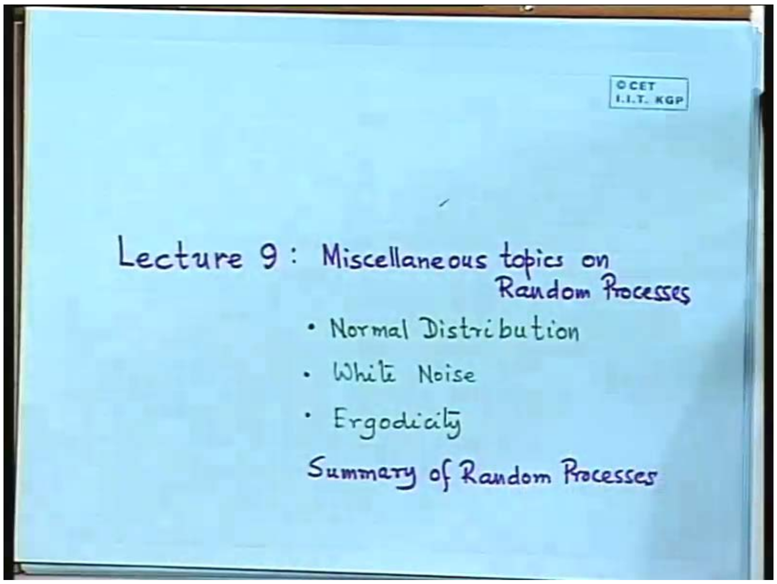
Estimation of Signals and Systems
Prof. S. Mukhopadhyay
Department of Electrical Engineering
Indian Institute of Technology, Kharagpur

Lecture - 09
Miscellaneous topics on Random Processes

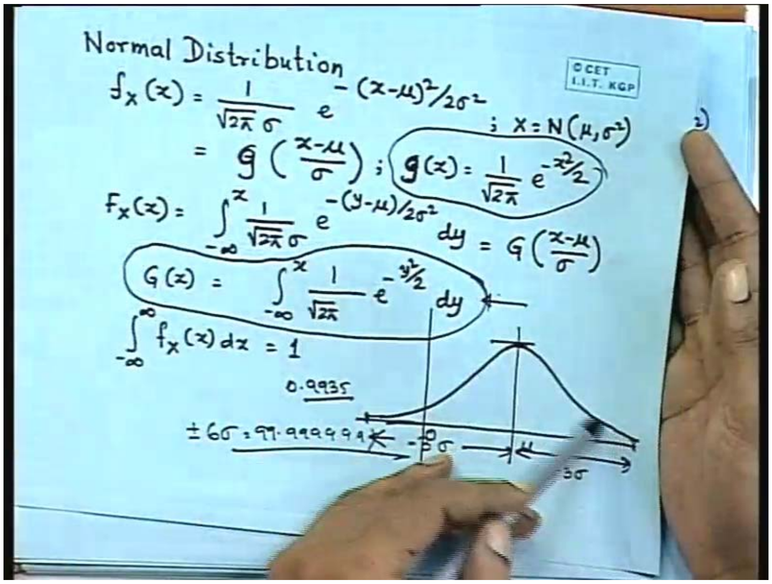
So good morning, today we will have a last lecture on random processes. And from next week onwards, we will move into actual estimation.

(Refer Slide Time: 01:04)



So, before ending I thought that, there are some topics left here and there, so we should discuss them little bit then, we should briefly summarize what we have done. And perhaps try to take a brief look on, what we are going to do. So this three things came into my mind, because many times we will be talking about them and there are certain facts, which would which we should know before we close our discussion on random processes. The first thing that is I thought, we should discuss the normal distribution.

(Refer Slide Time: 01:57)

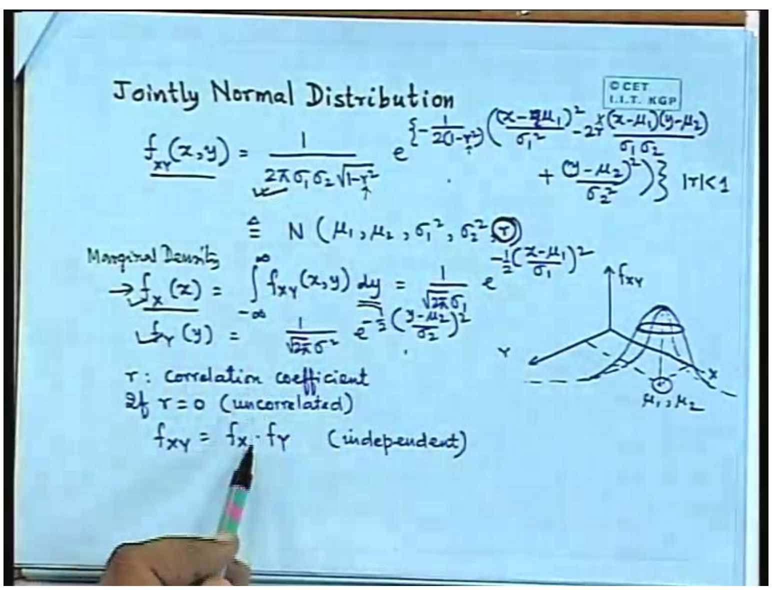


As such we have not discussed, any of the distributions as a case there are many examples in books. There are various kinds of distributions, which are used in various kinds of applications. For example, if you read a book on KWing theory, you will get mostly get; you will mostly get across the Poisson distribution, an exponential distribution. Which you read a book on reliability; you will come across the Weibull distribution. Similarly if you read a book on signal processing, you would very much come across the normal distribution. That is so let's take a brief look at the normal distribution, it is a it is a very special candidative which is very often quoted. So this is the normal distribution; I think you know it this is the normal distribution of a single variable and in fact.

Its values are sometimes tabulated in charts; in terms of a normalized normal distribution, that is a function this function is a normalized form and the general that is this is the normal distribution, with mean mu and variance sigma square. And as you can understand that, the whole distribution itself is given in terms of mu and sigma square. So if mu and sigma square are known, then the whole.. then the whole probability distribution is known. This is one of the biggest advantages of of the normal distribution, that the whole probability density function is characterized, in terms of what is known as first and second order moments. In general for a I mean a probability distribution is a much more detailed information, rather than the first and second order moment; but in the case of this distribution, if you know the mean and and the standard deviation, then you know the whole distribution. So it only that much needs to be known, that is one advantage.

This function is sometimes called error function; this function is.. is also in many cases tabulated. That is this minus in this is you will find that there are various you know, statistical charge in which this function and this function are generally, this function is tabulated. So if you are given any other mean, you have to first transform that your distribution to, to this distribution and then if you want to get get, the value because, this is a very odd integral. It is not easy to get the value directly, analytically. So this is sometimes tabulated and obviously; though though it looks like this it can be it can be verified, that this is itself a this is indeed, a distribution function and minus infinity to plus infinity a x is one. In fact to make it one, this this this factor has been added otherwise; this integral itself will not give one, so just to make it one this factor is added. So this is the normal distribution, which is somewhat known.

(Refer Slide Time: 05:29)



But we in general, I have talking about; will be dealing with various quantities. So we have to first understand, what is the jointly normal distribution. So jointly normal distribution is you know; so so one thing I did not draw is that the normal distribution curve. So the normal distribution is in general is a bell shaped curve, whose so that that that bell peak arrive.. happens at mu, may be this is the origin of the axis. It is not always necessary that, it will be at zero, maybe this is zero. And the sigma is such that you know, there is a in the sometimes I do not know sometimes in those whose those of you who are some familiar with management, sometimes have might have heard of the term, called six sigma. It is a it is a quality control practice and that name six sigma is basically derived from the normal distribution, okay? So if you see the normal distribution, then the then if you take plus minus six sigma suppose; the normal distribution has a standard deviation of sigma, then ninety-nine point I think three percent of the area falls within plus minus three sigma.

The this side is three sigma, this side is three sigma. So this side, plus three sigma, this side minus three sigma. If you take this limit, then the ninety-nine percent then, then the area under this becomes nine point nine nine three five or something. So ninety-nine and the whole area is, from minus infinity to plus infinity is one. So ninety-nine percent of the area falls, within that band, right. That is why in in many engineering applications, when we characterize, when we describe variations of parameters, then we often say that this is a plus minus three sigma limit. That is the parameter variation is in general random, in all engineering designs, there are some parameter variations, which occur. And you have to take care of such variations, in your design.

So so it is common practice to to characterize, that your design has been has been verified under what parameter variation. So sometimes people say that, it is verified under plus minus three sigma variations; now now in management, they are saying that, that you should have such practices, such that plus minus six sigma of your products, produced products will actually meet quality. There will be no all products that, so if you take plus minus six sigma, then this becomes some ninety-nine point nine nine nine nine nine percent. So when the say in management that, some companies you know six sigma and a has now become a quality standard. So if if if some some company six sigma compatible; then they are trying to make a claim, that their products that so many of their products meet the meet the stated quality, right? So that name in in management, also derive six origins from the normal distribution.

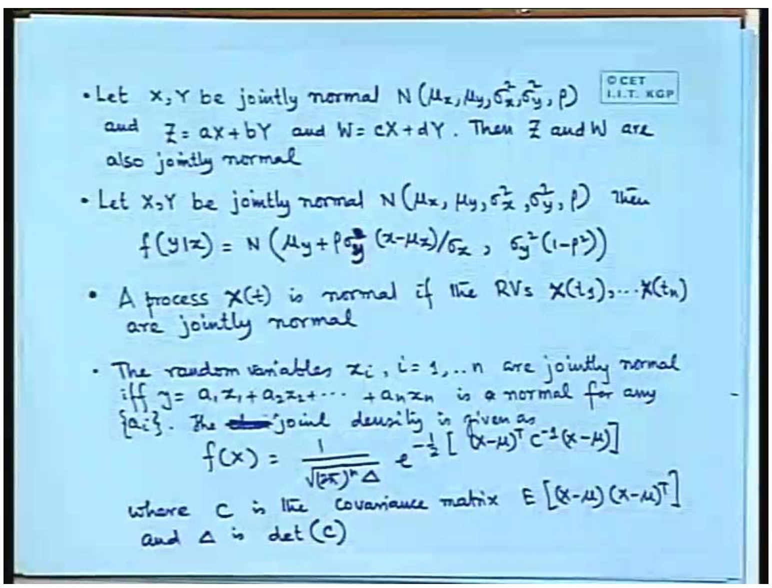
So if if in one dimension; it is the it is this curve, then in two dimensions this is going to be like, a dome which is circular in cross section, right. So have tried try to draw that. So if you have two..two random variables; which are suppose to be jointly normal, then you get a form like this. And the and the shape looks like this, where this is the mean. So on the x axis, you have a mean mu one, on the y axis you have a mean mu two. And that is characterize, so you see that it is similar and mu one, mu two, sigma one square, sigma two square, but there is a correlation coefficient. So two random variables; may be correlated to a certain degree and in general they are they may be jointly normal and correlated. And it turns out that, if you take thus, these are what these are called the marginal densities, if you recall. So from this you can get only the variation, the probability density of x by my integration, it getting with respect to y, remember. If you do that, then the individual ones become Gaussian densities; so both this and this are Gaussian, individually and they are jointly Gaussian too, right.

And and you can also verify that, if r is equal to zero which means that, they are uncorrelated. Then they then they also become independent because, if r is equal to zero then this term will vanish, then it becomes e to the power this. All this will become one and this term will become..simply; e to the power minus this plus this which means that, e to the which means that f x y equal to f x into f y. So then then there will become an independent also. So if they are uncorrelated, they are independent. This is a direct consequence of the fact that, the distribution function is fully characterized, by the second order moment. So... if the so if the second order moment satisfies the uncorrelated property, then the whole distribution becomes independent, which is not the case for general distributions.

Now, what is so great about this distribution? I mean why do we, why is it so common? That is because you know, what what are we going to do as really. You will see that, we will in our in most of signal processing and estimation theory, people generally work with linear models. That is because, it simpler otherwise; you know on one side you have this probabilistic uncertainty and random things. Now if you put non-linear systems into that, then it will blow up. I mean it is very difficult to get results. So people first try to see we would that, how let us have some randomness and then then try to see, how it transforms through linear systems. So people largely use linear models and try to get results, and you get e. I mean in in many applications, you can get reasonably good results, inside used and right.

Now it turns out, so what are we going to any any if the moment we use linear models? What can of computation, that we going to do? We are going to basically multiply things by constants and add. So we are going to do linear operations, so all variables we will be going through one after the other various kinds of linear transformations, when we are using a linear model in our computation. So now the point is that that, will mean that the probability distributions of the input variables, will get continuously, get changed. So one of the biggest advantages of the of the Gaussian variable is that, the they they maintain their Gaussianness with respect to linear operations.

(Refer Slide Time: 13:04)



This is one of the biggest reason, is why we always tend to assume Gaussian variables, when we talk especially when we discuss signal processing. Say almost I mean, some people think that white noise cannot be anything other than Gaussian, I mean Gaussian white noise becomes a term. Actually, there is there is there is no relationship between a noise being white and its pdf being Gaussian, they are show two different thing, that quite unrelated. But but sometimes I mean, we are we get so much use that, we always think that I mean if we have a noise, it must be Gaussian. It is it it become so common, I mean sometimes in books etcetera.

So so these are some of the properties, which which are useful. So first thing is the simple version of it which says that, it if x and y are jointly normal, then you make any linear transformation variable Z by ax plus by. And you take another linear transformation W, which is c x plus d y, then then then, Z or W Z and W are also going to be jointly normal. So once you have assumed one to be normal, if you even if you do linear.. I mean I mean I mean, arbitrary linear operations on them, they also remain normal. So so they remain normal means, they are they are still also characterizable; in terms of their means and standard deviations.

So whatever operation you do.. the linear the all you need to characterize, to get full information, is mean and and standard deviation. You you only need to know, how these are transformed during your linear signal processing nothing else. That is that is that is a big advantage, I mean that is why you will see that, you always try to characterize these things. And if we I have if we have assumed Gaussian; then we assure that if we can characterize this two things, we have characterized everything, that is the advantage. There is also the there are also other facts; for example, it turns out that a a large number of real random variables, had do tend to approximate the Gaussian curve reasonably. There are large number of random.., where I mean random variables, which we which occur practically, which will tend to occur the tend to approximate. I mean there there there if you if you if you actually, carry out an experiment and plot a histogram, you know I mean how do you plot this probability density function in practice?

You have to do a large number of experiments and then you have to plot, what is known as histogram. I think we learn that in school that, we we will characterize that. Okay, how many boys are there are within this group x one to x two? And then how many boys are there within x two to x three and then will plot them plot the histogram. So if we take a large range and if we take small small ranges, then it will it will try to approximate this. I mean that is a that is how we draw the probability density function. So that is also another reason, that it actually occurs in practice, in a large number of cases. So that is so people are I mean people are happy, that it is a practical distribution, which has nice properties. That is why people use it in this application; in in in other kinds of applications, people will use other other other distributions. But in our application, Gaussian will be useful.

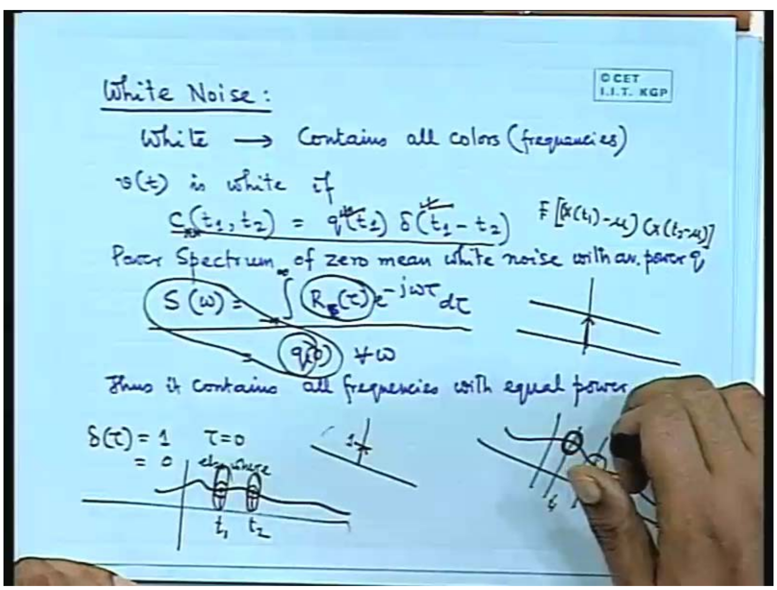
Similarly if x and y are jointly normal, then the conditional density of y given x also happens to be normal, okay? So in many cases, we we will see see this is a very common thing that will do because, we are in many cases we are see in general about uncertain things, what are we interested in knowing? Our whole course, it is about estimation and from the next class onward, we will try to see that how we can get the get a get a good guess about something in the phase of randomness. That is the whole, that is that is what the whole course is about. And one of the main things that you would like, to know is what is expectant that is in some probabilistic sense? What is the average? So so the so the expected value, in general is of great interest in this course.

And and and obviously; we are not trying to find the expected value in absence of any information, so we are we we in general will will try to get as much as much prior information as possible, and then try to find out what is expected, is it not? Because we are always trying to make guesses in certain specific cases. So we will always try to incorporate the the information about those cases; whether the signal is what is the I mean what kind of signal, we are applying suppose, we are trying to find out what is going to be the output, so obviously we will we will try to incorporate information about, what is going to be the input or right. So so we are always so we are not in general; only interested in probability densities, we are in many cases interested in conditional probability densities, because we already have some facts available. So we will try to incorporate them in our estimate, okay. That is what the conditional estimate does. It does not get y for I mean arbitrary values of x, but it finds out that value of x which is which has occurred, probably something else and then from there try to guess, what is the what is the expected value of y, that is why it is so important.

So it turns out that, this conditional estimate also happens to be normal. So again you can only you only need to characterize covariance and mean. So this is the this is this is the central reason, in why in signal processing people will always tend to use, a Gaussian noise. And a if you in general if you have in general, if you have this result; that if there are n random variables, if you make any arbitrary linear combination of this this random variable, say a one x, one plus a two x, two plus a three x three, Then for arbitrary a one, a two, this variable y is going to be normal, right. That is a that the that is a good thing which means that, which means that if you take a vector, whose all components are actually...normal variables, then you make any linear transformation of the vector multiplied by any constant matrix. The resulting vector is also going to be jointly normal and in all linear processing we only do that. We only multiply signals by constant matrices everywhere, right. So so we are going to always remain normal, that is a that is such a nice mathematical advantage. So this is this is one point which I wanted to make because, sometimes people wonder that, why is it that I am not taking a Weibull distribution of noise? If I if you take a Weibull distribution of noise, we will quickly get lost after one or two transformation, we we won't know what distribution the output will have, that is the problem.

Next we talk of white noise. So we will many we will in in many cases, again assume that the the input is white noise or or we want to make something white, we want to make the innovations white.

(Refer Slide Time: 20:17)



We will we will we will get these things in our course, very often so we will we should know what is white noise. So white noise is is something actually, this is a stated in terms of a okay, you can interpret it in the discrete context also. So that if a noise is called white, if it is if it is covariance, this is the covariance, covariance means expectation of x minus x t minus mu. Let us take a stationary process, so mu constant. So so covariance c t one, t two is what? It is expectation of x t one minus mu into x t two minus mu. This is auto covariance; you can have cross covariance also, okay in general. So it what are it is it what is this what is this, delta function in case of continuous; this delta will is the what is known as the draft delta, there are two kinds of deltas, one kind is called the draft delta and the other is called the chronica's delta.

The chronica's delta is the is the is the discrete delta function; which says that, for the chronica's delta, delta tau is equal to one, for tau is equal to zero and is equal to zero elsewhere. So it is plotted, it looks like a function like this, its value is one. In the case of draft delta, its it is its value is infinite but, but the but the integral under it is one, it is it is unit area. So so, what does it basically say? So whether, it is continuous or discrete? It basically says that, if you that the two values of the random, if you take two different random variables which are even a small time apart, they are going to be uncorrelated. Because this value will be zero, if t one minus t two equal to zero only then, this has a value. If t two is any different from t one; then this delta function is going to be zero, because it is zero everywhere else, other other than at zero. So which means that, this covariance function is going to be zero for all t one t two such that, t one is not equal to t two. It is non-zero only when you calculate, c t one t one otherwise, if you take any other t two it will be zero.

So the so all random variables, so so the so if you take a, if you take a random signal I mean, which is according that is one instant of the random process, then if you take the random variable at any t one and if you take a random variable at any t two, then these two random variables are going to be uncorrelated. And and this will happen for arbitrary, toward, right. That means, that things are really random I mean there is there is you cannot extract any information, if you if you get property of let let say, if you if you measure any one of them, you cannot say anything about the other one because, it totally uncorrelated. See correlation generally let see, guess what is those suppose two variables are positively correlated? Then on an average, if one is positive then, the other is also positive. So you can at least make, some guess about it. But, uncorrelated means, that the individual random variables at the various times are completely uncorrelated.

For example, what is going to happen two minutes later, even if you have data obtain the on that thing. If you even if you have data about the behavior of of an object for the last ten years; you cannot say, what it is going to do in the in the in the next two seconds, because it is because all the this whole behavior over the last ten years, every point of it is actually uncorrelated with that, with the next two seconds. So everything is uncorrelated so it is a its an ideally random process. In that sense it is it is ideal, you cannot extract any information about anything, from any other point they are..each one is independent, right. So it is in a sense, it is the peak of randomness, okay, as for as second order moments are concerned.

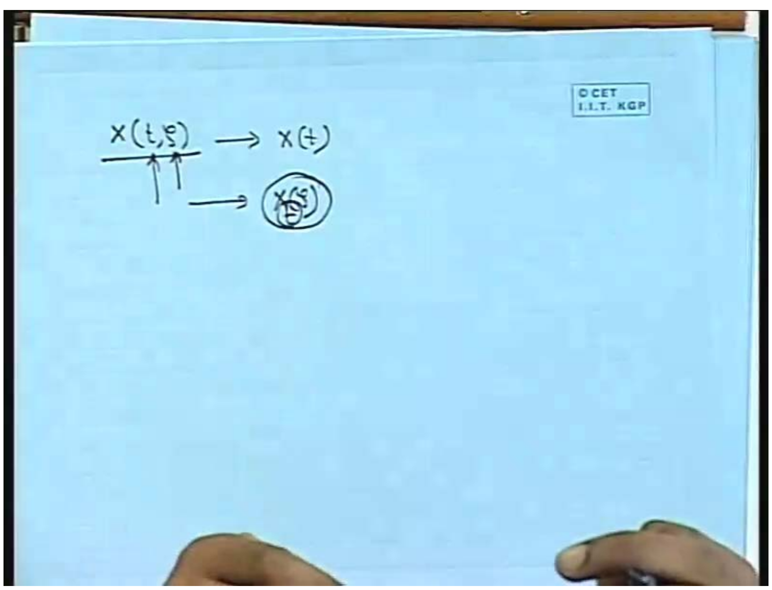
And and and why it is called white because..., in general when we say white light I mean that is how the name has come white light. Means it contains all all spectrum of the VIBGYOR, right. So an spectra are actually related with frequency, because color comes from frequency. So it means that, it contains all frequencies. Why it contains all frequencies?, that comes from this integral. So now if you see the power spectrum of power spectrum, as actually is defined like this, right. This should be minus infinity to plus infinity. Now this R tau is this, so only at tau equal to zero this is going to have, this is going to be non-zero, at every other point is going to be zero, this is going to be zero. So this integral value is always Q, actually Q zero,.. right. Which means that, irrespective of omega, so now see S S S is actually should be a function of omega, but it is not a function of omega, it is a constant. Which means that, at all omega, S omega is q. Which means that, power is uniformly distributed over all frequencies omega, so there is equal power in each frequency. That is why it is I mean, there is a people try to draw a similarity between white noise and white light. Because white light also contains all frequencies, that is how the name has come, right.

And I mean there are there are various, because of because of this noise property of the because of this nice nature. So you see that, we have to remember that the power spectral density of white light is like this, of of white noise is like this while its autocorrelation function is like this, okay.

So so this this we should remember, because this we are going to use in in many cases. This is this is but but one thing, I would like to mention again that, this you must remember that again that, white noise has no relationship with Gaussian. You can have.. you can easily have uniform white noise, there is there is no problem about it. Actually let us if you see, what is what does what does whiteness talk about? If you have if you have is if you have a if you have a random process; whiteness comes from the property of the behavior of one random variable, at any given instant of time, one and another random variable, at another given instant of time t two.

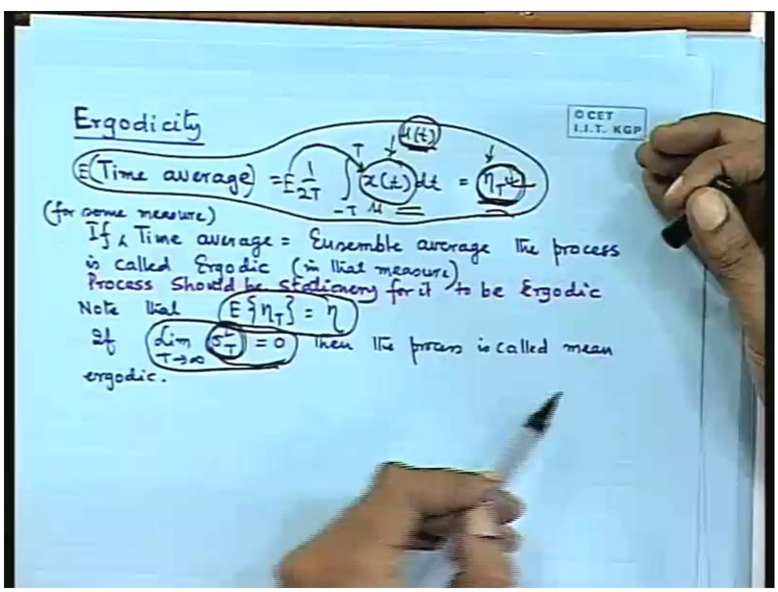
So whiteness trying, tries to characterize the joint properties of two random variables, at two different instants of time in fact, the because it is stated in terms of the correlation function. White the signal is called white, if in a if it has a certain correlation function. So whiteness merely says that, the correlation function is of a certain type, it says nothing about what is the probability distribution of this individual variables. The when we say Gaussian, we are saying the that if we generate several such random processes, and take e for each one of them, if we fix the time t one remember that, that if you general in general a a random process is a function of two things, t and xi.

(Refer Slide Time:29:05)



If you freeze xi, you get a function x t. If you freeze t, you get a random variable, x x i at t, right. So when we are talking, a when we are saying Gaussian, Gaussian refers to a random variable, it is a it it it it is a it is a density function. So it must refer to a variable. So when we are saying Gaussian; we are talking about random variables at a given time, at a given time. So if you take make many such random processes and at at every time sample at one second and take the value and then if we plot those values, then we will get a Gaussian probability density function. That is what we mean, but that statement says says nothing about, how these variables are going to behave if you take two random variables at time t one and t two. What is going to be their their joint behavior, that does not that that is not stated in the word Gaussian; Gaussian is concerned about a one time, while whiteness is essentially concerned about two different times or or or in general n different times. So there are two different things we only assume, Gaussian and white for analytical simplicity, but it it is not it is in no way implied that, I mean a all white noise needs to be Gaussian. So this is another point that, I wanted to make finally, you will sometimes again assume that things are ergodic now. What is ergodic means, you know one of the one of one of the big problems of especially, when we will do signal estimation.

(Refer Slide Time: 30:51)



We will in a in many cases; we have to try to compute statistical quantities, that will try to do things in a statistically optimal way. Now point is that, doing things in statistically optimal way is means that, we we have to we will have to extract statistical information things like, what is expectation and etcetera. Now the point is that, how can we get expectation while I mean how can we evaluate expectation, when we are doing something, let us say in real time. Because if you want to do expectation essentially, you will have to carry out many experiments, because expectation all the averages that we have talked about till now are all ensemble average, sometime, right. So ensemble average means that, you are suppose to carry out a large number of experiments, get a large number of outcomes and then average over them, right.

And the point is it is very inconvenient to to to do an ensemble average, when you are when you are doing something in real time. I mean you are you are already in the middle of doing something, how can you do an experiment? You cannot do an experiment. So what so, but you have to try to get an get an idea of statistical quantities, based on what you have. Now now when, you are when you are doing an experiment, what you have? You have functions of time. So you have spent time from you you achieve to doing, so you already have passed values of a signal. So you have one particular signal at various time instants, you have in general this is the situation but in general, you cannot do an experiments. So the question is that, so people ask that whether it is possible that, if we do some operation based on this time signals, then also we will get the statistical things, right.

For example, can I get the mean see the mean, is a mean of a signal is it ensemble average you have to do n experiments, see values then then add up. But what they are saying is that, can I calculate a time average and then, expect that the time average will somehow be equal to the ensemble average. Remember that, if this is a random variable, this is also a random variable naturally, integral of a random variable over a finite time. Now the question is that, can I expect that that this will be the mean of this? I am not doing ensemble average. I am have time values of the random processes, at various times. Can I expect that this will be the mean? Now firstly, you you will be re call that this thing applies only, for stationary process because, if if this mean in general this mean is a function of time in general, for general random processes.. expectation of x t will be mu t. It depends on at what t you are doing it. So obviously, if you do time average and and if it is a function of time, this will never match because because, the time average does not have time argument at all, this a constant number. So this this this discussion applies only to the case, when the mean is a constant that means to stationary processes, okay.

So so now question is that, when... even for those, when can you say that this will this will be that this will be the mean? So that first of all let us know know that suppose; this is this I am saying that, it is a estimator of the mean after all I I am doing some computation based on some random variables and then trying to say that, is it the mean I mean I mean I am trying to kind of estimate the mean through this. So this integral can itself be thought of as an as an estimator. Remember that is in the in the in the last class, we were doing that that if you take a random variable x one, x two, x three, x n and then take all of them and then sum and then divide by n then, it becomes some mu n? We were doing that problem, so that mu n is nothing but an estimator of mu. So here also this is a this an estimator of mu.

Now the now the question is what are the properties of this estimator? Is this estimator good, so the first property is that expectation of eta t that is if you take an ensemble average of this, it will the mean. Why? Very simple because if you take expectation of this; you will put expectation of this. And this expectation will now move here, move inside the integral and then, it can be replaced by mu. So it will be minus T to T mu d t one by two, T which is mu. So at least the expectation of this estimator is the true value, so it is in that sense, it is an unbiased estimator. But what is the use of the estimator being what is the use of the expectation being equal to the mean, in a given case will it be close that is what we are interested, right. So for that we have to know that, this is as T tends to infinity whether this standard deviation also becomes zero; remember we were saying remember that, if we have x one, x two, x ten then as we increase n standard deviation, become sigma square by n, we had we had done in the last class. So there as n nu increase, that sum is nothing but this integral in the discrete case.

So there we were finding that as if; you increase this capital T that is do the integration over large and large intervals, then you will get. You are you are you are you are you are time average will tend to, match the ensemble average more and more closely, in the sense that if you take a do a particular signal, then do it, you may be away but, you are not likely to be very much away.

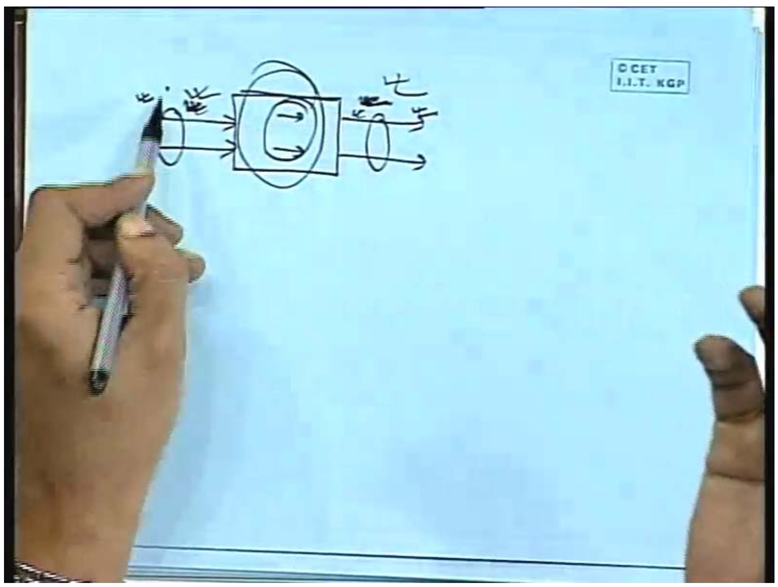
The the more interval, you actually integrate on the better, will be your case. But there are some kinds of processes, but you can actually show that by doing a sufficiently long amount of long integration, you can actually get the real true value. That is if you do it very long, then every time you will get the mean, so such processes are called mean ergodic processes, okay.

Now one of the advantages of that is that, see we as I said that, why we need to make this assumption sometimes because; we we will when we will compute real algorithms then, we will not have ensemble average, then and we will we will still need to generate estimates of the mean of signals. So we will say that, we will we will average them in mean time. And then under the assumption that the processes is is ergodic, we will say that, this computed mean is close to the distribution mean. We will we will we will make that assumption, that is why it is needed. Yeah? no no if if it is stationary, for being ergodic, it must be stationary because; otherwise if if the if the mean is a function of time and then you are averaging over time, then you will let you will definitely not catch the this thing. But but but all stationary processes need not be ergodic. No no standard deviation.. will be zero; if standard deviation of this will be zero, standard deviation of the random variable is not zero, standard deviation of the estimate standard deviation of the time average minus, the mu is goes to zero as T is, T tends to infinity. So as you are doing a larger and larger average.

So that is all that I wanted to mention new facts. So now so we have done some I mean covered, some basic ground on random processes starting, from I mean I must admit that I have been somewhat fast, but remember that we had started from the bare definition of probability, which is taught in school and then I did not expect anything else. So from there we have come here, so we necessarily had to be fast anyway; we have we will try to make up. However this is to some extent I mean, this is this is necessary to first to.... because first to understand the nature of probability and the nature of some of the result, that we are getting and why we are doing it. Because now in our in the in the course of our discussion, we can freely talk about conditional expectations, we can freely talk about base kinds of estimators.

So that is that is why, it was it was needed and so now in the next part of the course, we we we still have not covered certain things like; you know in estimation in many cases, we are concerned about we are we are we have algorithms which are called records, right. See we are what are with what is what is our central concern.. in estimation, that is we have we have certain signals, which are coming from processes in estimation. And we want to based on those signals, we want to compute other quantities, those quantities can be many things. So in our setting, we generally have a dynamic process, either known or unknown.

(Refer Slide Time: 41:14)



And we have several, signals going in and several signals going out. There are also some there may also be internal signals; so these signals some of them may be may be may be measurable, some of them may be not. So number one, they are we will we will discuss, this in more detail. This is just going to be a five, this is just five minute introduction. So so what is the situation? That you have a dynamic system, you have certain input signals which may or may not be measurable. Even if they are measurable, they may be measurable with certain uncertainty. For example, all sensor signals suffer from some amount of uncertainty and then, based on this signals we ask some question, about what? That could be about different things. Could be will say that, given these signals, can we guess some of this internal signals? That is one problem, which is called state estimation, which will do in great detail. We can ask that given these signals in the past, can I estimate this signal in the future? Can I say that, if I have samples of y, from zero, one, two, three, up to capital N; and if I have the samples of u, from zero, one, two, three, up to uN plus two, can I guess what is going to be y and plus two? See, I am standing at N.

So y and plus two has not yet occurred. I am trying to say that; I have some measurements on the past, and some inputs applied and some some of the outputs, I have measured. And now now I am asking that, if I apply the input at N as as this much and the input a at N and at N plus one this much, what is going to be my output at N plus two? This is the question that I am asking it has not yet occurred. So it is not here occurred means, that obviously y is is not only effected by u, because it.. if it was only effected by u then, you would then there is nothing uncertain about it. If you know the system, and if you know the know u and if you know that, it y is only effected by u, then you have a deterministic scenario. So y is not only effected by u, it is effected by some other random quantities and those you are talking about them in future. So you do not know, what they will be so that is the uncertainty so in the phase of that uncertainty, can you guess, what is what y is going to be? This is a very useful question, to ask in many situations.

For example, I mean several situation for example; if you could estimate the stock market tomorrow, you would become a billionaire. If you if you can heat one air craft with a with missile, that also involve this prediction because, the because the pilot is maneuvering. See the pilot is ... trying to save himself, so he is sometimes sometimes accelerating, sometimes not accelerating, sometimes in the x plane, sometimes in the y plane, sometimes doing s maneuvers various things. So you do not know what he is going to do, but you have to still guide your beside. So can you can you guess, where the plane is going to be over the next two chambers? This is this is the question if you want to do; I do not know whether you have seen; I I was very impressed by a by a a national geographic clip, when it where one scientist showed, I mean they were having a problem with this thing, on active noise cancellation. So active noise cancellation is when you cancel an one noise, by by generating another noise. So that they destructively interfere, see noise is nothing but sound waves. So if two waves constructively interfere, then they then the noise will get.. I mean.. increased, if that if the destructively interfere, the noise will go.

So the question and a noise is a random thing. So he was showing that, he was creating a noise by a by a machine. And here put some you know some thermocol type balls, on on the surface. And because the surface was noisy; so it was so it was vibrating on the balls were jumping and the intensity of noise you can easily visualize, by the by the amount by which the ball was jumping. So then he switched on, his noise canceller and immediately the balls, I mean it came to one tenth of the height and I mean visibly, the the noise signal was gone. Now, how is it done? If you want to cancel noise, then you have to estimate, what is going to be the noise in the next instant and then generate an anti-noise, right. So there are very practical applications, where such things will be needed.

Similarly another another very common thing that common question, that may be asked is that. Given this and this, can you guess what is this this is a? For those who are in control, this is a fundamental question. I mean all control designs starts with a model. If you do not have a model, you cannot design a controller, right. But the but the trouble is that, in many cases when you are... when you are trying to design controllers, you you do not even have a model. Because of the fact especially, for unstable plants, how can you even get this and this? I mean there are there there are some plants, which will not even operate unless, it has a controller. But for a controller you need a model, so I mean how do you get a controller without a model? I mean I mean, this is this was a situation which occurred in our let us say in our LCA design. See LCA is a near craft, which if it has to fly it needs an it needs a very sophisticated controller. But the point is that to to be able to design the controller, you you need to know the model of the aircraft, when it is flying. Do you see the complexity of the situation? So there are there are there are various ways of solving this problem.

And so what people do is that, they have first flow in it in very restrictive conditions. Using in fact in fact they have they have, there are there are aircrafts which can be made to behave like, another air craft. People have companies which with aircrafts for example, a very fast air craft I mean I mean something like an F sixteen can be made to behave like, a slower air craft just by control. So using such such methods, you first get the somehow get the model of LCI in a very restrictive flight condition and somehow construct the controller. Then with that controller, fly the LCI in some restrictive condition. Then you get more data, because to be able to get the model to be able to estimate the model, you need data. You won't get data unless you fly it, okay? So in fact this exercise is still going on, I mean there as they are able to fly the craft then they are getting. Then they are gradually now getting speed, increasing speed, coming almost towards one mark and getting more data. And then refining their models and refining their controllers, so that they fly it in more, so this thing goes on. So so this itself is a cycle, which every control designer has to go through. Especially for in a for stable processor are are are I mean very simple processes; in which you can analytically design a model, this is not a problem, but for most complex plan this is a big problem.

So so we are going to ask such questions, in in the rest of the course and try to get and we are we are also going to be, we are we are we are not only going to ask the questions, that is we are not only looking solutions to these questions, we are also looking good solutions to these questions, not just solutions. So whenever whenever we will we will produce a quantity, some quantity; we will have to examine, how good that quantity is in a given sense. And then may be optimize and and get an algorithm which gives, a the the the best such quantity, okay? And all so this is what we are going to do in the rest of the course, and we will start with input output things. So will basically start with filtering, so statistical signal processing and then gradually move towards state estimation. And then finally go to system identification and some other problems,

thank you very much.