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

Lecture – 03
Lecture 02A - Motivation and Overview 3

Let us get started. So, yesterday we had a brief discussion on what is time series and probably what is time series analysis about. Today I am going to talk about what are the specific challenges, again today you will not see any equations per say that is a typical strategy to make sure that to make you believe that time series can be learned without equations. But that is not the case gradually will spruce up the number of equations that come up and so on, anyway not to scare you first, it is important to gain a conceptual understanding of what the subject is about and then that will also give you an opportunity to decide for instance whether this course is for you and. Secondly, maybe it will give you some ideas which you have never come across so that is the objective with which we are going over an equation less and more of a conceptual understanding of time series analysis.

What we will now specifically focus on is what are the challenges in time series analysis and that is what is the thing that we need to know because then it will tell us as to what kind of complexities we will have to face and what are the problems that we will encounter. So, the key challenges in time series analysis are randomness and uncertainties these are the 2 things that makes a subject very special very challenging and it has continued to intrigue researchers for several decades, you can if you look at a history of time series analysis in the open literature dates back to at least 1920s and then there was a flurry of activity a lot of momentum somewhere in the 30s-40s and so on and what we are learning today is essentially the essence of whatever has been worked out somewhere from 30s to 60s with of course, a lot of other embellishments and developments in the subsequent years.

These 2 are actually divine because they help us carry on with our jobs, they help you know get jobs we because imagine if randomness or uncertainty was not there in the world then a single person would have been capable or a few people would have managed to predict everything and then the rest would be jobless there would be nothing

to be excited about and so on. The good thing about this randomness and uncertainty is that nobody knows the truth which is a very nice thing all of us are actually trying to inch closer to the truth, we believe that we are somewhere around the truth and then somebody comes along and says no you are millions of kilometers away from the truth and that saga continues. So, in that sense it is divine and that is why I say every morning you should worship uncertainty and randomness because without which we would not be carrying on with our lives; of course it has its own cons, but we will not worry about that.

Why are these 2 things making the time series analysis very challenging first of all what is meant by randomness? All of us have some misconception or some conception about this term randomness right, let me actually take a sample on what thoughts you have when you hear the term randomness maybe randomly pick one student here maybe somebody in the other hall wants to answer also he or she is welcome. So, let us ask around at least a couple of students as to what they he or she thinks about randomness anyone.

Student: Which is not (Refer Time: 03:54)

You should raise your hand I said in yeah, sorry.

Student: Which is not (Refer Time: 03:58)

Like what we went today, he says yeah my goodness. So, much planning you can still have uncertainty so; obviously, I mean it is not that we have not tested, we have tested many a times, but still uncertainty can come in and say I am the boss. So, it is got not to do with planning anything else

Student: There is no order.

There is no order.

Student: Something after a b c in canada.

Okay, that you cannot really say d will come, it could be z also, I am sorry.

Student: I cannot predict it.

I am sorry

Student: It cannot be predicted.

You cannot, what is that it? It cannot be predicted would mean what?

Student: Actually what I mean to say it cannot be predicted 100 percent suppose all that

What is that it?

Student: (Refer Time: 04:54).

Are you referring to it as a part of IIT, it. What is that? I know what you mean, I just want you to be more clear, that is all.

Student: (Refer Time: 05:08)

Sorry yeah.

Student: The variable, the like or balance variable (Refer Time: 05:19).

Anybody from that hall, who wants to answer on what is randomness because this is a key to understanding randomness, is a key to enter the world of time series analysis.

Student: You do not know what the result is going to be before you do it (Refer Time: 05:43).

Not bad, any other answer from that hall.

Student: You would except that range of values instead of a single one and these values may be distributed within a certain theory.

We are already into the world of approaches, mean you are probably a step ahead in terms of explaining how to deal with randomness, we are not asking how to deal with randomness, we just want to know what is randomness, what do we mean for example, by a random variable or a random signal these are the terms that we keep hearing.

Somehow we have some notions about random signal for example, a few thing that random signal is not predictable and so on, so, one final answer from any of the halls.

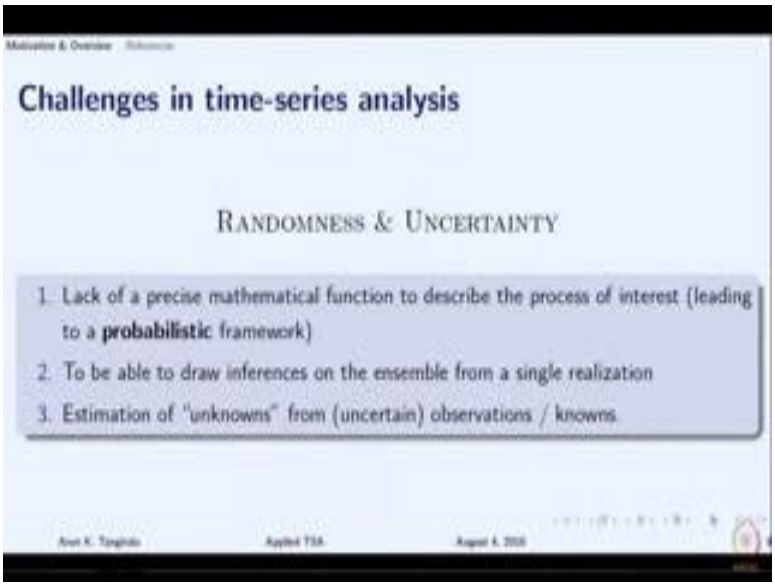
Student: (Refer Time: 06:40).

Sorry, may or may not happen; now that is the standard misconception. So, thank you, now all of you have given some inputs which are partly right, it is not fully right, does not fully explain what is a randomness, there are several ways in which one can describe randomness, but definitely randomness when we talk of a random process or a random signal it does not mean that a signal may exist may not exist or a process may produce an outcome or not it is got nothing to do with that. So, the randomness is not about the event itself event per say, but the randomness is about the outcome of the event. So, one has to be able to distinguish between the outcome and the process itself a random process by definition exist forever it always exists because if it ceases to exist there is going to be a contradiction which we will realize later on.

A random process exists forever its outcome in the simplest term if you want to say is unpredictable, I mean that is a vanilla definition, we will actually refine that definition as we go along, but from a prediction viewpoint and it is useful now to start developing this perspective of prediction from a prediction viewpoint a random process is one or a random signal is one whose outcome you can say or a value that the signal takes at the next instant is cannot be accurately predicted.

You have to notice that I said cannot be accurately predicted I never said cannot be predicted at all. So, there is a difference between accurate prediction and just not be able to predict at all as you will see later on when we get into the math of time series analysis on the status of it, you will realize that there is a huge difference between these 2 situations.

(Refer Slide Time: 08:38)



Now, from another perspective, randomness actually means that there is no precise mathematical function to describe that process, some of you have actually pointed that out which is correct I cannot for example, somebody said well if I am given that a b c has occurred then I cannot say accurately the d is going to follow a b c something else can come up.

This is a hallmark feature of any random process it our inability it basically makes us it puts us in a weak spot we are unable to actually predict accurately, it does not mean that we cannot predict at all for example, if you have seen it is raining, if you know the rainfall pattern in a particular city then if I ask you whether it is going to rain or if you have seen the temperatures dropping, the air is cooling down, clouds are in the sky and if I ask you whether it is going to rain then; obviously, you can say no, I do not know it is a random process, it may or may not rain that cannot be the stands yes that is true, while it is true that it may or may not rain. As practitioners we cannot take that stance will you do that in reality you will say I do not know if it is going to rain or not despite I see clouds black cloud forming in the sky the air is cooling down and the temperatures have dropped we will not take the stance that no, it may or may not rain. I mean we are not running into a lagoon like situation where everything comes and then the song is over and the clouds go away, it is not like that.

It is you can predict, but there is going to be an answers uncertainty associated with the prediction, even there we are not 100 percent confident that it will rain, it can happen that despite all these conducive situations for rainfall that the clouds will go away, it would not rain, it can happen. So, it is that other thing that can happen makes the process random we cannot say accurately on the other hand if I say that I hold this and I leave right how many of us believe that it would not fall up and we will rather fly upward a fall down and fly upwards anyone in this hall or the other hall believes that it will fly upwards anybody you probably do not know I am a magician.

Maybe I have a button here which will actually let it go up, if I release it, what is your prediction? You will fall how sure are you mostly somewhere you want to give yourself a chance correct so; obviously, if I release it, it has fallen. So, you are right.

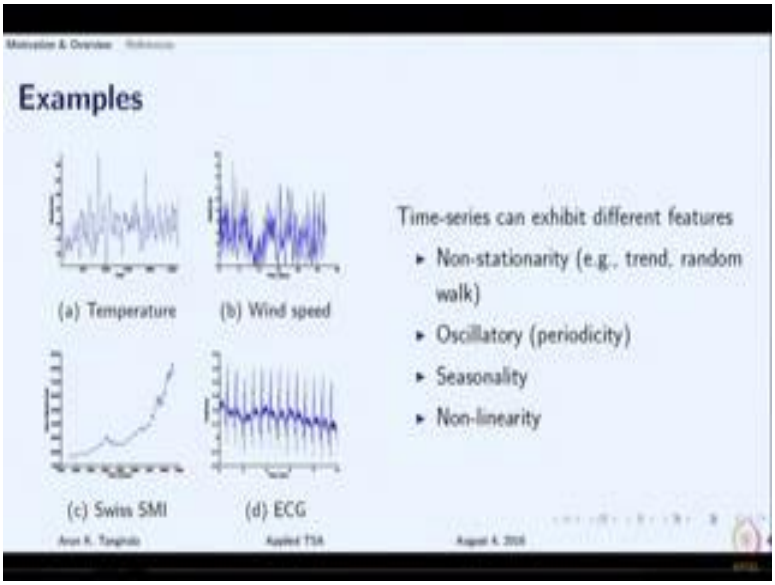
So that you do not want to call a random process because you believe that there is a more chance that it can flow, it can fly upwards, we call that as a deterministic process. In fact, a better technical term for that is predominantly determinesed that is yes there is somewhere you know somebody said mostly yes somewhere there is some uncertainty troubling our minds, but a pretty sure that is how an engineer should be in the sense that is the stands certain engineer should have to be practical. There is no ideal deterministic process that is for sure, a deterministic process is one which can be accurately predicted whose course can be out accurately predicted there is no it is only an idealism, but that idealism helps us in defining what are non deterministic processes. There are different terms that you will encounter in the literature non deterministic random stochastic and so on, there are of course, some subtle differences between randomness stochastic we will not observe those differences today or even in the course.

The point is that in general, any process will be characterized by some uncertainties; the question is to what extent. Now all of these discussions that we are having and more were actually a part of the initial developments in time series analysis many researchers really spend enormous time in trying to come up with a definition of random process, we will come up with, we will discuss the technical definition later on and that definition today is attributed to wold, w o l d he was a pioneer in random signal processes and time series analysis. His definition of what a random processes or what a non deterministic process is kind of the gold standard today, we will come across that a bit later on. But the fact is that a random process is one which cannot be predicted accurately, but it does not

mean number one that it may exist or not that definitely is not there that is out of the definition it is not that.

But there is a situation that mean there are a class of random processes which cannot be predicted at all what we mean by this is given any amount of history of the random process you cannot actually make an improvement over the average the simple average is a best prediction for such processes, we call such processes as a ideal random processes engineers call them as white noise processes for a reason that we will know later on. So, there is an ideal random process and then there is an ideal deterministic process. In reality we would be dealing with either an ideal random process because we do not know, we will only figure out from data that is what time series analysis is about. The first step typically in time series analysis is about figuring out whether we are dealing with an ideal random process; that means, which offers no scope for prediction or it offers some scope for prediction and also to figure out if there are any deterministic trends for example, if you were to go back to the stock market index series that we saw sorry that we saw yesterday.

(Refer Slide Time: 14:47)



If you look at the bottom left, here the Swiss stock market index that we saw you see that there is a trend, correct we do not count those strengths as a part of the random process with, we imagine this random process to be made up of 2 components, a deterministic component and a stochastic component. So, you can see over and above the trend, there

are some irregularities correct, those irregularities are you know constituted or they we assume that they constitute the stochastic part and the trend that you see is attributed to the deterministic part of the or the nature of the process. And you will see many such series where it is the series is a mix of deterministic and stochastic and also many series where you do not see a deterministic component you deal with as a pure stochastic I am not saying ideal stochastic, a pure stochastic where there is or there is not a possibility of making a prediction.

Generally the best prediction that you can make of any random process is the average that is if you are not given any history at all, when there is no scope for improving on this average, we call such processes random processes ideal random process. If the history of the random process has something to tell you improve upon the prediction that is improve upon the average then we say that yes you know I am going to build a model and make some forecasts. Any model that I will build for it for a random process we will still fall short of predicting accurately that is the main I would say that is the core of wold's definition of a random process that is what distinguishes between a deterministic and a random process. For a deterministic process, if I were to give you the infinite history we are talking about theory.

So, we will talk about infinite history if I were to give you the infinite history of the deterministic process you will be your prediction of what the process is going to be the next instant is going to be spot on accurate. Whereas, with a random process even if I were to give you infinite history your prediction will still be inaccurate; that means, there is going to be an error between what you predict and what you will observe of the process at the next instant. So, that is the fundamental difference between a deterministic and the stochastic process.