Applied Time-Series Analysis

Prof. Arun K. Tangirala
Department of Chemical Engineering
Indian Institute of Technology, Madras

Lecture – 88
Lecture 38C - Goodness of Estimators 2 -3

The most important property, desirable property of an estimator is consistency; this one being efficiency other being consistency. In any estimation exercise even today when we keep showing we keep coming up with different methods of estimating models and so on even in our cutting edge research everywhere when we come up with a method of estimation we have to prove either; I should not say by hook or crook, but by theory of by simulation that this estimator is consistent. If we cannot prove that then the estimator is not really acceptable

What is consistency measuring? It is a very simple concept, but related on some maybe some complicated basics of convergence. I have talked about this before consistency.

(Refer Slide Time: 01:12)

It is rest on the notion of convergence; what is the basic idea here? Its basic idea is as follows: I have N observations I have estimated a parameter theta. So, we denote that as theta hat N subscript N. And this sample size can go from one to infinity. So that means, my sequence can begin from theta hat 1 theta hat 2 and so on; and then this can continue.

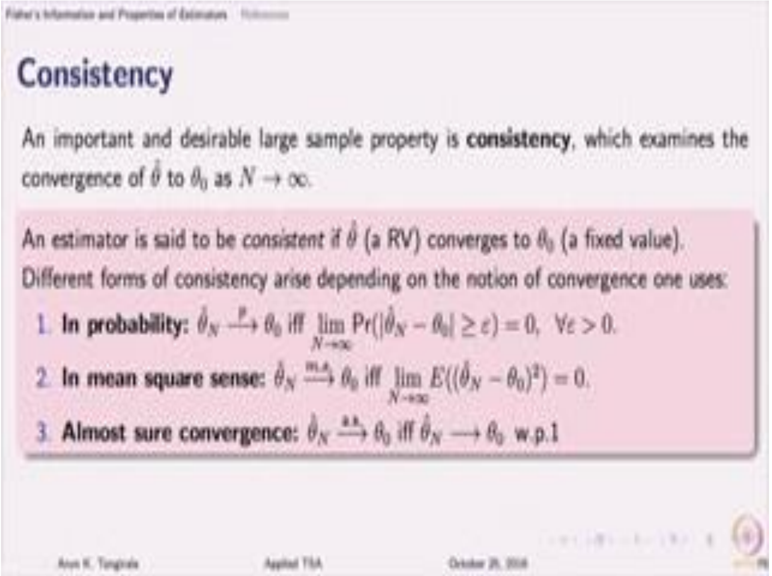
So, imagine that you are constructing a sequence like this based on the number of data points that you have, each estimate that I am constructing is a random variable; remember theta hat is a random variable. Consistency is asking two questions in one shot; what are those two questions? One question is whether it converges; whether this sequence of random variables does it converge. That means we need to have some knowledge of how to treat sequences of random variables and what is meant by convergence of that. Two it is asking whether it converges to the truth. In general when you think of the sequence of random variables you would imagine that it would converge to some random variable. How can it convert to a deterministic variable? But it is possible because after all deterministic variable is also a random variable with variance 0.

That is the beauty of working with random variables, you can bring the deterministic world also into it and also and always argue that a deterministic variable is a special case of a random variable or a limiting case of a random variable in the limit as sigma square goes to 0. What we are interested in parameter estimation is the convergence of this sequence of random variables to some constant; that constant is the truth theta naught. In general what is required? When did we talk about this convergence of sequence of random numbers earlier, do you remember? When we are talk of linear models right, models for linear random processes we asked whether that summation that we had does it actually converge to a random variable, we demanded stationarity but more than that first it should converge to a random variable.

So, consistency requires us to understand some basic ideas about convergence of sequence of random variables. We are used to and we have learnt before primary school or even in your first and second year convergences or convergence of sequence of deterministic numbers and so on.

And it turns out that they are different notions of convergence. Even in the world of deterministic numbers there are different notions of convergence. You have point wise convergence, you have uniform convergence, and then you have convergence in a limit and so on. Likewise here, when we talk of convergence of random variables there are many notions of which three in fact I should say four notions are prevalent, three will discuss will I am just mentioning here the fourth one is convergence in distribution which will talk about tomorrow.

(Refer Slide Time: 04:45)



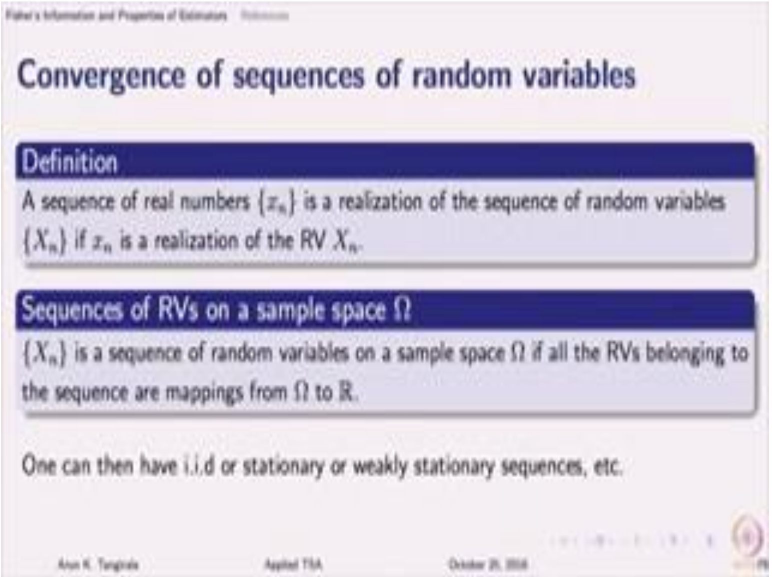
So, the three notions of convergence that one comes across very often in the world sequence of random variables is: convergence in probability, then convergence in mean square sense, and then con almost sure convergence. They mean different things, but all of them are talking about whether this sequence converges to the truth, how close this sequence converges to the truth.

Here the random variable that we are talking about is a fixed quantity and each of them is using a different sense of closeness. Ultimately, it is asking if your random variable goes and sits very close to theta naught and if that closeness actually becomes exact that difference between where it sits and the truth shrinks vanishes as N goes to infinity; that is what consistency is asking.

To understand these concepts I will just go over very quickly the convergence of sequences of random variables and then we will wind up. So, first of all a sequence of real numbers is said to be is the realization of the sequence of random variables if each element in that sequence is a realization of that corresponding random variable. It is a very formal way of saying that I have a bunch of numbers and this number I can call as a realization of the sequence of a random variable, if each of the numbers in that sequence is a realization of the individual random variables that is all.

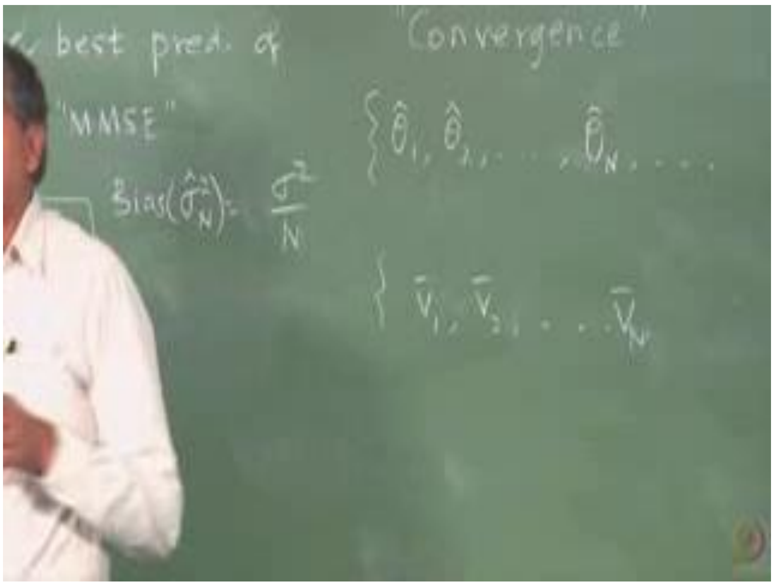
And then there is a notion of sequence of random variables on a sample space omega. I will give you an example do not worry, if this is too heavy the example will lighten it up. So, the sequence of random variables on a sample space omega.

(Refer Slide Time: 06:41)



This just formalization; in your mind you have to imagine that there is a realization right, there are many many realizations that is triggering, that is corresponding to this sequence.

(Refer Slide Time: 06:58)



For example, let us look at these estimates of sample mean. So, I have v bar 1 v, bar 2 and so on right v bar N, what are these v bars here sample means obtained from 1 observation 2 observations and so on.

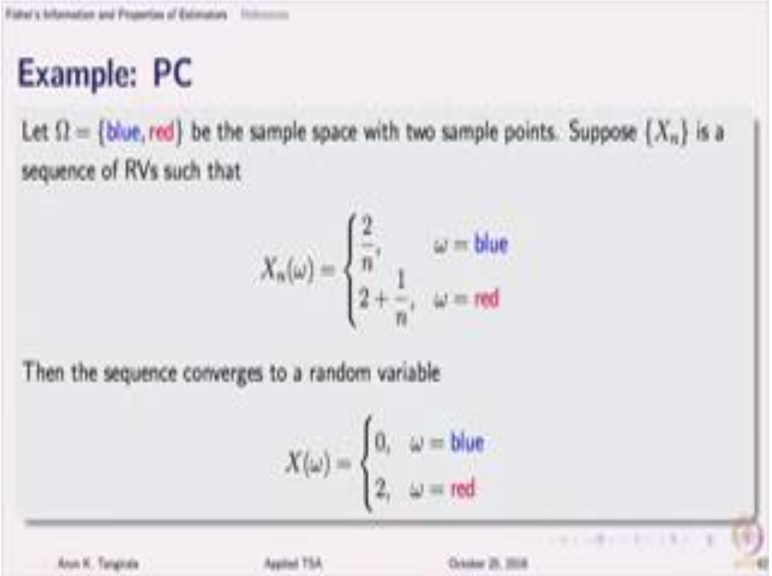
Now, imagine that each of this is being generated, where are these coming from they are coming from your data realization that realization is across some sample space which means you have one data record, but that is only one of the possible data records there are many many many data records that are possible. This sequence here that you are constructing maybe form one realization, you are taking 1 observation of the realization 2 observations of that realization and so on.

But you can construct many of this depending on the realizations that you have and all possible realizations of this sequence you can construct on that sample space, but we will work with a simpler example shortly. First let me define what is point wise convergence its fairly straight forward to understand imagine that there is some event when we say sample space imagine that there is some random event something like I am going to close my eyes and pick an individual and I will go by the colour of the shirt or the colour of the dress something like that or colour of the hair or eyes and that is going to decide what sequence I am going to generate.

And imagine all such possibilities of this random event. In fact, let us think of a situation where the colours can only be red or blue if the colour turns out to be red colour of the dress turns out to be red I will generate one sequence. If the colour turns out to be blue I will generate another sequence. Imagine doing this for every possibility of this event and generating a sequence.

We say that the sequence is point wise convergent if for every possible outcome of the event that is triggering this random sequence this sequence itself converges to another random variable; let me give you an example. So, as I said let there be only two possible outcomes for an event blue or red it could be maybe balls or the colour of the dress or something or the colour of the hair.

(Refer Slide Time: 09:29)



I will generate the sequence this omega is the particular outcome of this event, so this is actually omega is a discrete valued random variable it is not the frequency please do not get confused.

So, if I use if I run into the blue colour then I will generate the sequence as 2 over N if I hit upon red colour the omega is only to trigger the sequence. Once I have hit upon blue colour the sequence is generated according to this. If I hit upon the red colour the sequence is going to be generated differently.

So, now both these possible sequences should converge, they should converge to what to a random variable. You should be able to find the random variable; that this will converge to that is all point wise convergence is asking point wise convergence saying- for all possible events that trigger your sequence every sequence corresponding to that event should go and converge right and they all should converge to a random variable like they all should meet in the ocean. Imagine there are many rivers that are going that are being generated each river is being triggered by one outcome of the event.

So, in this case what does the sequence corresponding to the blue event converge to as N goes to infinity? 0. What does this sequence converge to corresponding to the red event? 2. So, all I have to do is to establish that there is a random variable that exists a random variable with possibilities 0 and 2 and that is all. So, that is very easy, obviously you can define a random variable with two possibilities you say in this sample space there is a random variable which takes on a value of 0 and takes on a value of 2 depending on whether the blue has triggered.

(Refer Slide Time: 11:35)

So, blue triggers a sequence right and red also triggers another sequence both the sequences go and converge to a random variable which takes on a value of 0 when omega is blue and takes on a value of two when it is red. You just have to establish that there exists a random variable. It is possible one can give examples where such a possibility is ruled out, I am not going to give you that this example is only to illustrate the idea of point wise convergence.

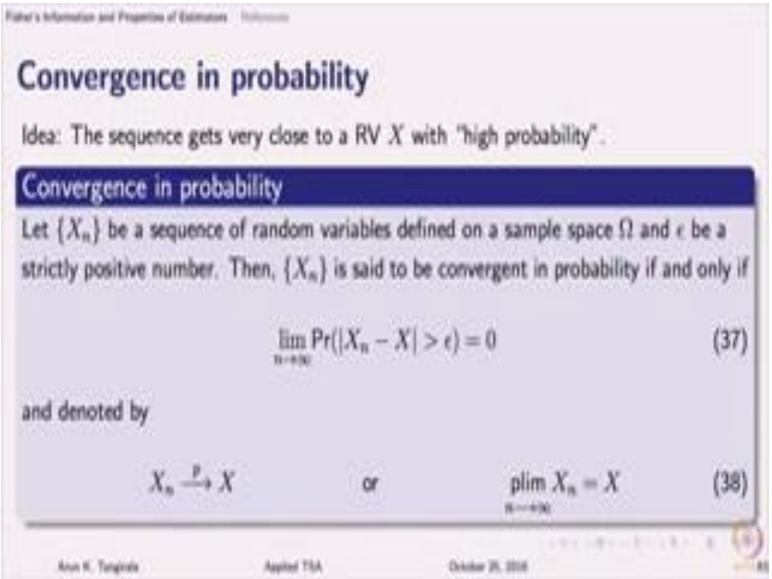
So, the simple summary is there are a set of events that are driving your sequence in this case what is driving your sequence your data your experiment, if the experimental realization is of one kind you get one sequence, if you had a different realization your friend would have another realization you would that your friend would have a different sequence. Imagine using r norm, what is the trigger for r norm there is a seed. So, that seed is your omega. So, one seed is generating one sequence of v 1 bar v 1 bar and so on v bar another seed is generating another v bar; point wise convergence says all such possible sequences should converge number one and whatever values they converge there should be a random variable which should take on these values for every seed that is all, and you can give examples where this is not possible all right. So that is point wise convergence.

The other convergence that we are talking of is in probability, the point wise convergence is perhaps the strictest convergence that you can expect, because you are saying every possible real sequence should converge and it should converge to a random variable, whereas convergence in probability lightens up that requirement it says that this these sequences do not have to necessarily go and hit a random variable as long as they are you can say that the probability that they are outside an epsilon radius of that random variable goes to 0 as N goes to infinity then you are.

So, you will find many such examples in the internet. So, one of the examples may be related to the failing of the device right and so on. So, you will find many such examples, but the simple implication of this is based on the notion of closeness between two random variables in a probabilistic sense.

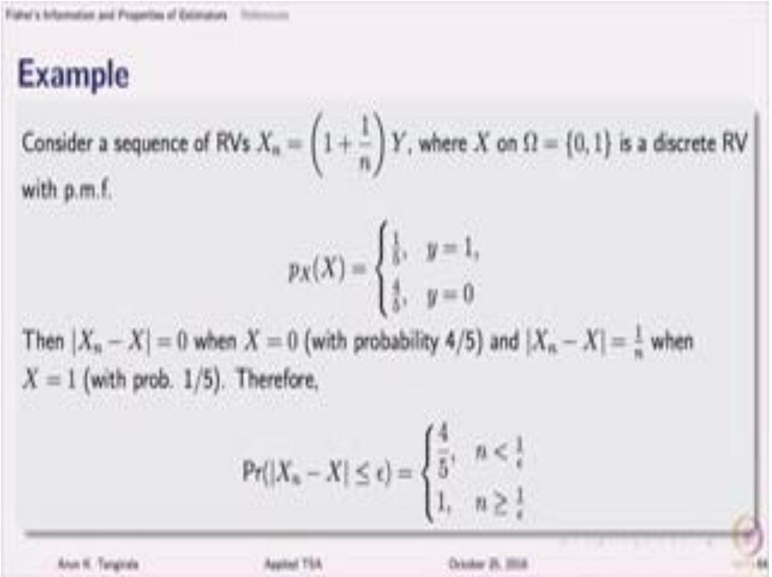
So, all this is saying is that for finite N the sequence will be within some epsilon radius, but that the probability and there may be certain sequences that will go sit outside the epsilon radius, but as you increase the sample size those events that were resulting in a sequences outside the epsilon radius will come into this disc and as N goes to infinity the probability of finding a sequence outside the epsilon radius will go to 0.

(Refer Slide Time: 15:09)



So, it is only saying in a very probabilistic sense you can think of it like a very lawyer's statement or a legal statement which does not give you a lot of room for suing.

(Refer Slide Time: 15:23)



And here is a very simple example suppose I have a sequence that I am generating this way 1 plus 1 over and times y let us say x N is being generated this way; just a contrived example. And this is being defined on again a sample space with two possibilities: if 0 turns out to be the outcome of that event then x, actually it should be 1 plus 1 over an x naught y there is a mistake. So, x will take on a value here also there is a typographical error take on a value of one with probability 1 over 5 and will take on a value of 0 with probability 4 by 5.

Depending on of course now omega being 0 or one x N will take a different route. For example, if x N is 0, what we are interested in is not x n but the distance between x N and x, what happens when x is 0? When x is 0 x N is 0; please remember it should be 1 plus 1 over N times x, when x is 0 then the distance between x N and x is 0 with probability 4 by 5. Again here there is y, but it should be x there should be a there please note that correction.

So, what is the probability with which x takes on a value of 0? 4 by 5. Therefore, the distance between x N and x is always going to be 0 regardless of N, whereas the distance between the sequence that a number in the sequence and x is going to be 1 by 2 when x equals 1 which occurs with probability 1 over 5. So, the events that are triggering the sequence here are only two events I mean the outcomes they are only 0 and 1, and that is resulting in these two distances between x N and x.

Now, if you look at the probability that x N minus x is less than or equal to epsilon it occurs with probability 4 by 5 when N is less than 1 over epsilon and it occurs with a probability 1 when N is greater than or equal to 1 over epsilon. What happens when N goes to infinity? What happens to this probability? Which one applies? 1. It says that the probability of finding x N within epsilon radius of x is going to be 1 which means the probability of finding x N outside an epsilon radius of x as N goes to infinity is 0. So, we say then that here this sequence converges to that x in probability. So, some of these concepts are a bit heavy unless you go through these examples again you may not follow, I think in the interest of time and also the heaviness of this concepts.

I will stop here, tomorrow we will talk of mean square convergence and almost sure convergence, then consistency will be hopefully consistent and clear then we will conclude with convergence and distribution.