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

Lecture – 91
Lecture 39C - Goodness of Estimators 2 -6

Alright, so let us move on and I am going to go pass the summary. Now we will turn to the most important purpose of estimation, which is making confidence statement about the truth. Until now we have talked about how good an estimator is; there was a question that was asked at the end of the class yesterday - if I estimate let us say some parameter using maximum likelihood then is not it the best estimator why do I have to really worry about its properties. Whether it is MLE or least squares or any of such estimation methods, they are estimating parameters in an indirect way, they are not explicitly guaranteeing that your theta hat will go and sit at theta naught; we talked about this when we introduced estimation theory itself.

We said that ideally we would like to pose the estimation problem such that theta hat is very close to theta naught in some sense, but we cannot pose such a problem because it is imposed, I do not know theta naught. Therefore, we (Refer Time: 01:26) the indirect means in least square we try to minimize the approximation errors, in maximum likelihood we maximize the likelihood. All of these are indirectly getting you the estimates, optimal estimates; they still have to be subjected to whatever we have talked about; whether you get unbiased or you get minimum variance, whether you get consistent estimators and so on.

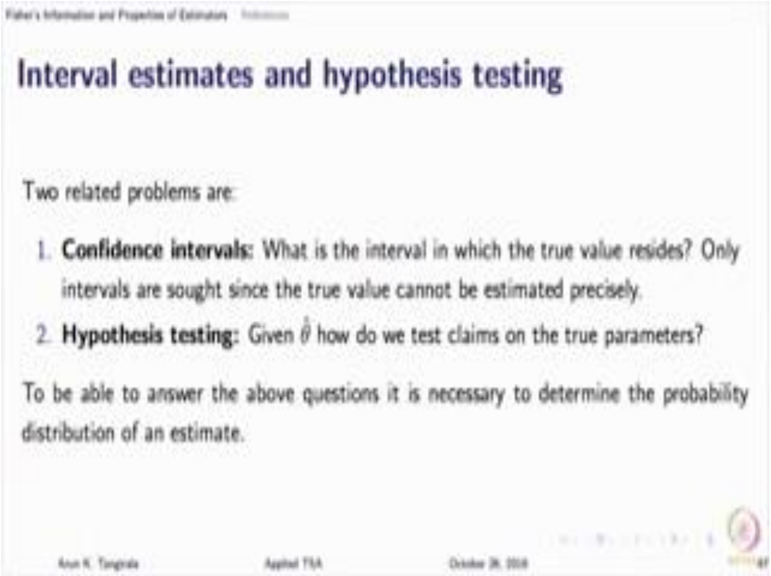
So, all the properties of the estimator that we have talked about are your quality inspectors, they have to come in the moment you developed an estimator and that is the reason we are discussing these up front before even discussing methods of estimation least squares or MLE and so on. Of course, we have introduced MLE to dry home the concept of (Refer Time: 02:10) information, but otherwise we have not talked about methods of estimation first that I would say is the typical procedure followed in many text books, that first methods of estimation are introduced; no I personally feel that first we should be aware of what are the matrix for evaluating the goodness of an estimate and those are the properties that we have talked about.

Now, we are coming to the kind of a final statement a final purpose of estimation which is what can I say about the truth. After all why I am estimating the parameter; I want to say something about the truth correct. Generally we are used to giving some number which we call as point estimate, but now unfortunately because of randomness in data and fix sample sizes, I can only; I cannot pin point the truth, I can only give an interval and how the question is how do we construct this interval.

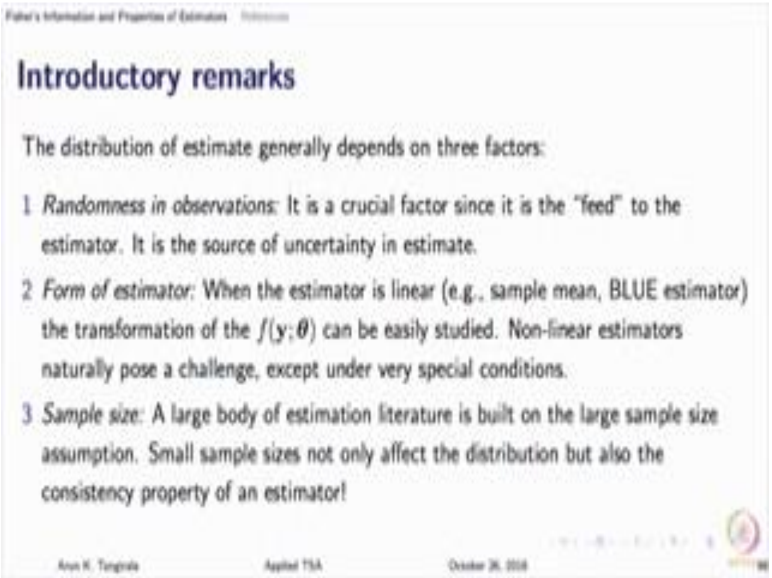
In order to construct this interval, we need to go through another concept called convergence and distribution that is we need to know the distribution of theta hat; what are all the possible values that I can get for theta hat, for a fix sample size. Now we are you can say; we will come back to this same sequence, but for a fixed sample size what is a distribution; that means, if I fix the sample size like we did in the simulation let us say I freeze one sample size, I look across realizations and I look at the distribution that will get me the possibilities for what I have actually estimated.

From this distribution, we can construct what is known as a confidence region. So, it is a two step procedure that we are adopting; first we estimate theta hat when they calculate theta hat, step two we compute the pdf of theta hat and then from the step two we construct a confidence region and this confidence region will allow us to test many statements of hypothesis. So, in general you will see hypothesis testing I hope that all of you by now have gone through the video lectures on hypothesis testing; if you are not it is time to do that now, because I am going to just give you a very breezy review of hypothesis testing, but if you have sat through the lectures, you would have realized that we talked about two concepts - hypothesis testing and confidence intervals and this confidence intervals can be used in hypothesis testing and hypothesis test can be conducted without construct in confidence interval also. We will mostly focus on confidence region, but we will go through an example on hypothesis testing as well.

(Refer Slide Time: 04:56)

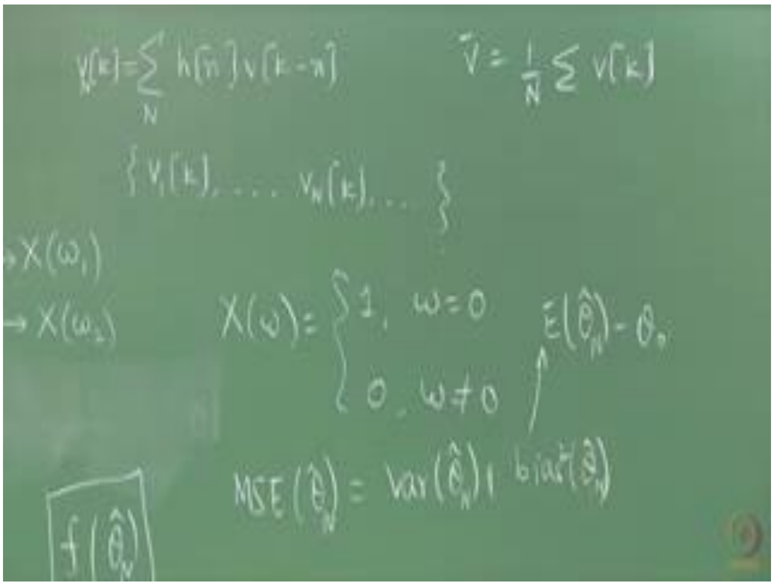


(Refer Slide Time: 04:58)



So, the goal is now to construct interval estimates and as I said earlier, the key for constructing interval estimates is a distribution and the distribution that is f of theta hat. This depends on the sample size remember we are worried about a fixed sample size.

(Refer Slide Time: 05:20)



So, this distribution here if vector of parameter it is going to be joint pdf; depends on the first and foremost is a randomness in the observations in your data. Secondly, the form of estimator whether it is a linear estimator or a non-linear estimator and thirdly it depends on the sample size, why let me give the classic example of the central limit theorem. What does the central limit theorem says; if I combine linearly combine a bunch of random variables that have identical distribution and that are independent; regardless of the distribution of the variables that I am combining, the resulting variable will tend to have a Gaussian distribution; it is a tend to have a Gaussian; that means, not for all sample sizes you will get a Gaussian distribution, only for large sample sizes you will see a Gaussian distribution; however, if you are combining random variables that have a Gaussian distribution themselves, then you are guaranteed to get a Gaussian distributed random variable regardless of the sample size.

So as you can see that the sample size place a role; the nature of the randomness plays a role and the form of the estimator; what does central limit theorem assume, where is central limit theorem applied for example, in deriving the estimate distribution of the sample mean, what kind of an estimator is this? Is it a linear estimator or non-linear estimator; it is a linear, linear and observations. I can straight away apply CLT to this assuming that v k is falling out of a Gaussian white noise process, I can straight away say if it is Gaussian white noise, v bar is always Gaussian distributed regardless of n, but if v k is falling out of let us say uniform white noise process, pause on white noise process and so on then only a zen goes to infinite, v bar will have a Gaussian distribution; so that is nice.

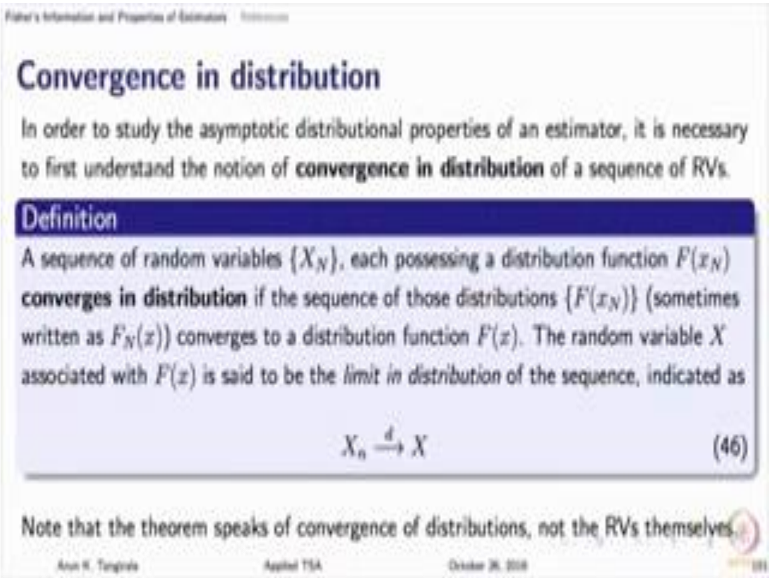
So, I can actually CLT is a cool limit theorem, but unfortunately I can apply that only to linear estimators, if I have a non-linear estimators such as the one that I have for moving average model estimates and so on, we will show that auto regressive model estimates obtained using least square methods will also result in linear estimators, but in general I may encounter non-linear estimators; I cannot apply CLT. Then we have two choices, one is to look what are known as asymptotic results that is a zen goes to infinite what happens? Predominantly in the literature that is a problem that has been studied, what is a distribution of an estimator as n goes to infinite, that is called an asymptotic distribution and where you are asking now I have a sequence of estimates to what distribution, to what random variable do they converge in the sense of distribution. We are not asking it to converge to some random variable that distribution of this sequence that we have a zen goes to infinity.

(Refer Slide Time: 08:42)

So, in place of theta 1; imagine that now I am constructing sequence of distributions right, I can do that even in simulation I can fix a sample size walk across realizations concern the pdf.

Imagine that now you have a sequence of distributions, this sequence what does it converge to CLT says it converges to a Gaussian; if theta hat has been obtained in a particular way, but there is no universal theorem that will tell you what will happen in a general case. For certain estimators asymptotic results have been obtained for example, if you take maximum likelihood estimators of parameters of a pdf or a model asymptotic results are available, but essentially we are talking of convergence of distribution, I will come back to this point later on; let me make this statement and then we will I have already talked about central limit theorem.

(Refer Slide Time: 09:45)



So, we say sequence of random variables possessing distributions function converges in distribution to some random variable, if you are able to show that there exists a random variable with that. First of all it should converge and then you should show that there exist a random variable which has the distribution; convergence in distribution is the weakest form of convergence because it is only talking about distributions converging, it is not talking about the random variables themselves converging to a another random variable.

So, central limit theorem is one such statement and I am going to I have talked about all of that. For tomorrow's class please if you have not gone through the lectures on hypothesis testing at least take up some relevant one or two videos, go through hypothesis testing because what we are going to do is just walk through an example for 5 minutes and then also about 5-10 minutes on confidence intervals and then we will start of one least squared methods, so please come prepared for tomorrow's class.

Thanks.