Applied Time - Series Analysis

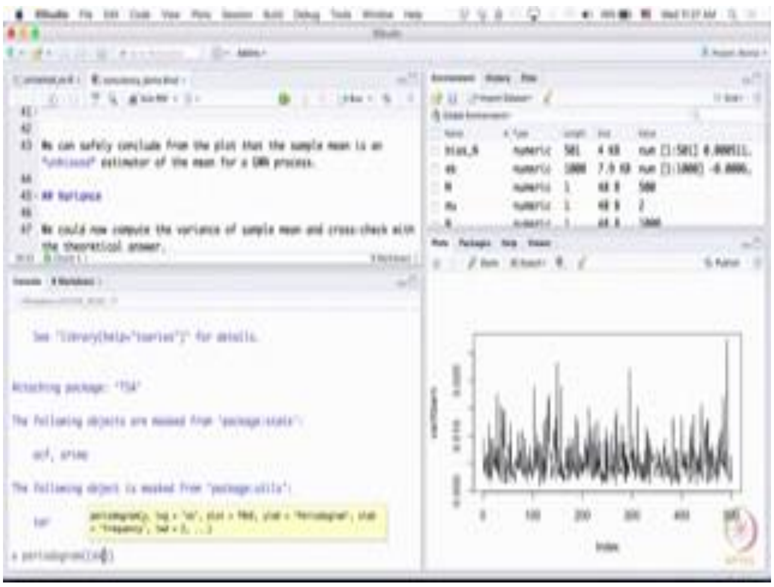
Prof. Arun K. Tangirala
Department of Chemical Engineering
Indian Institute of Technology, Madras

Lecture – 90
Lecture 39B - Goodness of Estimators 2 -5 (with R demonstrations)

Consistency is a very very important because a time and again I keep saying this because it assures the experimentalist, but as you draw more and more observations you are getting closer and closer to the truth that it has to be assured. In fact, it turns out that there are estimators which do not conform to the consistency requirement and the classic estimator is a periodogram that may come as a surprise, but the straightness is not complete if I use periodogram to estimate these spectrum of a random signal we are not talking of deterministic signals if I use the periodogram what is periodogram you take the d f t square it and divide by the sample size, that estimator does not go and sit at the truth or even get closer to the truth as you keep increasing the sample size.

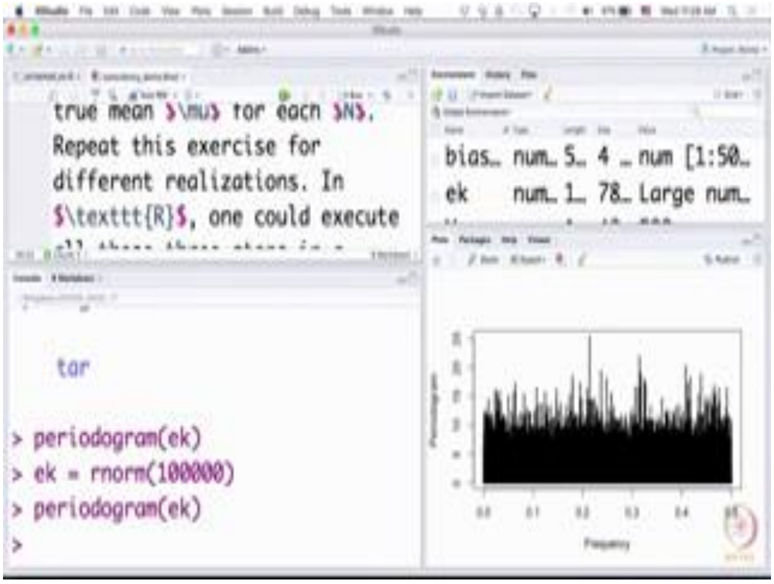
Let me just illustrate we will of course, see this later on, let us take the white noise process.

(Refer Slide Time: 01:18)



Let us generate for example, 1000 observations of a Gaussian white noise process standard Gaussian white noise and I need to. So, let us plot the periodogram of this finite length realization of a white noise what should be expect to see, what is a theoretical spectral density periodogram gives you a empirical spectral density. So, what do we expect to see?

(Refer Slide Time: 01:56)



Student: (Refer Time: 01:57).

What is theoretical spectral density of a white noise?

Student: (Refer Time: 02:02).

Flat correct.

You only see the base being flat that is not a big deal because what it says is that it is non zero, it is non-negative value, is it anywhere close to being flat. So, you may say no I look I have not taken the entire population that is a theoretical spectral density. So, then we can go back and generate we can add 2 more 0s. So, that we take a 100,000 observations and let see the estimate improves, did it improve? You cannot say it got worse it is only looking more dense because a frequency spacing as come down. The number of points on x axis as increased, but of course, you can say that it has not gotten better why is this happening because this periodogram estimator you can say therefore, is not a consistent estimator.

Although we are not looking at variance or we are not looking at anything else the fact that your estimator is not getting better and better as n goes to infinity your estimates are not improving. On the other hand, what about sample mean, you can easily show. So, let us do that with the package here. So, let me tell you what I am doing here the purpose of this document is essentially to illustrate the concept of convergence and probability and so on, I do not illustrate all concepts of convergence, but let us at least look at some concepts of convergence and our focus is on the estimation of mean using the sample mean estimator.

Let us generate a Gaussian white noise realization of sample size m. So, that is our reference process here Gaussian white noise process. Let us compute the difference between sample mean and the truth because what does convergence and probability say convergence and if the sample mean.

(Refer Slide Time: 04:07)

Is convergent in probability then the probability of finding X bar N minus the truth outside an epsilon radius should go to 0 in the limit as n goes to infinity right mu is the true value because this simulation I have to fix the truth in theory it can be symbolic.

I have to show that at least I have to be able to see some simulation that the probability of this magnitude of difference being outside an epsilon radius epsilon is an arbitrary small number, but not 0 that probability should keep shrinking. We would not compute the probabilities you can in fact, the package that I said convergence concepts that package computes the probability for you it is a very simple formula. But what we will do is instead of plotting the probability we will just plot this difference and see whether that difference actually keeps shrinking and stays within some bounds as we increase the sample size.

What do we have to do? We have to do 3 things, first thing is we have to generate the realization of a sample size N that is the first step, second step is to compute this difference for a given sample size and 3, repeat this for all possible realization, it is impossible to generate all possible realizations we will generate about 500 realizations and we will go up to n 1000 we cannot go do this for n going to infinity, we will go from sample size one 2000 fortunately, you can do all of this in a single line in r and I given that code here and that code is here.

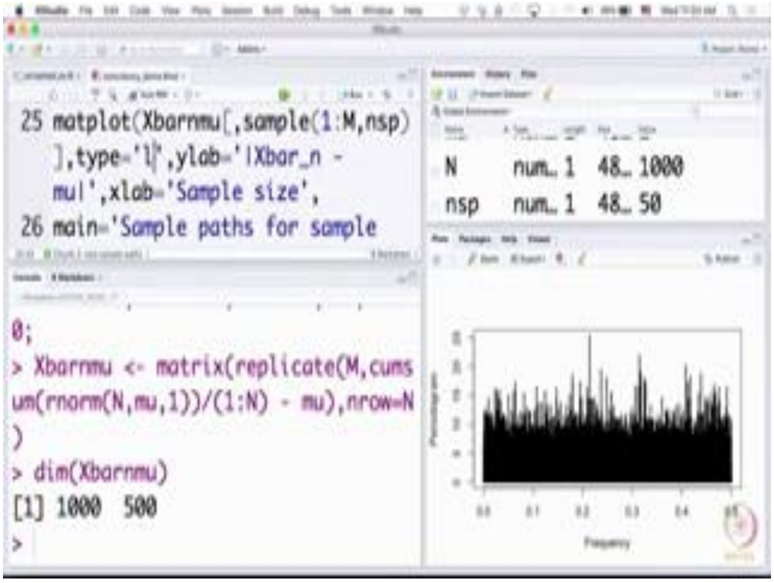
First I set the parameters for my simulation thousand data points up to maximum of thousand you can change that. In fact, n should be radius n max, m is a number of possible realizations that I am generating here I am sampling 500 different points and I am fixing the true mean to be 2 and the number of sample paths. What we mean by sample paths is remember this x bar is nothing but your some estimate for every realization as I keep changing n you have a path and I am going to. So, we have 500 such sample paths here plotting all of them will make it look very dense we will just plot some randomly selected 50 realizations.

Now what you do is compute this difference in this single shot inside you have cumsum, what is cumsum doing?

Student: (Refer Time: 06:52).

It is adding up for different n and then do not forget to divide by the length of these sequence so that is for a single realization then you repeat this that is what replicate is going to do then you want arrange the results in a matrix so that each column corresponds to a realization and when I execute this particular chunk of code sorry. So, let me go here and execute.

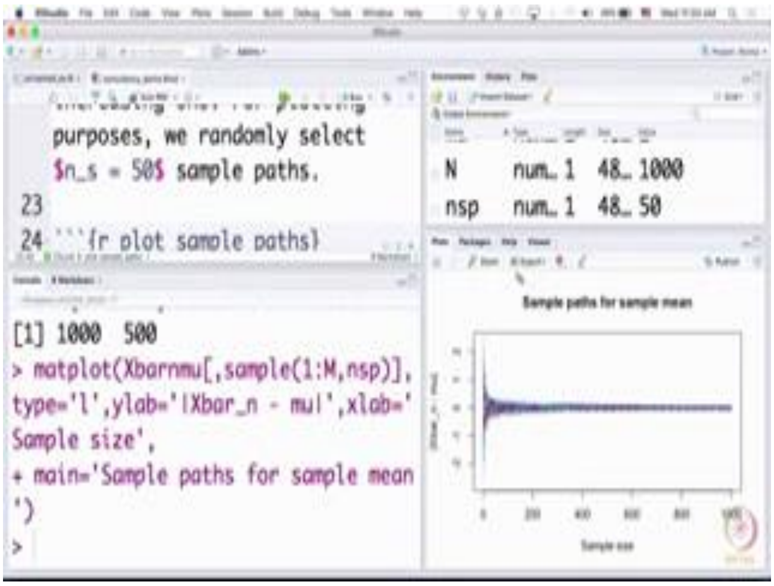
(Refer Slide Time: 07:22)



It has executed and if everything has been done correctly, you should get this dimension 1000 by 500 at least qualitatively what you have done is correct.

Now, what is a next step? To plot the resulting sample paths, so I am doing here the same thing. So, I am just using mat plot instead of using plot matplot works on each column. So, use matplot, but do not forget to mention that you want to a line plot, if you want a scatter plot on top of line it will take time to plot, the moment you say line plot it plots it very quickly. So, let us execute this chunk here.

(Refer Slide Time: 08:05)



I will just zoom this out so that you can see what is happening. So, this is the sample path that you have as you can see.

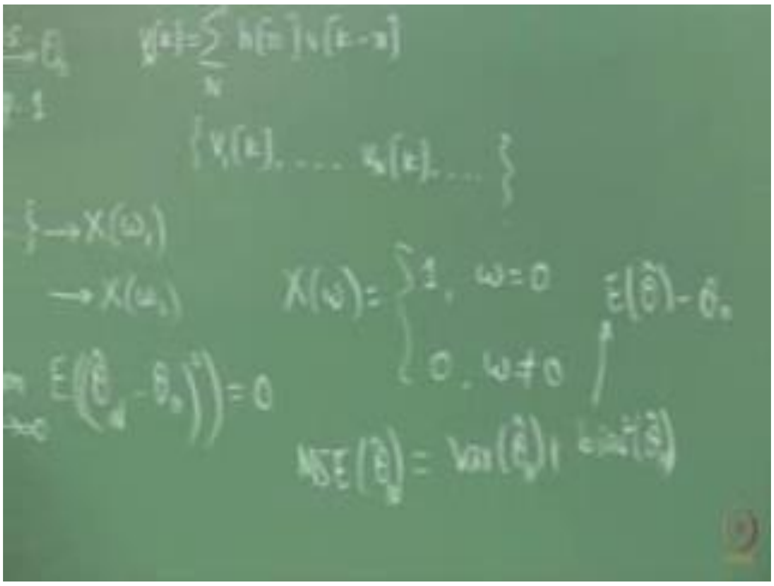
What is on the x axis is the sample size which is n. So, if you are standing here then n is one and you are computing the difference between x bar computed from one observation minus the truth which is true and that difference; obviously, is going to be quite large.

As I keep increasing the sample size the distance between the estimator estimated value and the truth keeps coming down and different sample I am only showing you 50 sample paths ideally you should go back and do it for many many sample paths you can see that they are all within some bound and if you plot all the 500 realization there may be some outside a specified bound I am not specified the bound here, but some differences between the estimated value and the truth can go out of bound right. What probability a convergence probe in probability assures you is that that the probability of finding such sample paths outside a specified bound will keep going will keep falling down as n goes to infinity.

The rate at which it falls down or and so on is not specified and that depends on your epsilon also it says that [FL] that it does not tell you when whereas, the almost sure convergence can be interpreted as follows there not much different, but there is in one respect big difference. In almost sure convergence also it is talking about how close the estimate is getting to the truth as n goes to infinity for all possible sequences, but it says there exists a finite sample size after which you all the sample real all the estimates will stay within a bound.

But again that does not tell you what is the finite sample size where as the convergence and probability says it can only happen at as n goes to infinity where as almost sure convergence says there exists a finite sample size it does not tell you where what is that, but there exists a finite sample size after which you are at truth, at the truth. So, practically you may say what is a big difference, but you know it is assuring you in different ways that is all it is it is just the difference different ways in which it is assuring convergence what it makes them different, but all of them are guarantying some form of convergence. And this is the minimum requirement converge convergence and probability is the minimum requirement, is it clear now. At least you know we have illustrated this for sample mean for a Gaussian white noise process you should go and change the distribution or you should work with the different estimator, clear on with this document, I will post the document just play around see if you know as we increase the sample size does it get better if you change the distribution or some other or some other estimator and so on.

(Refer Slide Time: 11:44)



Now, the other thing that I wanted to show you through this simulation is computing bias we have learnt how to compute bias theoretically right what is a bias defined as, the bias is defined as expectation of theta hat minus theta naught this is your bias. Theoretically it may be possible for some simple estimators, but increasingly today people are turning to simulations Monte Carlo simulations. So, if you where to be using Monte Carlo simulators to actually compute this bias, how would you do it what is a simple procedure? What is the simple procedure?

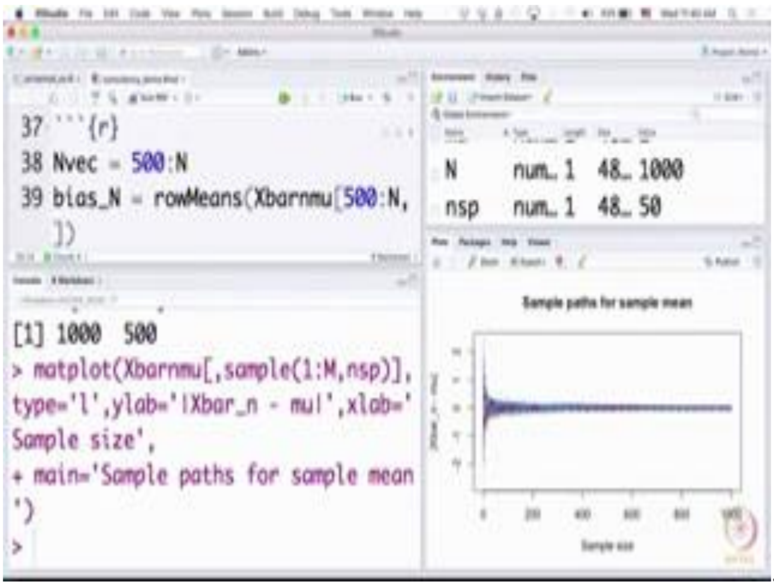
Student: (Refer Time: 12:16).

Generate many realizations of your data for each realization compute a theta hat remember bias is only looking for some fix sample size you do not have to vary the sample size right, bias as got it is not looking at a sample size direction at all it is looking at the direction of realizations. So, this expectation here in simulation should be replaced with what with averaging across many realizations and more realizations you have better is your calculated expected value in this example we have used 500 realizations so we can compute the bias.

We can compute the bias and I have done that here and that is given in this chunk of code all I am doing here is I am not considering small sample sizes. Remember we have simulated from 1 to 1000 max I am looking at five sample size of 500 or more and I am plotting the bias for each sample size essentially what I have to do is I have already computed the difference right I have already computed the difference between x bar and the truth and all I have to do is now take the average across the realization.

X; the x bar minus mu is arranged as with sample size across rows along rows and the each column corresponding to a realization. So, what I am I doing here? I am computing the bias and I am calculating the average of this x bar minus mu across realizations.

(Refer Slide Time: 14:05)



.

Student: (Refer Time: 14:12).

Sorry.

Student: (Refer time: 14:15).

Yes. So, if I am standing at a row I am fixing the sample size correct for a fix sample size I have to average across realizations. So, what row mean does is actually it stands at one row and take takes the mean for a row, do not get confused with that row means would mean that you are standing at a row and it is looking at the average for a row yeah it can be a bit confusing, but that is what row means fine let us actually. So, this is the bias that you get for different sample sizes, what is a theoretical bias that we should see?

Student: 0.

0; obviously, because I am averaging only across 500 realizations so, what you see on the x axis is sample size and I pick any sample size, it should be 0 that is what theory tells me regardless of the sample size the bias is 0 because it is an unbiased estimator. On the other hand now you should repeat this exercise for sample variance what would happen to sample the bias in sample variance. For small sample sizes the bias would be large and if it is asymptotically unbiased estimator this bias will go and sit at 0 for large sample sizes.

Play around with this document so that you understand this concepts much better more modern way of doing this here we are able to generate different realizations because the data generating process is in my hand, I have an evaluation equation or a random number generator. What would you do in practice? In practice you have only a single realization and let us say you are estimating some complicated parameters that is some parameters which as a very complicated formula, how do you generate artificial realizations? Because the data is come from an experiment you cannot ask the experimentalist to do some 500 or 1000 experiments, it is not possible.

What you do is you turn to what is known as bootstrapping techniques, these bootstrapping techniques are very powerful today and increasingly becoming a very popular there is a whole theory around these bootstrapping techniques. The idea in a bootstrapping technique is to generate artificial realizations out of what has been given, whatever has been given you simplest idea is suppose I give you a realization of white noise. What you do is, you re-sample you collect all the observations put them in a bag because it is white noise I do not have to worry about time sequencing I just collect all the observations of the given sample size just put them in a bag and now start re sampling and then there are 2 options when you are re-sampling you can re-sample without replacement, with replacement with replacement is recommended.

Because without replacement can cause problems what is that because the probability distribution can change you do not want to do that with replacement is always preferred and there are some very nice packages in r or even in mat lab and so on which do the bootstrapping for you essentially the idea is re-sampling.

That is what you use in practice. So, there are 3 approaches to examining the properties of estimators, one is theory but theory is useful when the estimators are simple expressions then simulations when you are illustrating some concepts or when you are you are simulating a process to generate the data then comes bootstrapping which is more relevant to a practical situation where the estimator is complicated and you have only one realization of one data record so that in that order you turn to examining the properties of an estimator, any questions on this simulation? So, take this document and try it out with other for example, you can calculate variance also in the same way as we have calculated bias.