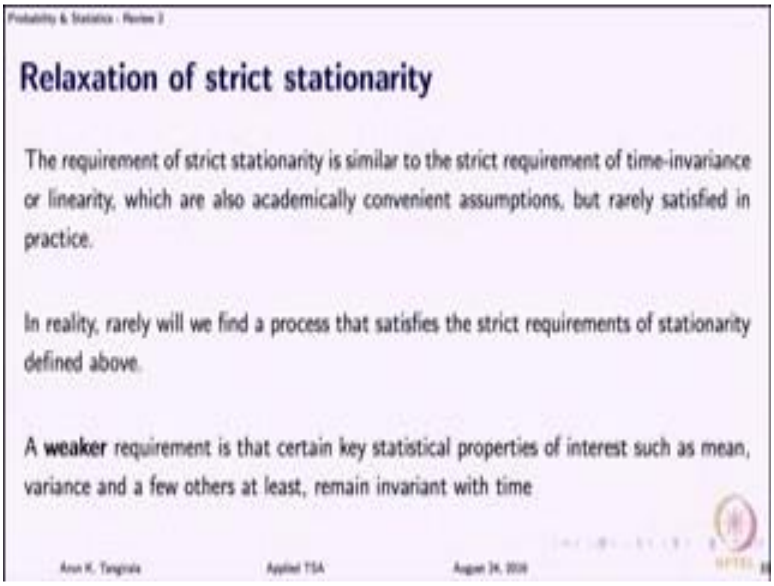
Applied Time-Series Analysis
Prof. Arun K. Tangirala
Department of Chemical Engineering
Indian Institute of Technology, Madras

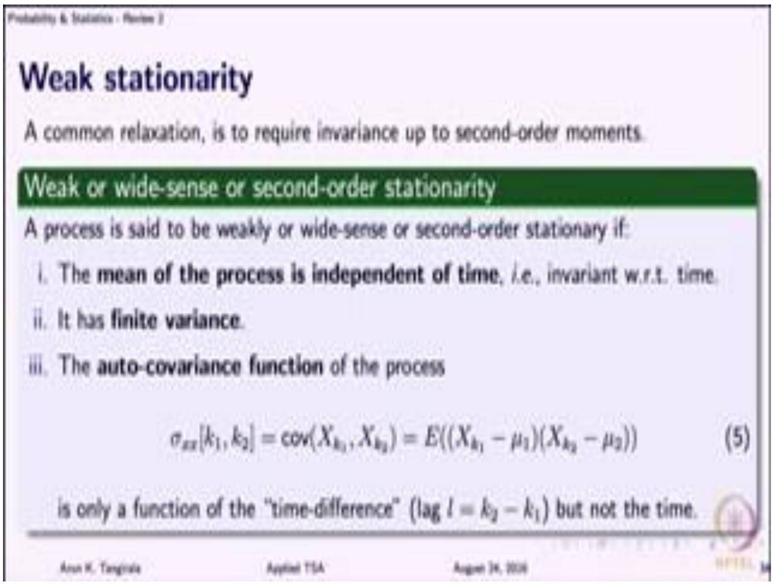
Lecture – 23
Lecture 10B - Introduction to Random Processes-3

(Refer Slide Time: 00:17)



So, generally the weaker versions of stationarity are defined by requiring that certain key statistical properties remain invariant with time and we know by now that the some of the key statistical properties are mean, variance, but what we do not know is something called auto covariance. We have talked about mean, we have talked about covariance that is because we were talking of general random variables, but now we are talking of random signals where we are looking at an ordered collection of random variables and it is important for us to look at these jointly and that is where the auto variance notion comes into picture and leads us to this notion of weak stationarity.

(Refer Slide Time: 00:56)



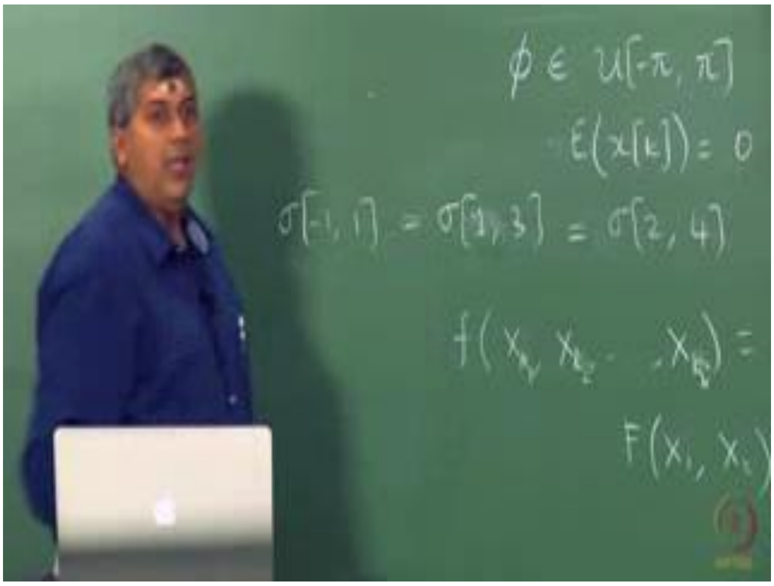
So, this weak stationarity has different names it is called white sense stationary, stationarity or second order stationarity; as you see from the definition why it is called second order stationarity. This order here is not in terms of distribution it is in terms of moments we are looking at second order moments. So, there are three requirements for a process to be called as weakly stationary or white sense stationary or in a second order stationary the first requirement is that the mean be independent of time.

(Refer Slide Time: 01:47)

So, the in the example that we just discuss we had mu k as minus 2 a by pi omega correct sin omega k correct good.

Is this weakly stationary it does not satisfy the first condition which is the mean to be independent of time on the other hand. So, this was a case when phi was uniformly distributed between 0 and pi, if on the other hand I have to remember my t s advice hold on.

(Refer Slide Time: 02:23)



So, if on the other hand phi for the same problem phi was uniformly distributed in minus pi to pi, we have already calculated the mean have not we, what is the expectation 0 right. So, expectation of x k is 0, at least it satisfies the first condition. It does not mean it is weakly station, there are two other requirements that it has to satisfy correct; then the second requirement is at the process to have finite variance.

So; obviously, we are excluding class of processes that can have infinite variance are their real life processes that do not satisfy the second requirement, yes there are many processes that have so called Cauchy distribution like an earth quake process for example, has infinite variance it does not really fall into this realm of weakly stationary processes, but we will not worry about that, we will fortunately we have a large class of processes that still satisfy the second requirement. You must appreciate that these two requirements are not as stringent as the requirements of street stationarity correct, the street stationarity or even the nth order stationarity in distribution requires at the distributions remain in variant, but here we are only saying let the averages and the variances remain invariant with time and let them we finite.

Of course in the first we have implicitly also meant that the mean should be bounded as well I have not stated that, but you can note that down; mean should also be finite. If the mean is infinite what happens; is a second condition satisfy not. So, it is kind of understood, so then the third requirement where now for the first time we come across this notion of auto covariance function. If you look at this definition, we will go over this definition in much greater detail bit later, but up front in order to give a complete definition of weak stationarity; I am introducing the definition of auto covariance here itself it is like a curtain riser, but we will look at this in detail later on.

So what is this auto covariance, why is it called auto covariance? Because you are looking; first of all it is a covariance then as a name suggest auto covariance function is a covariance and now you understand why we studied covariance in great detail correct. It is a covariance between any 2 observations that I randomly choose; located at instants k 1 and k 2 alright and this auto covariance in general for any process can be a function of where these x k 1 and x 2 are located where these 2 observations are located and that is why we have written the covariance as a function of k 1 and k 2 and a subscript, there is a double subscript indicating that you are looking at observations from within the same signal; which means there is a possibility of looking at covariance; computing covariance between two observations from two different random processes.

That is where we run into cross covariance function, but will discuss that later on. So, the notation has to be carefully understood. So, the third requirement is this auto covariance function be only a function of the time difference; pretty much like what we demanded for the nth order or the second order stationarity in distribution. Why do you think this is necessary, why are we demanding that this auto covariance function be only a function of the time difference and not k 1 and k 2; any explanation; you can think about it its fairly easy the same kind of arguments that we used for requiring stationarity.

Student: (Refer Time: 06:47).

Right, what kind of relationship?

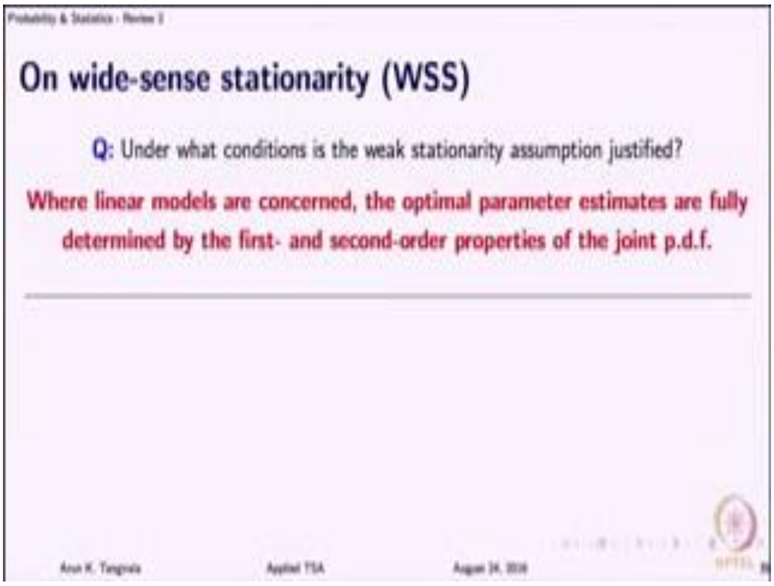
Student: (Refer Time: 06:55).

Very good, so why is that required? Why is it required? Practicality, it is shear practicality right. I mean; I collect a bunch of observations that is one data record and I am going to examine the dependency ultimately what I am going to do I am going to use his model for forecasting. So, if my statistical measures are telling me there is a dependency of the past the present with 2 samples in a past let us say then I would like that to remain invariant that dependency to remain invariant with time, what happens if it is not that model that I have identified is only for that combination of observations, I cannot use that model in future. So, each time we are going through the same argument ultimately we want to make sure that the model that I build that the analysis; the inferences that I draw from the analysis of single record whole good for future and probably for the past as well right as definitely.

So, here we have restricted our self to linear relationships straight away because we know covariance is a measure of linear dependence, straight away it gives you a feel that this definition is sufficient as long as you are in the linear world or this is definition this relaxation of strict stationarity is good is find and a sufficient when your developing linear models alright. So, the white sense stationarity therefore, has three requirements mean should be independent of time, variance should be bounded and auto covariance should only be a function of this so called lag, what is the lag; the time difference. This also means that - if I am looking for example, at covariance between let us say now x 1 let me say here 1 comma 3; if it is white sense stationary or weak weakly stationary they should be equal to sigma 2 comma 4 and this can also be equal to sigma minus 1 comma 1 and so on.

Any two observations that are spaced l samples a part should have the same linear dependence, what is this l? k 2 minus k 1 and that is what we call as the lag. So, this auto covariance function is not a mysterious one, it is after all at it is heart a covariance measure that is all, it is auto because a signal you are looking at the same signal, it becomes cross when you do not claim that auto you actually get into two different your analyzing two different series correct any questions. So, these are the three requirements for a process to be called widely sorry wide sense stationary or weakly stationary and also second order stationarity why do we call this a second order stationary because the covariance is a second order movement of the joint period correct.

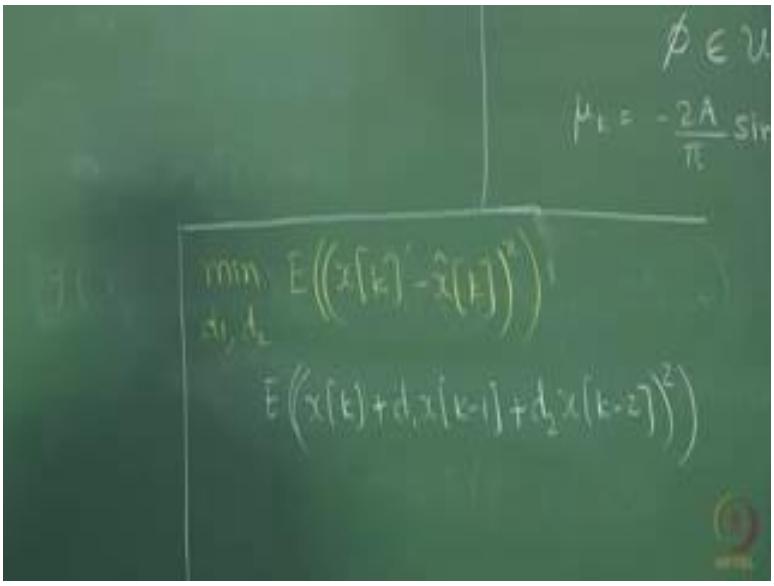
(Refer Slide Time: 10:25)



So, under what conditions is this weak stationarity assumption justify when we are living in the linear world. How do we justify this? Of course, we have already seen covariance is a linear measure and we know that in correlation is one of course, we have talked only about covariance very soon, we will talk about auto correlation which is nothing, but a standardized measure, we have seen the relationship between linear regression and correlation they are synonymous to each other. So, intuitively it is very straight forward to see that this wide sense stationarity is justified in the linear value, but let us understand is with the simple example. So, suppose I have a stationary process some station; I am given it is stationary and I am going to bill a predictor of the form that I have given; that is I am going to predict the random process at k using a linear function of 2 immediate observations in the past and let us say the coefficients are minus d 1 and minus d 2, there is a reason why I have used minus d 1; you can use d 1 as one does not matter.

Now, what I want from you is the answers to the optimal estimates of d 1 and d 2 and predicting a stationary process using this linear predictor and I would like you to give me the optimal estimates; optimal in what sense that minimizes the mean square prediction are, we have already seen that.

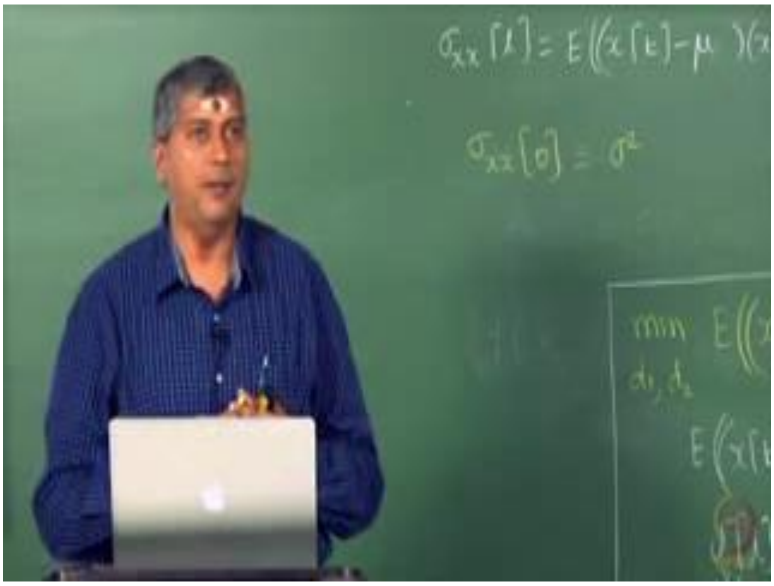
(Refer Slide Time: 12:10)



In other words you have to tell me what are d 1 star and d 2 star that minimize expectation of x k minus x hat of k x square and the answers could be given in terms of the statistical properties. You will realize the purpose of this exercise once you get the answers and then it is kind of a corroboration of what we have answered the top. Assume x to be 0 mean for convenience, remember that since there are two unknowns, you will have to come up with two equations right at least 2 equations and it is even sufficient if you tell me those 2 equations, the hint is that you will have to use the definition of auto covariance that you have introduced earlier.

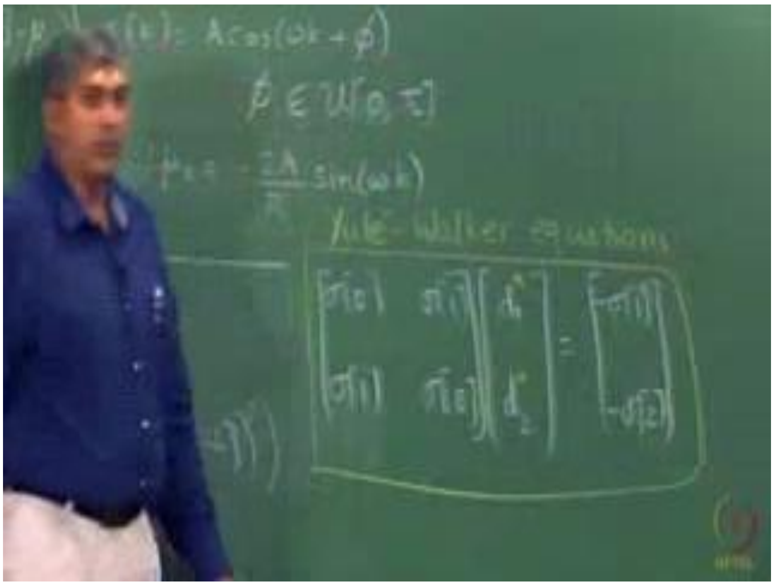
So, that when you are looking at expectations of products of two different observations, you will write them in terms of auto covariances; you have those equations sure.

(Refer Slide Time: 13:43)



Or I can even write on the board for you sigma x x at any lag l is expectation x k minus m k; x k minus l minus mu k minus l. Since I am given it is stationary, we can actually assume that mu is the same and that the auto covariance is only a function of the lag l and given there it is stationary; it is very important and you are additionally given that mu is 0. So, for this example sigma x x at l is nothing but expectation of x k times x k minus l, any answers from here or the other hall. Let me ask you let me write here matrix remember will we have two equations and two unknowns.

(Refer Slide Time: 14:40)



So, two unknowns are d 1 and d 2 ultimately, you should be able to write your equations in this form as a set of 2 equations in a matrix form and then eventually it boils down to a linear algebra problem. So, you will have to tell me what are the entries of this matrix and this vector here, if you proceed in a step wise fashion then you will not get lost; first expand this quadratic and then apply the expectations and rewrite your objective function in terms of the parameters and auto covariances and then differentiate the objective function with respect to the parameters d one and d 2 any answers the expectation of x square k would be the variance itself any answers from the other hall; yes.

Student: (Refer Time: 16:09).

Sorry.

Student: (Refer Time: 16:13).

Here.

Student: (Refer Time: 16:14).

Sigma let us draw the subscripts, so that is easy sigma here sigma of 1 alright minus sigma 1 alright and then.

Student: (Refer Time: 16:36).

So let us see if there is at least we need one more person to confirm this answer. Now what is sigma naught square hold down, so that is probably got to do with the confusion in the notation; sigma x x and 0 since I have not a written in the subscript perhaps is a confusion sigma x x of 0 is nothing but sigma square. So, either you specify the lag or you drop the lag and then you write it is in the sigma square. So, that is the convention; this confusion will prevail for one or two lectures and after that you will be comfortable, so either you use sigma square or sigma at (Refer Time: 17:26) clear.

Any other person, with the same answer; is it correct is fine very good. So, others of course, can verify you can always discuss with me after the class, but this set of equations are correct. So, you see we have actually moved from the world of random processes to the world of linear algebra and that is always going to be the case. Now we will as you will learn throughout the course keep visiting these kinds of equations for model building, what you see straight away is the optimal estimates. So, if you call them as d 1 star and d 2 star, the optimal estimates are actually only dependent on the second order movements.

So, it is if I want my model to remain invariant with time or let me put it with this way; if the auto covariance is are invariant with time then a model parameters are going to be invariant with time which means that the model that I estimate from a finite record of data is potentially useful for future as well; that means, some kind of stationarity is obey. So this hopefully drives home the point that we made earlier that the weak stationarity or white sense stationarity is good as long as we are in the linear value. The moment we move to the non-linear regime, then other movements come in that is a higher order movements come in; what kind of higher order movements set in depends on the nature of the non-linearity.

We are not in a position to comment on that because as a said it depends on the nature of the non stationarity and non-linearites as well. So the point to remember is; it is sufficient to assume a process to be weakly stationary as long as I am working with linear models. By the way the bunch of equations that you see here where conceived by 2 different people Yule and Walker and this equations that you have are named after them, somewhere in mid 1920s these are called Yule Walker equations and they are very frequently used to estimate what are known as auto regressive models. In fact, the predictor that you see here in this example comes out from what is known as an auto regressive model of order 2, but we will discuss that in detail later on.

So, write down the purpose of this example was to show by the sufficient to have invariance of second order movements; up to second order movements as long as we are working with linear models good. Now again like we saw in the case of prediction we said a linear predictor is in general sub optimal predictor compared to the conditional expectation, but their identical when x and y are jointly Gaussian distributed; likewise here white sense stationary process or a weakly stationary process is also strictly stationary, if the observations or if the random process itself is so called Gaussian process; what is a Gaussian process the joint distribution of any set of observations that you pick have a joint Gaussian distribution; does not matter how many you pick, you pick 2 then there should have a joint Gaussian; you pick three observations jointly there should have a joint Gaussian and so on.

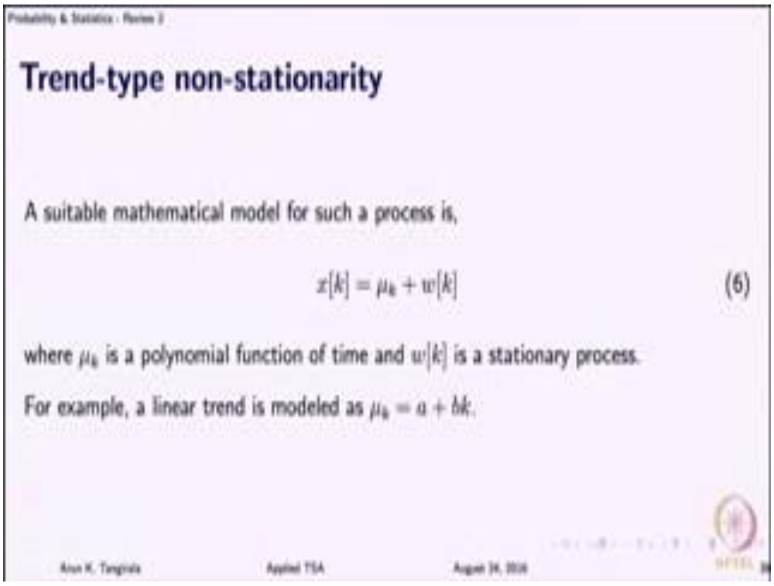
Which means every observations should follow out the Gaussian distribution as well why is this so, why Gaussian process which is weakly stationary strictly stationary process as well because the Gaussian distribution is completely specified why the mean and covariance and bit and the whites and stationarity demands that the mean and covariance remain in variant with time and; that means, the distribution itself will remain in variant with time therefore, you have the result it. So, Gaussianity has a lot of nice properties and that is why that is the de fact to assumption, default assumption that people make; if they have to complete the degrees and if you do not want to get in to complications alright.

So, very quickly let us talk about non stationarities and will continue with discussion tomorrow, we have talked about street stationarity, we have talked about order stationarity in distribution and we have also talked about second order stationarity in movements. So, it definitely it is worth and there is a merit to discussing the different at least some of the most popular types of non stationarities that one encounters; it is impossible to discuss all types of non stationarities. Generally speaking; that is broadly speaking, you can classify the non stationarities into 2 types deterministic type of non stationarities and stochastic type of non stationarities. In the deterministic category you would run into non stationarities such as polynomial trends like we saw for the air line passenger data, there was a linear trend; it could be quadratic also within test, but definitely there was a trend.

What is it mean? The stochastic signal is riding on a horse that is a deterministic function of time, it is a polynomial function of time or it could be any other deterministic function of time, it could be for example, may be a sign wave as well, but then those are special cases, you could also that is what we call as periodic, you could also have variants non stationarities and variances being deterministic functions of time, it may not be just the signal being signal having a trend. On the other hand in the stochastic type of non stationarities that trends that we spoke of just now could be a function of stochastic function or could be the outcome of a stochastic process that is leading to the non stationarity and we shall discuss only one class of this stochastic type of non stationarities known as a random walk processes or the integrating processes.

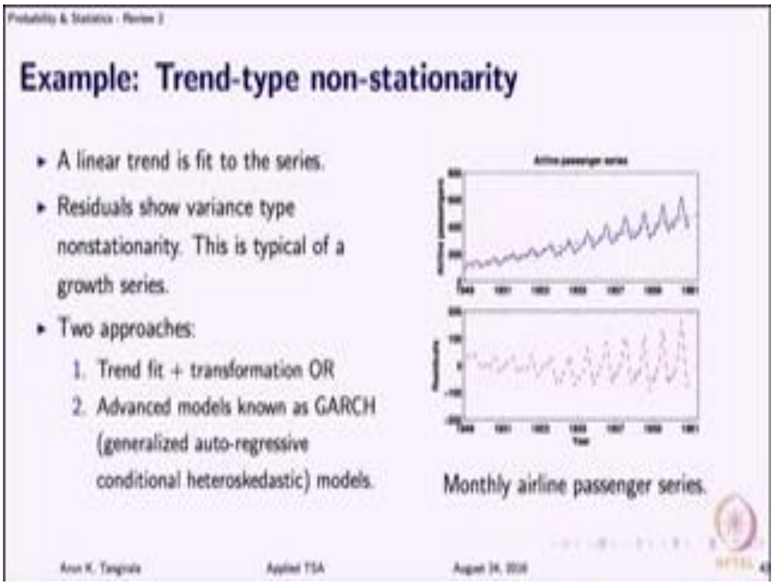
There are of course, many more types of stochastic processes, but large class of the stochastic processes that we encounter particular in economic rates and even in engineering can be model as random walk processes and we will also see that this random walk model can handled to a certain extent the deterministic trend type of non stationarities. So, to give a very quick example of a trend type non stationarity and now I have come back to the deterministic type of non stationarities.

(Refer Slide Time: 24:30)



For example you could have x which is your random signal as super imposition or super position of two signals; one is a deterministic function of time and the other is a stochastic process. So, mu k is the polynomial function of time, if mu k is a linear function then we say it is a linear trend.

(Refer Slide Time: 24:51)



Let me give you an example; the air line passenger data that you see here on the top we have also look at this yesterday, if you see there is a trend correct and on the phase of it; it appears itself a linear trends you can probably say it is a linear trend type non stationarity, but one has to only confirm by fitting trends and seeing if a linear trend is a able to explain the trend or a quadratic is necessary what I have done is I have fit a linear trend and now I am showing you the residual whatever is left over after fitting the trend; what you see is it still stationary it is not because the variance is changing with time at least if you your we have handle the mean non stationarity that we have captured with the variance non stationarity remains and this is typical of any growth process.

That has to be handle and there are 2 different approaches to handle that either you can then a perform a transformation to convert this non stationarity process to a stationary signal or you can fit what are known as gorge models generalized auto regressive conditional hydro stochastic models too much, but I mean if you do not know anything you can use it to scroll summer or that is what we use do to with terms and biology like protozoa and so on when we (Refer Time: 26:10) I had very difficulty great difficulty understanding having a feel for platyhelminthes and so on, but now probably I will appreciated by that anyway. So, we will not talk of gorge models in this course I am just saying that there are different approaches to handling non stationary processes.

We will continue our discussion tomorrow where we will conclude with ergodicity and then go into auto covariance.

Thank you.