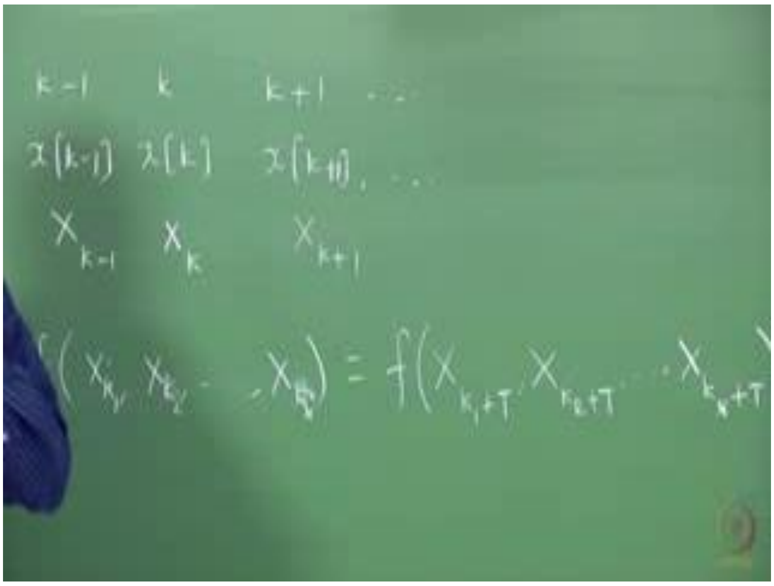
Applied Time-Series Analysis
Prof. Arun K. Tangirala
Department of Chemical Engineering
Indian Institute of Technology, Madras

Lecture - 22
Lecture 10A - Introduction to Random Processes-2

Very good morning to all of you, we will continue with the discussion on stationarity and also draw conclusions on it and today we will also discuss the notion of ergodicity, a bit more in detail, we have talked about it before, but first let us talk about stationarity and if you recall the reason we are looking at this condition of stationarity is practically speaking, the model that we develop from a finite set of observations collected over a period of time should be usable at some other time as well. So, we are requiring that the process satisfy this, of course, in reality processes may not satisfy this then we will have to find remedies for that and will briefly talk about some non stationary processes as well today.

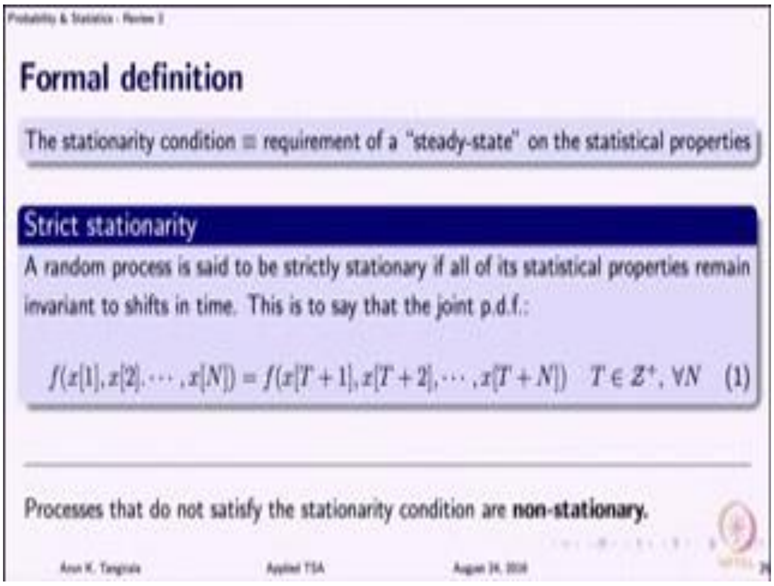
Let us recap what strict stationarity is about, essentially the joint distribution of the joint p d f of this n observations that we have collected, should not change with time and this is regardless of how many observations you have. So, in the strictness sense, every observation should follow out of the same distribution that is what stationarity means there is one point that I want to talk about notation here; of course, we are using x as I have said build shift over to v shortly.

(Refer Slide Time: 01:39)



But apart from that strictly speaking, what is happening is here you have a random signal. So, this is at kth instant, the observation is x k, at k minus 1th instant we have x k minus 1 and at k plus 1 we have x k plus 1 and so on.

(Refer Slide Time: 02:08)



Now what I have written on the slide in the sense the notation that I have used in describeness strict stationarity is with some abuse of notion what I mean by that is the notation x k denotes the observation strictly speaking, we should associate with this a random variable for example, upper case X, subscript k, upper case X subscript x k, k plus 1 and likewise here x subscript k minus 1 and so on.

The x k for us generally denotes the observation or the value that the signal takes at the kth instants and associated with this kth observation is random variable, remember we do distinguish between the random variable and the outcome that the random variable takes therefore, what you see on the slide in the form of definition of strict stationarity should have been strictly written as f of x1 x2 comma up to xn or you can say x k1 k2 whichever way you want to look at it kn should be identical to f of x k1 plus time t comma x subscript k1 k2 plus time t up to x kn plus time.

In other words the period should have been written in terms of the random variables rather than the observations that I have written there, but I will introduces notation a bit later and we will revered back to the notation that we have seeing on this slide just for the sake of convenience. In other words at some point in time you should be able to see x k as an observation x k as also that is this x k as an observation and also as a random variable and sometimes as a random process depending on a context you should be able to interpreted if there is an ambiguity of course, I will clarify the same so that was the part on a notation. Now obviously, any process here coming back to the discussion any process that is not strictly stationary is said to be non stationary pretty much like what we have in a linear world with define what is linearity and we say whatever process or system does not obey which our system does not obey the linearity definition is said to be non-linear.

In reality are all process is there any process that is strictly stationary the answer is most probably no, it is very rare to encounter as situation. In fact, in principle you cannot say that any process is strictly stationarity, strictly stationary for example; you take a dye and generate a random signal by keep keeping on throwing the dye and recording the face value so that is the random signal that you generate let us say I look at this random process now is it strictly stationary, can I say it is strictly stationary? If yes then under what assumptions, what do you think? So, I am just throwing a dye and I am recording the face value, this is the random process, the process of throwing a dye repeatedly, what do you think?

Student: There should be (Refer Time: 05:39).

What does it mean on the dye physically? What does it mean by unbiased?

Student: Probability of (Refer Time: 05:50).

You say, you meant to say under those conditions we can assume it to be strictly stationary, any other answer?

Student: Should be at same condition.

Sorry

Student: (Refer Time: 06:10) conducted on same condition.

Under same condition, what does same condition mean?

Student: (Refer Time: 06:17).

Correct, exactly. So, everything kept the same, the surface should be the same, throwing angle should be the same, the force should be the same, the dye should not undergo deformation, but can we say that it is not really possible. So, I mean we cannot guarantee that the angle would be the same, the surface probably it is the same, but if you look at it a nanoscale, probably the dye would undergo some miniscule deformation; the surface probably changes its morphology and so on. So, as you zoom in and zoom in and zoom in then the process is not stationary. But this kind of idealism is good is necessary so that we know theoretically the rigid frame work in which we are operating.

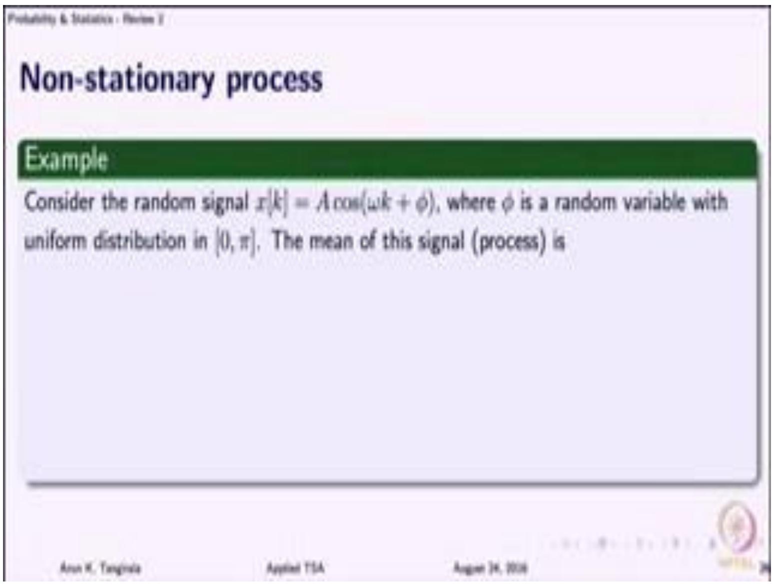
The hope is that small deviation from this will not cause too much harm to the theory that that is the hope and what is small, what is large is kind of subjective in some sense, but there are also statistical ways of saying what is not strictly stationary and what is stationary and so on and we will learn that as we move along. So, there is theory and then there is practice and one has to be able to appreciate both the theory and the practicing part of it where we will not satisfy the theory, but it is in the sense, there is a there are always going to be practical aspects.

Let us look at a simple example of a non stationary process, before we move along, we will talk about will return to the stationarity part, but just to get a feel of what is a non stationary process, we have looked at the signal before. If you recall when we were computing expectations, we are taken up this example, but there is a slight difference between the conditions in that example.

(Refer Slide Time: 08:14)

Now the signal is given on the slide x k, it is a periodic signal, x k is a cosine omega k plus phi a is amplitude, omega is a frequency, it is this is the discrete time s a cosine and phi is the phase or the phase shift and the randomness in x k comes about due to the randomness in the phase phi which is uniformly distributed in the interval 0 pi.

(Refer Slide Time: 08:47)



Now, what you think is the mean of the signal can you calculate the expected value of x k the purpose of calculating this expectation is to see even if the at least if the expectation is invariant to it time, if it is in then we can straight away conclude that the process is non stationary generally it is very hard to check for non stationarity, but we can perform some basic calculations such as on mean come computing the mean variance and so on and if we somehow chance upon finding that the mean or the variance or later on auto covariance that we introduce will change with time then we can straight away say that the processes is non stationary. So, as a first strep let us compute the expectation of x k we have already gone through this calculation only difference is now the interval over which the phi is distributed I will give a minute to calculate the expectation go ahead.

It given right that phi is uniformly distributed, anybody with an answer from the other hall, anyways any difficulty in computing expectation.

Student: (Refer Time: 10:28).

Sorry.

Student: (Refer Time: 10:30).

Sure.

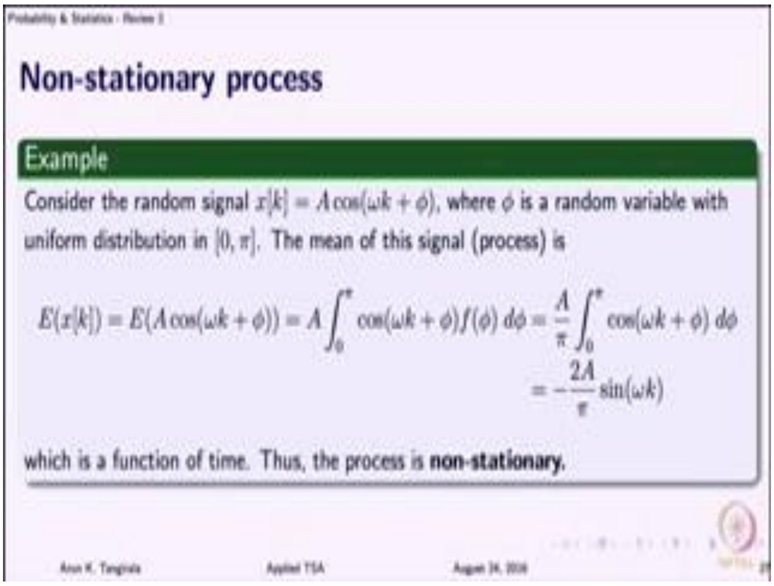
Student: (Refer Time: 10:34).

Already now 4 answers have come out. So, which one is the correct one? Shall we go by majority? Please check your answer once again and give me the correct one, any answer from 241, you have, you are ok? You have some difficulty computing the expectation? You do not want to expect anything done; you have the answer, what? Say answer.

Student: minus 2a by pi sin omega k.

Minus 2a by pi sin omega k, anybody with the same answers, good. So, we will go by the Lokh Sabha route of course, this is math; you do not need Lokh Sabha route. So, that is the answer, right, yes correct. So, minus 2a by pi sin omega k, very good.

(Refer Slide Time: 12:02)



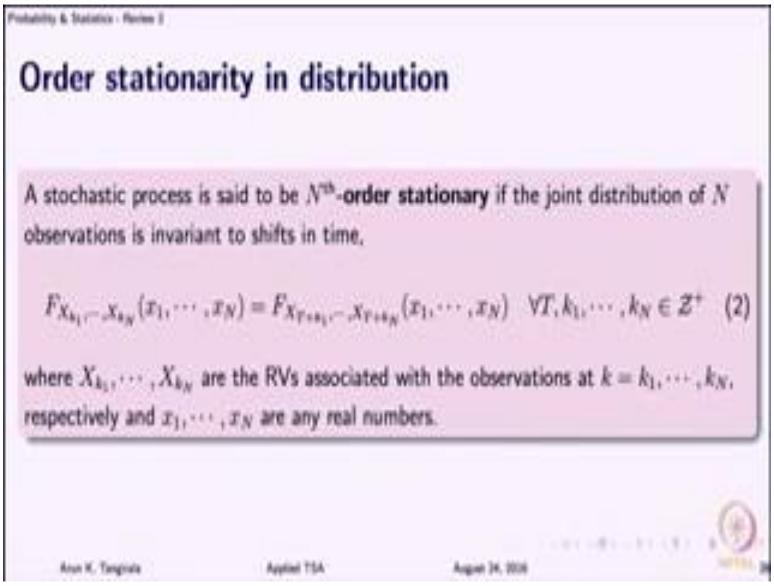
Now; obviously, here the mean is a function of time correct; that means, at any time k has that is as I move across time when I say time here the sampling instant k the mean is going to change and that is the sin of non stationarity. So, straight away we can conclude that this process or this signal is non stationary. In the example previously that we had looked at when we were computing expectations; we had a slightly different interval, if you remember, we had interval between minus pi and pi, am I right?

There of course, the mean turned out to be 0, but that does not mean that the process is stationary, says that in other words if the phi, the phase here was distributed uniformly in interval minus pi to pi, we would get the mean to be 0 which is invariant with time that does not mean that the process is stationary, it only means that it is stationary in mean, we say it is mean stationary, the variants could be changing with time or any other movement could be changing with time. So, as you can see, it is not easy in practice to verify non stationarity, even by hand it is difficult and from data it is going to be even more difficult why? Because the data that we have will be only a single realization, at least by hand we are able to compute across all realizations, but when it comes to verifying stationarity in practice, it is going to be even more difficult because we will be working only with the single realization, is that clear.

Therefore when you look at the literature on tests for non stationarity, there is no test out there which will universally test for all kinds of non stationarity, one has to specify a priory what non stationarity we hypo, one is hypothesizing and then conduct the test for that we need to know what are the different kinds of non stationarity is that exist, but will come to that shortly, but this is something to keep in mind when you are looking out for test for non stationarity you should be first clear what is the null hypothesis in such a test whether its assuming there is a linear trend whether its assuming the non stationarity is of integrating type like random work type or quadratic trend or a mix of both. So, there are few different tests be very clear on what the test is actually testing for before you use the non stationarity, it has many applications will pro generate non stationary random signals. Any questions on this example? Fine.

Now what we do is we take this strict stationarity and come up with a slightly is not exactly diluted, but slightly different version of stationarity, it is a slightly definitely a relaxed version of the strict stationarity that we had seen and we will see even more relaxed versions this is not the one that that we are going to work with, but for completeness sake let us understand this kind of relaxation or dilution of the strict stationarity definition.

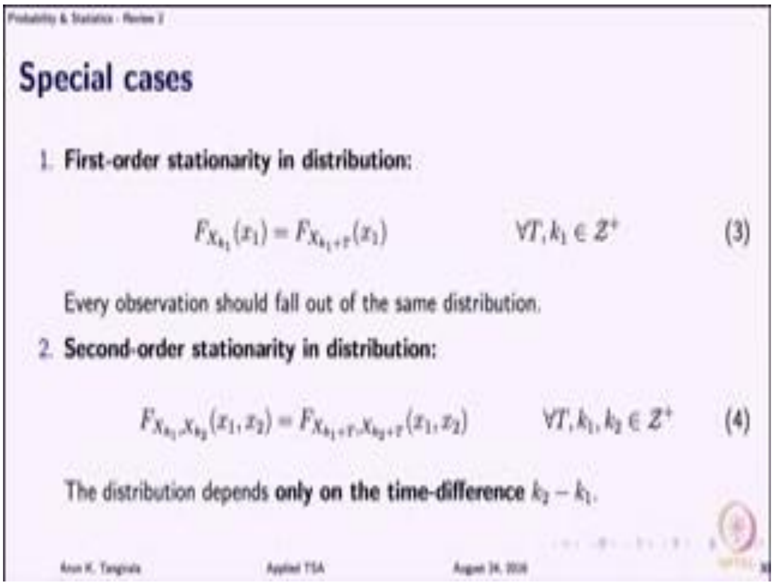
(Refer Slide Time: 15:30)



This is a this stationarity is called order stationarity in distribution very soon we will talk about order stationarity in movements where we will talk about stationarity in the second movement stationarity in the first movement and so on, but this is order stationarity in distribution, what do we mean by order here? It is a number of observations that you are looking at jointly essentially you apply the same definition of strict stationarity, but not for any n, particularly for only for a specified n, if you compare with the definition, we had here, the strict stationarity demands that the distribution whole for all sets of observations that all sample sizes that you are at whereas, order stationarity is looking at specified user specified value of n.

When the distributions are identical, earlier you saw the statement in terms of p d fs, now you are looking at statement in terms of c d f, that is only difference, why do we bring in c d f so as to be so as to handle discrete valued random processes as well, p d fs are fine or valid for continuous valid random processes, but c d fs are valid for both therefore, it is a slightly general version and you see the notation is now more appropriate, I would say and it uses this notation that I have talked about earlier. So, the random process is said to be nth order stationary in distribution when the distribution of this n observations that you have collectively looking at remains invariant with time, it is a very simple definition. As special cases one can look at first order stationarity.

(Refer Slide Time: 17:20)



That means I am looking at one observation at a time, I am analyzing single observation and in this case what we are saying is when I pick a single observation, the distribution should not change with time and when I look at second order stationarity in distribution then I am looking at the joint distribution of 2 observation and that should remain invariant with time and you can go on, you can say that the joint distribution of three observations should remain invariant with time and so on.

As this strict stationarity consumes or subsumes all order stationarity remember, it says that it should be stationary for all sample sizes again these are all idealizations that are important, but there is a particular point about the second order stationarity in distribution when we require that the joint p d f of 2 observations remain invariant with time. It also implies that the distribution is only a function of the time difference that you are looking at rather than the time instant because it should be independent of time. So, the only way this distribution can be different is when you are looking at for example, observations x1 and x2 and x1 and x3; obviously, this statement does not require that the joint distribution of f of x1 x2 should be equal to f of x1 comma x3, this is not necessary, what it says is instead this can be equal to this should be equal to what some other f of x 5 comma x6 or x4, does not matter.

(Refer Slide Time: 18:54)

2 successive observations should have the same joint p d f, it does not matter where those 2 successive observations are located in time it could be, now it could be later or it could be even more later, does not matter any 2 successive observations I take or any 2 observations that are separated in time by k2 minus k1 should have the same joint p d f, this is a stronger requirement than the first order stationarity, is it clear?

This is important to understand because very soon we will talk about what is known as a second order stationarity in moments or wide sense stationarity or weak stationarity, there are different names to the same thing which will have this kind of a property, but in moments only here, we are saying distribution should remain invariant with time as long as the observations are separated by certain time, what we time difference or what we call as lag.

When we move on to wide sense stationarity, we will require that the moments remain invariant with time and they be only a function of the time difference, we know that moments are subsets, they there being generated from the distribution so distribution is the Pitamaha, moments are all children of that of the distribution. So, here we are saying the parents should be invariant.

So, even to verify, to satisfy the distributions, to remain invariant with time like this, here is quite difficult, it is impact; it is quite unrealistic to expect. Now again we ask here this question whether this is necessary, remember we asked a similar question in predictions what was the question that we raised at that instant when we talked about predictions I said conditional expectation is the best prediction, but then we realize that conditional expectations are hard to compute because they require the knowledge of joint p d fs, correct and then we asked well conditional expectation is a non-linear function of the predictor or the regressor, do I need such a non-linear function? Can I work with linear functions and then we had some discussions saying yeah in practice I would I would like to test the use of linear predictors if they do a good job then it is great, correct if they do not then will worry about non-linear predictors and so on likewise here may be I do not need this strict stationarity.

Even I do not need perhaps the second order stationarity; forget about strict stationarity, I probably do not need the second order stationarity in distribution. I can focus rather on moments and that is what the relaxation that generally one looks at. The questions is what is that relaxation exactly and when is that relaxation justify, when can I work with the relaxed sense of stationarity?

All of this is to make sure remember, every time you have to ask yourself why on earth we are having this discussion because we want to make sure that the models that we work with or the inferences that we draw, need not be model alone, it can be features, it can be whatever classifiers that you have developed, on a single realization should work for future as well and should work in the past as well that is the reason why we are talking all of this.