sigma, those exact value will be given by I am directly quoting from Papoulis, but what does the right hand side bound gives. It gives sigma square by epsilon square. In this case epsilon is 3 sigma. So, sigma square by 9 sigma square, it is 1 by 9 and 1 by 9 is close to 0.1 which is much higher than this.

So, just remember this probability, this left hand side is always less than equal to the right hand side. So, right hand is a bound, but it does not mean that left hand side is close to it because left hand side is probability as such can be very small compared to the right hand side. It will depend entirely on the probability density function given to you and epsilon that we choose and all that. Main importance of this bound is that this is independent of the probability density function that is applicable to x because this bound is always satisfied by any random variable of any arbitrary probability density function.

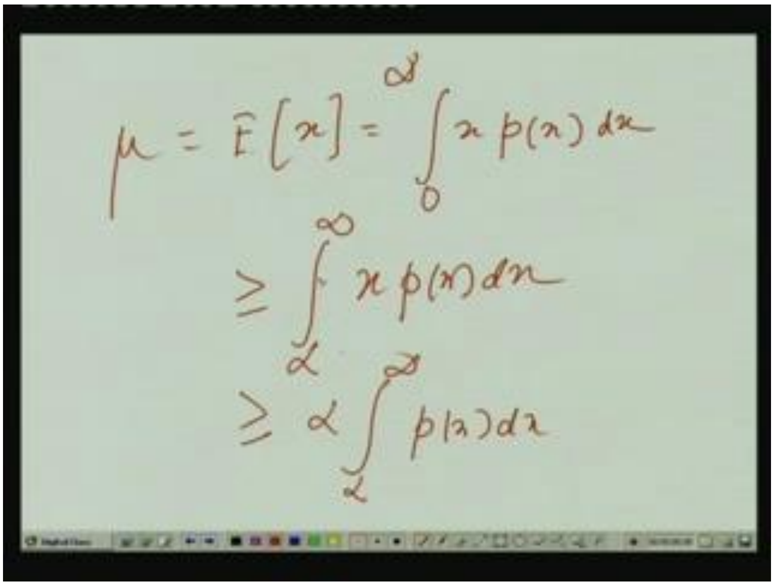
(Refer Slide Time: 30:17)



Now, consider a more general result. Suppose this is given that x random variable, but P of x equal to 0 for x less than sorry less than 0, not less than equal to. That means, x always takes values either 0 or positive term or x always takes positive values which are either positive or 0. X is always non-negative. In that case, then for any alpha greater than 0, this is also you choose some positive number. Probability of x taking values greater than equal to alpha, it is always less than equal to mu by alpha, where mu is the mean of x. This is interesting. There is no sigma here.

Please you see the difference. In this result I am only considering random variables which take non-negative values. There is either 0 or positive numbers that is P of x is 0 for x less than 0. In that case, I am saying that if you now choose any alpha which is positive, then the probability that takes values greater than equal to alpha, it simply is less than equal to mu by alpha. So, in this relation, there is no variance that is coming into operation, but this is more general result. We will come to that soon. How to show this?

(Refer Slide Time: 32:45)



Proof is very simple. What is mu? Mu is E of x. What is E of x? E of x is x into Px dx, but integral should not be from minus infinity. It should be actually from 0 to infinity.
So, I correct it. This is 0 to infinity. Now, remember within this range 0 to infinity, x is taking always values which are positive or 0, and p of x is always non-negative. So, that integral is always positive and this product x into Px is always positive, but that is being integrated. Of course, the result is positive. Mu is positive. It has to be. This means, if you now do the integral from alpha, sorry there is some types of confusion between alpha and infinity.

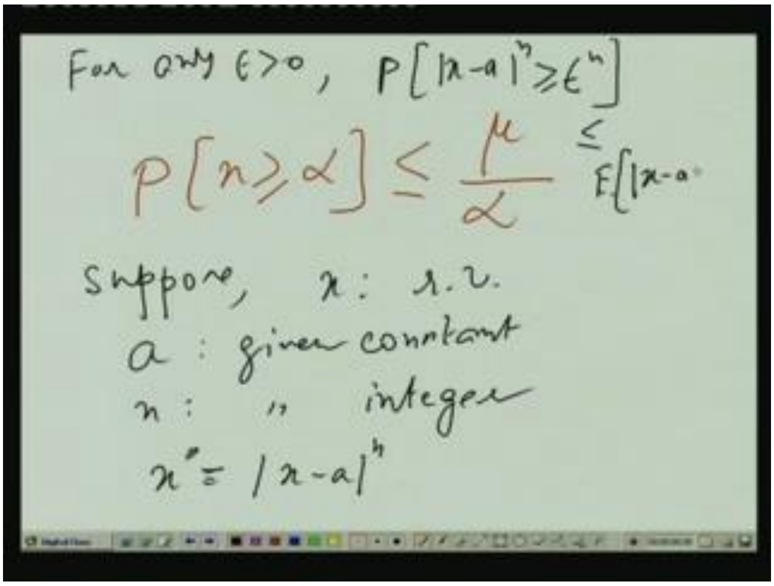
So, this alpha and this is infinity. The alpha to infinity where alpha is positive, if we just carry out this integral, then this must be less than equal to the previous integral which was very larger range from 0 to infinity. Isn't it? So, this is less than and equal to the previous one and now, within this integral we look at the function x. It starts at alpha some positive number and goes up, up, up, up, up, up, up to infinity, and multiplied by the respective value of Px.

Now, if you state and if you do not allow x to go up, but we hold it at alpha only, but let Px continue as x moves from as that is we carry another integral like this, where as far this function P of x is constant, we let x move from alpha to infinity, but as far as x in the previous integral is constant, we do not allow to vary from alpha to infinity. We will hold it at its minimum value alpha. Then, obviously since everything is positive, the integral now will be less in that value than the previous integral because earlier both x and Px, they multiplied and integrated x was first taking value alpha and then, going up, up, up, up, but now I am multiplying Px only by alpha. Even if x is varying from alpha to infinity, Px is multiplied only by alpha all along and not by some quantity which is becoming higher and higher.

So, obviously, the net value of the integral will go down. So, this will be further less than equal to this and outside alpha. By the way if I ask you why I am putting less than equal to and not less than in both cases? Answer is, it is just for mathematical correctness. For instance, initially I said that x into Px dx integrated from 0 to infinity, that should be greater than equal to the same integral, but from alpha to infinity why greater than equal to. Because it may so happen that your probability density function of P of x is chosen such that from 0 to alpha, its value is 0. So, obviously the integral itself becomes identical to this. This will follow next integral.

There is x Px dx from alpha to infinity because being in the range between 0 to infinity, P of x is 0. So, there is no point in carrying out the integral there. It will take the zero values. In that case, equality will be given up. Then, here how you can take alpha out?
It may so happen in a different context that as you go beyond alpha, Px comes down to 0. In that case whether x is increasing or x is kept at alpha, it hardly matters. That is why less than equal to. What is P of x? It is dx from alpha to infinity. This is this side is then equal to alpha into there is nothing probability of x taking values greater than equal to alpha. That is why you integrate Px with respect to x from alpha to infinity and that will give the total probability of x taking values greater than equal to alpha. So, that proves inequality.

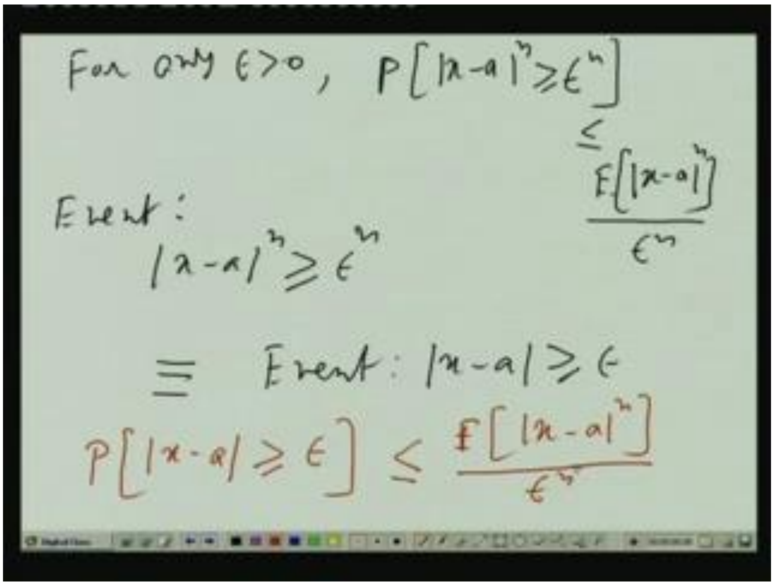
(Refer Slide Time: 38:28)



Now, suppose there is a random variable, a given number, a given constant n given integer. You consider a random variable x prime which is nothing but say mod of x minus a whole to deeper n. So, obviously mod of x prime is a random variable that takes only positive values. That means I can apply the above result. If I choose that is the above means, then that for any epsilon greater than 0, any positive number positive epsilon greater than 0. I can write from this that probability of this random variable x minus a to the whole power n. This is random variable. I call it x prime. This is greater than equal to epsilon to the power n. That should be from this result less than equal to expected value of this random variable x prime divided by epsilon to the power n, that is expected value of this random variable x prime that x minus n mod to the whole n like here. Expected value of x was mu.

Now, a random variable is x prime which is equal to mod of x minus into the whole power n. So, I am taking the expected value of that divided by the quantity to the right of this greater than equal to sin. Earlier it was alpha. So, alpha came as the denominator. Now, it is epsilon to the power n. So, it will be epsilon to the power n.

(Refer Slide Time: 41:38)

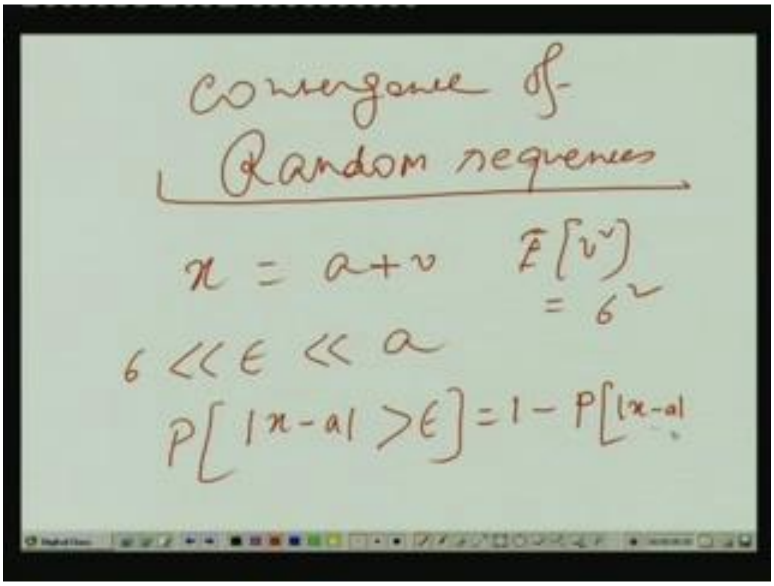


What does that mean is this is fine, but my claim is probability of actually let me write, remove it about it way, it is not probability. This event is identical to this event because whether you take mod of x minus a or mod of x minus a into the power n mod of x minus a is always positive. Epsilon is positive. So, if it is given that mod of minus say whole to the power n, it is greater than equal to epsilon n. So, it is the continuous are positive. It obviously means mod of n is positive, n is number, n is integer does not matter whether it is positive or negative. Though essential thing is that mod of x minus a and epsilon both are positive. If the above meaning that is given, then mod of x minus a to the whole power n is greater than equal to epsilon to the power n.

This obviously means mod of x minus a is greater than equal epsilon, conversely if this is given that mod of x minus say is greater than equal to epsilon. Obviously if you raise the both left hand side and right hand side to any power n, this is both are positive. Both mod of x minus a and epsilon are positive. If you raise to any power even, then left hand side will be greater than equal to right hand side. So, the two events are same. One implies the other. That means from other I can also write that probability of since this is interesting in the Thebycheff inequality, our left hand side was this. Only a was replaced by mu.

What is more general? P of probability of mod x minus say greater than equal to epsilon , but on the right hand side you are free to choose n, but left hand side, then it is always less than equal to right hand side for any n. So, you can put n equal to 1. Still it is satisfied. You can put n equal to 2. Still it is satisfied so on and so forth. You can now see that if n equal to 2, if you take n equal to 2 and a equal to mu, that is mean of x. In that case, you simply get back your Thebycheff inequality, right. So, this is much more general result.

(Refer Slide Time: 45:14)



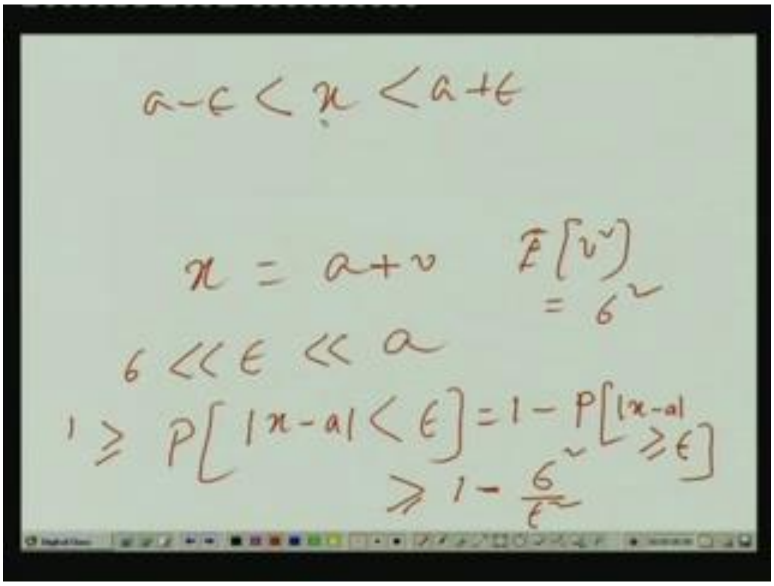
Now, in the limited time that is left I will just get the concept of I will just start this discussion on stochastic convergence, convergence of random sequences, but again I will need some further discussions before we can take that up. You know that given a random sequence, I mean finding out a notion of its convergence is fundamental probability to the theory of probability as I told in the beginning. So, we again try to develop this notions using those Thebycheff inequality and related concepts here.

Suppose you are just conducting a very simple experiment, you are measuring the length of something, and actual length is may be a, but what you are getting is not a, but you are getting some error in electric. In our communication electrical language, electrical engineering language, you call it noise. So, you are measuring, you are observing is x which is not equal to a. A is that sure length. It is not known to you. It is unknown. You are trying to find out what is actually a. Now, unfortunately every time we measure, you get some different readings, and this is a plus v. V actually is just a random error. It is like a noise. It is not a systematic error.

If it is a systematic error, it can be corrected or at least it is deterministic. So, it can be subtracted. It is not in your hand. It is just a random thing, so it is like a noise term or error term, random error term. So, x is equal to a plus v. So, instead of finding out a, you are only getting x. Then, you can easily see from our previous discussion and from here that if v is such a random variable, v is a random a variable such that its variance is much less, then x will be close to a always. This is what we have elaborated so far in our discussion that a was a mu that time. That means the random; the x minus mu, the deviation will not go far. It will be within some small range around mu because the variance of v or variance of x minus a or even variance of x is much less.

If that for Thebycheff inequality, it shows that suppose it is given that is E of it is given that v square is sigma square and if it is given that sigma is much less than some number epsilon which is again much less than a, that is sigma is much less than a and you can find out take some number epsilon in between, so that sigma is itself is much less than epsilon which is again much less than a. You can then write within Thebycheff inequality that probability of less than equal to that time, we will consider it to be greater than equal to. Now, you consider it. This is now necessary greater than epsilon which is nothing but 1 minus this is, sorry.

(Refer Slide Time: 49:29)

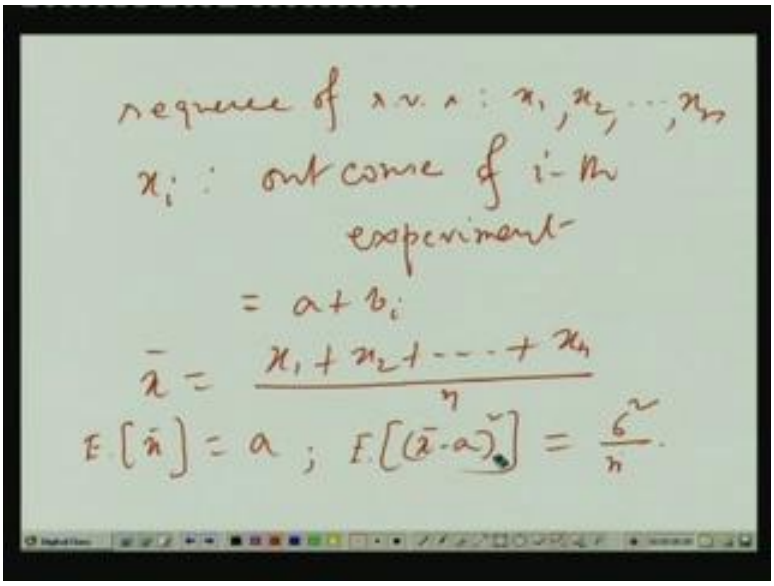


Obviously, this is I mean this thing probability of mod x minus x greater than equal to epsilon was less than equal to sigma square by epsilon square. So, you put a negative sign, so less than equal to becomes greater than equal to and add 1. So, greater than equal to simply becomes greater than. So, this is nothing but greater than sigma square by epsilon square. Since, sigma is much less than epsilon, what it means that x minus a that is, sorry this will mean that x will be in this range with high probability that the probability of that will be greater than 1 minus this, just a minute. It was mistake here. I left out one term epsilon. It is actually than equal to 1 minus this.

So, sigma much less than epsilon, this is almost close to 1, but of course probability cannot exceed 1. This is always less than equal to 1 from left hand side, but if sigma is really very close to very much less than epsilon square, this is just 1 minus some very small amount 0.000 something. So, we can loosely say using probabilistic language that x will almost certainly, almost surely will be within the range a plus epsilon to a minus epsilon and since, epsilon itself is much less than a as I told in the beginning, I assumed x will be very small and x will be very close to a or almost equal to a.

So, in that case it will be safe to take the particular x that you measure as a good estimate of a. You do not have to repeat the experiment again and again. We do not have to look for further trails because if sigma square is less than epsilon square, this itself is very close to 1, because it is less than, this is almost 0. So, its probabilities between 1 and some point at which it is very close to 1. So, you say that x is almost certainly within this range a plus epsilon to a minus epsilon. So, since again epsilon is much less than a, x is within a plus to a minus you can say. So, whatever x you get, you can simply take into the estimate of a. No problem, but suppose this is not true that sigma square is not really or sigma is not really much less than a, so you cannot apply this. Then, what we do that is what the sequence of random variable comes actually that suppose we perform the experiments several times, n times.

(Refer Slide Time: 53:05)

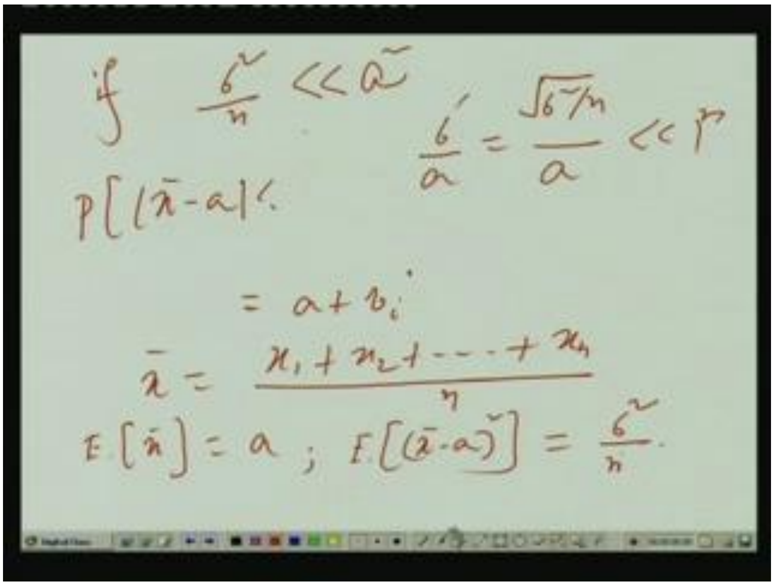


First time you get x 1. Whatever you observe if it is name random variable x 1, then you observe x 2, then you observe x 3. So, you will get a sequence of r v random variables, x 1 x 2 dot dot dot may be x n. Xi outcome of it h experiment. This experiment and xi is equal to a plus some noise term giving the ith experiments. The noise term that comes that is vi. So, actually if you again repeat the experiments, first I mean you are making n observations x 1 x 2 x n. After half an hour may be you can make again another x n observations. So, in the first observation, vi took some value. Next time ith step, you will get some other value of vi. In that sense, vi is a random variable of variance sigma i.

In fact, to make life simple, you can assume that vi whether it is v 1 or v 2 or v n, they all have the same variances. After all it is coming from the same error process. Now, after making n trails that is you got some values for this n random variables. Out of this sequence of your x trails, first trail x 1 you got the value of x 1, then 2, then x n you measure. That is together. This consists of one experiment. Next, you again measure x 1 x 2 x n a likewise, but I am just bothered about only experiment where you measure some value of x 1 only, x 2 and x n. Stop you are not measuring again. What you are doing? You are finding out a sample estimate which then again is a random variable.

Now, obviously what is the mean of this? Mean of xi is a because vi is 0. Mean may be I forgot to mention that this error you know it can take positive value or negative value. Its mean is 0. That means, each xi has a mean a. So, what is the mean of x bar mean, that is a n times n into a divided n. So, E of x bar is same as a, but what is x bar minus a whole square variance. That will be you can see sigma square by n square. That is a beauty. Sigma square by sorry n not n square n. So, variance is coming down. Earlier it was sigma square and I said if sigma is not much less than a, we have a problem. We cannot simply measure once and take it to a good estimate of a. In that case, I measure it n times and take a sample average. This sample average if I call it random variable, its variance is not sigma square, but sigma square by n and if sigma square by n.

(Refer Slide Time: 56:55)



Sorry I am about to finish just one minute. This is much less than a square because it is sigma square. Suppose this is much less than this, then obviously we can invoke that inequality. If it is much less than square, then whatever value of x bar you measure by this sample average, that is good enough. That will be close to the mean a because variance what does it mean. The variance, new variance if you call it sigma prime which is nothing but sigma, sorry square root of sigma square by n. What is mean? Mean is a. So, that means this ratio is much less than 1. So, I can use that previous theory and I can say that probability of you know less than equal to that is or may be less than some epsilon. It is almost close to 1. So, whatever you measure for x bar, this is good enough. So, that is all for today. So, in the next class, we go further into these notions.

Thank you very much.