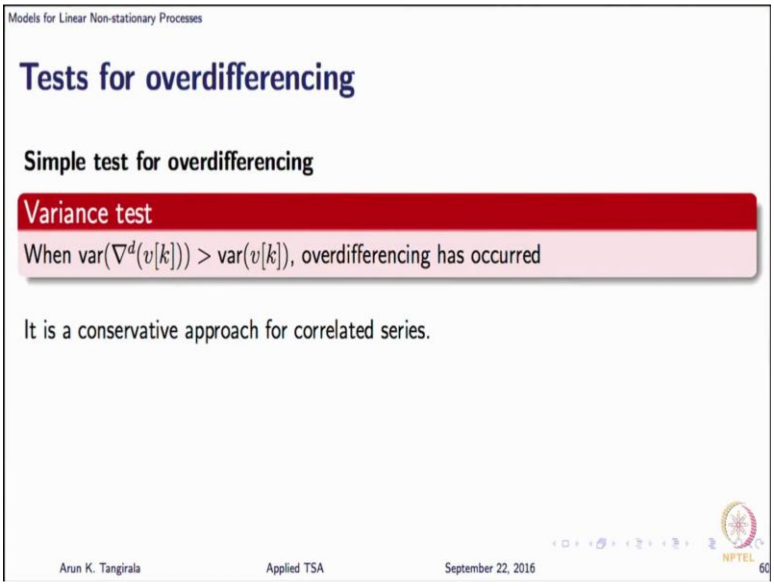
Applied Time-Series Analysis
Prof. Arun K. Tangirala
Department of Chemical Engineering
Indian Institute of Technology, Madras

Lecture – 58
Lecture 25B - Models for Linear Non-stationary Processes 6

So, how do I make sure that I do not end up over differencing? Then there are couple of test to help you detect that your over differencing, but the even without using this test if you exercise enough caution you can avoid over differencing. For example, keep looking at the ACF, conduct unit root test before you proceed to differencing that will by a large avoid this yield step of over differencing. But let us say that I do not know I have written some automated procedure and so on. One of the simple ways of testing for over differencing is to look at the variance of the differenced series.

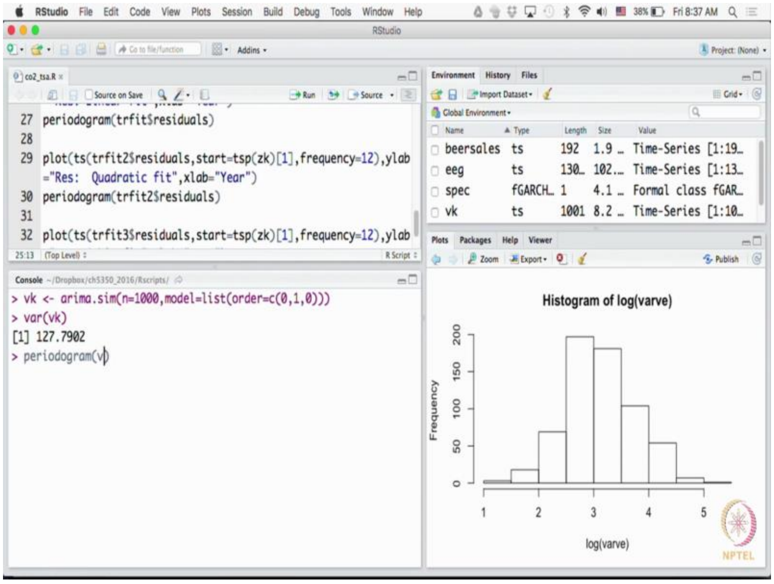
(Refer Slide Time: 00:57)



If the variance of the difference series is greater than the variance of the original series then you have over differenced. At least if the given series has white noise like characteristics you that is why it is a conservative test. For example, imagine that you had in place of w like in the previous slide we have e k in place of w k, then what you have here? E k; what is the variance of v d? Sorry, what would be the variance of v d to sigma square e are you planning to cal cancel out the variances? Ok correct. That is a correct position like this where. So, perfect at least the reaction is perfect. So, the variance of v d is to sigma square e where as the variance of the original series is sigma square e.

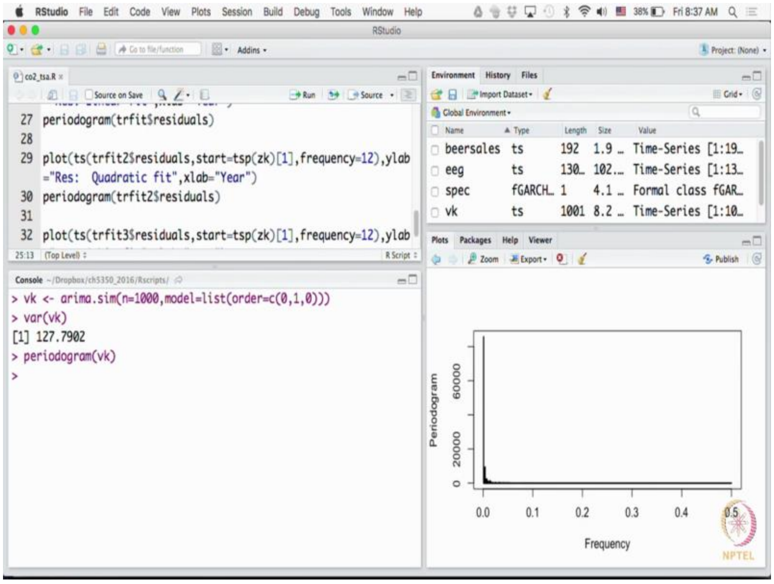
On the other hand if truly in place of v k you had an integrating process then differencing will reduce the variance of the series. At least you can look at it from a sample variance view point. For example, I do not know I let us go and look at what happens here.

(Refer Slide Time: 02:22)



So, let say I have here and ARIMA process, so I am simulating in AR 1. Let us look at the sample variance and not looking at the theoretical variance. So, this is the sample variance, right very high.

(Refer Slide Time: 02:44)

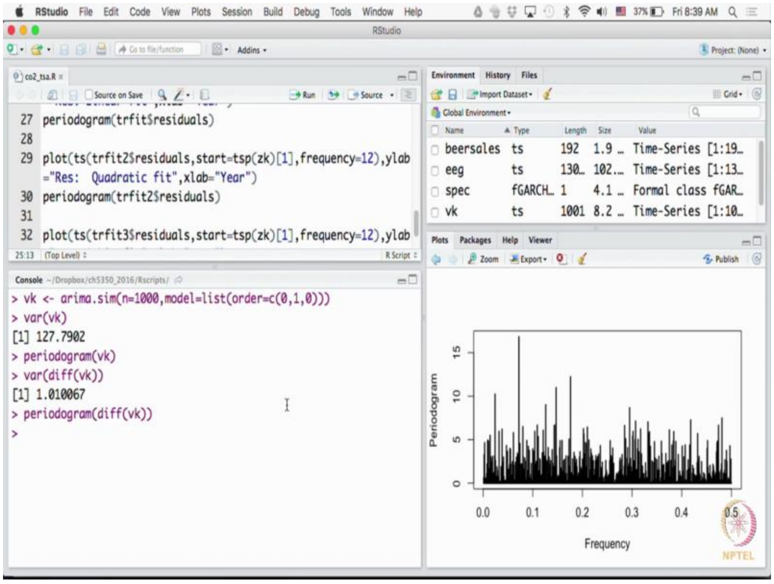


In fact if you also look at the periodogram of v k in find this is the typical characteristic of an integrating process or a near integrating process, because I do not know how many of you remember even from a (Refer Time: 02:26) background integrators are this ideal low pass filters. You recall, so this integrator is a ideal low pass filter it has allowed only low frequency is to go through and therefore the series has predominantly low frequencies; predominantly does not mean that other not there but it just over shadows everything.

Now, if I look at the variance sample variance of differenced series, what you expect? Do you expect to shoot it up or come down? Drastically it should come down what have you done by differencing that is why the filtering perspective is just beautiful and I am slowly showing the seat for frequency domain. Yes, so we have already argued that differencing operation is a high pass filtering operation; that means, it allows only the high frequency is to go through. Therefore, what have to done here? You have taken a process which predominantly has low frequencies and passed it through a high pass filter. What does a high pass filter do? It takes out the low frequencies, it just says you are not allowed to go in, only people with these ID cards are allowed to go through.

As a result now the differenced one has got in read of the integrating effects. So, look at this a same operation that you have been looking at it from an integrator and differencing perspective, now we have we can give a frequency domain flavour to it.

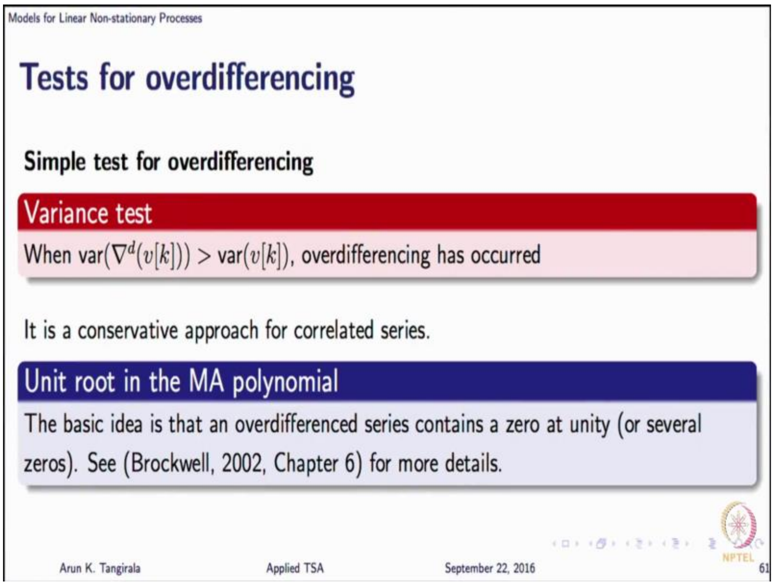
(Refer Slide Time: 04:32)



Now you can also look at the periodogram of and a promise is going to look bit crazy, but it definitely not going to look as clean as this showing heavy low pass low filtering characteristics. Now what do you see? You see all frequencies contributing to the power more or less. Theoretically you can show that since you have take an a pure integrating process we will show theoretically next week that what one should expect is a flat spectrum for the differenced series.

Now, what I can ask you to do as homework is go back and create and integrating process of order 2, and see how gradually the differencing operation is slowly taking out the integrating effects. Of course, do not do the reverse mapping that is always when you see low frequencies predominant that you have an integrating effect, but at least there is a strong case for differencing that is something that you can keep in mind.

(Refer Slide Time: 05:36)

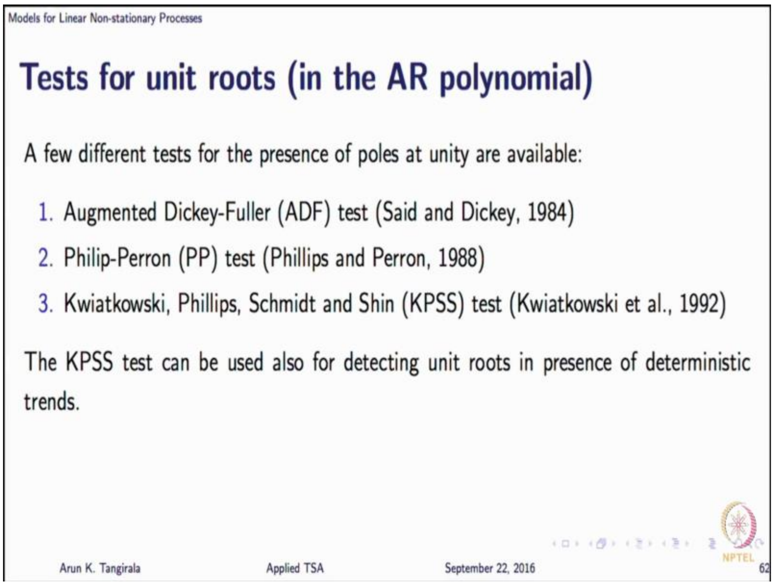


So, the other test that is available is the test for unit root in the moving average polynomial. This is also widely used and there are some test that are available I am not going to go further into a in detail on this aspect, but this allows you to do two things: one tests for over differencing, if you have over differenced what have you actually done? You have introduced and artificial 0 at a unit circle or at unity. And that is what this test will tell you and remember we said we can also use differencing to handle trends.

Suppose there are no integrating fix but you found the trend, and let say you did not pay attention to the trend you just looked at ACF and we seen this earlier when I have a they when the series has a trend the ACF decays very slowly in that case also. And without looking at the series in a rush you simply went I had not difference state. And we have said that earlier that is not the correct way to do with because it introduces an artificial 0, unnecessary 0 at the unit circle or at unity. And this test can be used to distinguish. Therefore, between series with integrating effects and series that has trends.

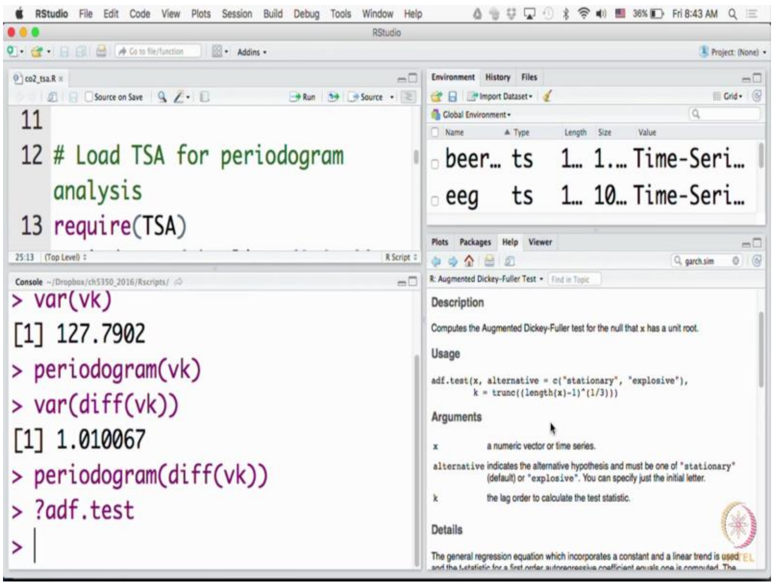
So, coming to formal test for unit root, now we are talking of unit roots in the sense poles on the unit circle, there are in the literature many, but among which three or quite popular the augmented dickey fuller test.

(Refer Slide Time: 07:10)



And the Philip-Perron test and the KPSS test I am not going to read out the author's names, but there are a few other tests also available. These three test are coded in our, you have ADF test and then you have pp dot test and then you have KPSS dot test. For example, here we can quickly run the ADF test on the series that we have created; I am sorry I am going to increase the font size here.

(Refer Slide Time: 07:44)



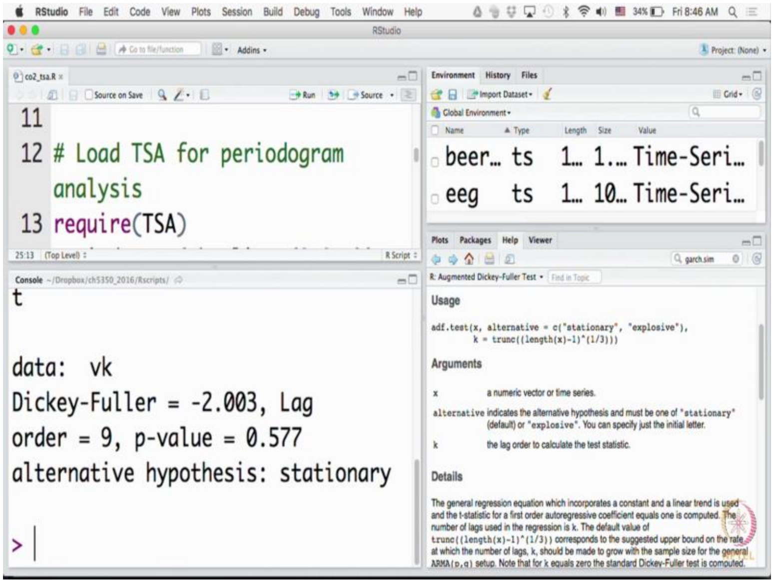
So, we can run an ADF. In fact, you should look up the help on ADF test. So, if you look up the help on ADF lit says computes the augmented dickey fuller test for the null back x has unit root. Whenever you conduct a hypothesis test make sure your clear what the null hypothesis is, it may be counter to what you think. So, here the null hypothesis is that it as a unit root. And going by the p value the significant level you can set for example, here (Refer Time: 08:24) significance level is set to be 0.05, but it does not worry about a significance level it says. I will give you the p value you use at your significance level if the p value is less than the significance level then what do I do? Reject the null hypothesis.

Remember if the p value is low null hypothesis is must go. So, you said your significance level after the p value is obtain and figure out if the null hypothesis is to be rejected, but you should also look at the alternative. It says the alternative is that to be one of stationary or explosive. Why these two alternatives are given? So because the integrating process is not an explosive series, always variance seems to change with time, but because of the nature of the mean non stationarity it does not run away from you.

But it is not stationary either; it is like a cat on a wall. And in the deterministic well such processes are called marginally stable processes. We have integrating process even in the deterministic world. In the systems theory language these processes are called marginally stable processes, you just marginally stable they do not belong to either category. So, the integrating process is neither stationary nor is it explosive, and now you want to test this null hypothesis that it is integrating against one of the two alternatives. And the two alternatives that you have are is that it is stationary; that means the pole is well within the unit circle and the other alternative is that the pole is outside the unit circle. That is how you can look at.

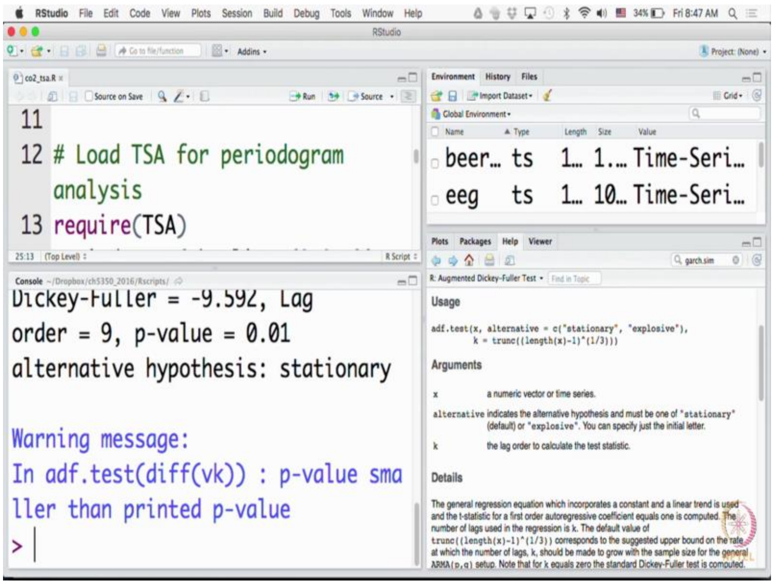
So, the default is stationary it says that. So, let us ask here that is asks a test what it has to say for v k, what you expect? Common this is at least we can say we know we are the creators of v k. What you expect the result in terms of p value let us say. Always when the two options there is a division; what is in null hypothesis? 0 and pole, at a pole at unit circle. Does v k has v k come out the process with a pole on the unit circle? Yes. So, what you expect for the p value lower high (Refer Time: 11:08) that is all. We just that to think logically do not think that you are in a pressure cooker and we have to go bang here and there.

(Refer Slide Time: 11:16)



What you say of the p value what you see, fell very high; very in the sense does not matter high or very high, but generally the significance levels that we choose the alpha which are the type on errors are 0.05 or 0.01 and the p value is higher than that. Which means there is no evidence to reject the null hypothesis that it comes sort of unit circle.

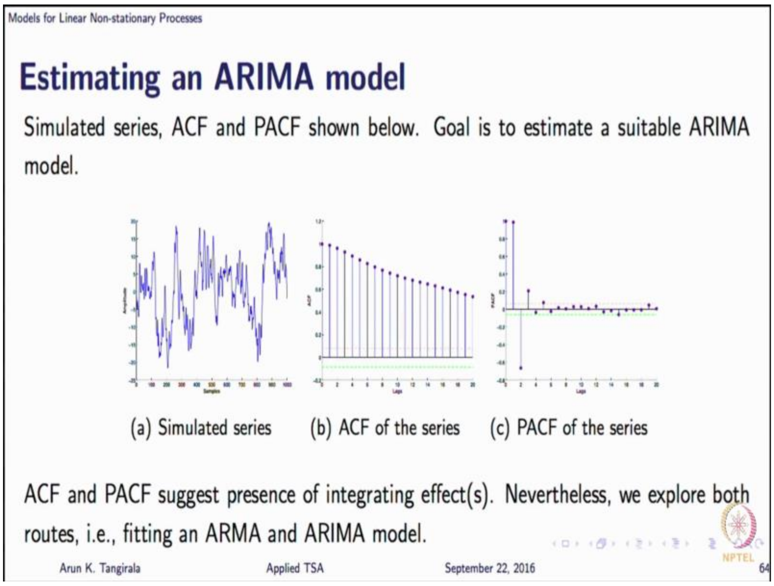
(Refer Slide Time: 11:42)



Obviously will be curious to know what does it have to say about the difference series, what you expect know election result to be? So, now its 0.01 and typically if you choose alpha as 0.05 then you have to reject this null hypothesis. Of course, it just sits on the border line there alpha equals 0.01, but the p value has to be lower than your alpha it cannot be exactly equal to that.

Anyway, this is the way you formally use a unit root test. And the ADF test has its own demerits; and as you can see chronologically the ADF test came earlier and then the Philip-Perron test and then the KPSS test. So, people have been busy doing their PhD is on these kinds of problems. And subsequently also few other test have come about. The KPSS test can be use for detecting unit roots in presence of deterministic trends as well.

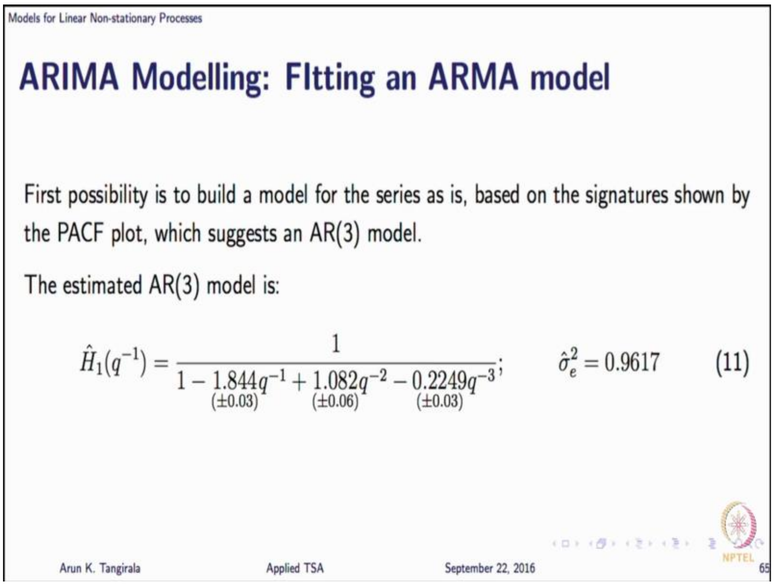
(Refer Slide Time: 12:53)



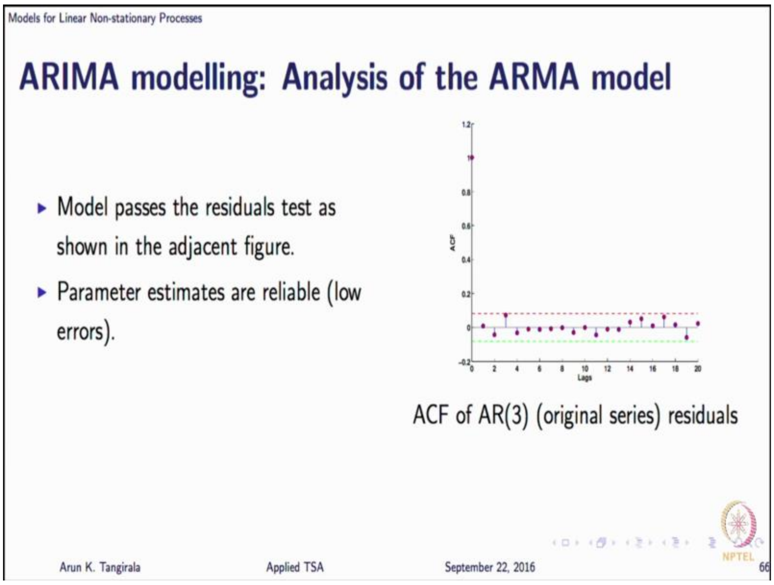
So, let me very quickly go through take a couple of minutes and estimate, go through this example of estimating ARIMA model and then I will conclude that various stabilizing transformations that will end the non stationarity business. So, here is a simulated series and we want to see what it means to fit an ARIMA model, what are the things that we are watch for? You have seen this kind of plots before, simulated series ACF showing very slow d k. So, there may be possibility that there is an integrating effect. Of course, you can run the unit root test and not doing that that is a simple exercise.

So, the PACF shows that yes that there are auto regressive effects and if you at to fit in AR 1 model it say the p ACF at lag 1 shows that you may end up with the pole on the unit circle. But, if you are to fit an AR model for the given series what is the order there it is suggesting third order. It is suggesting third order there are three significant one at least to begin with.

(Refer Slide Time: 13:54)

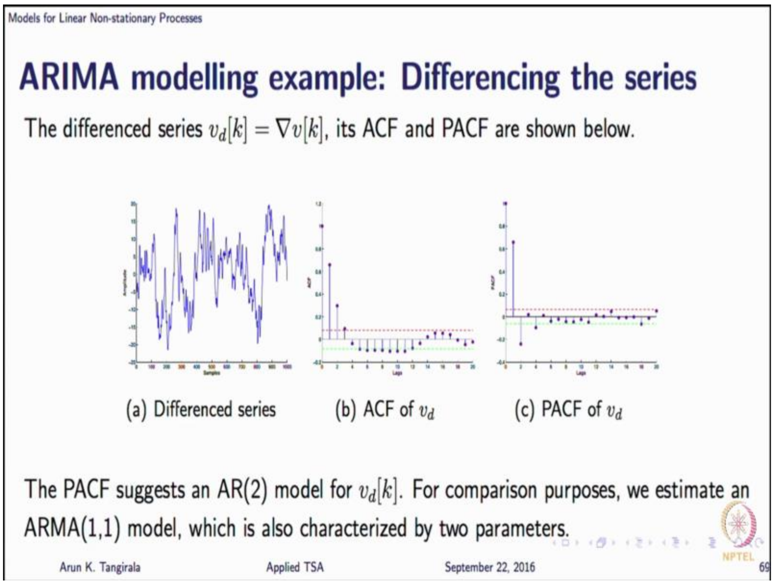


(Refer Slide Time: 14:00)



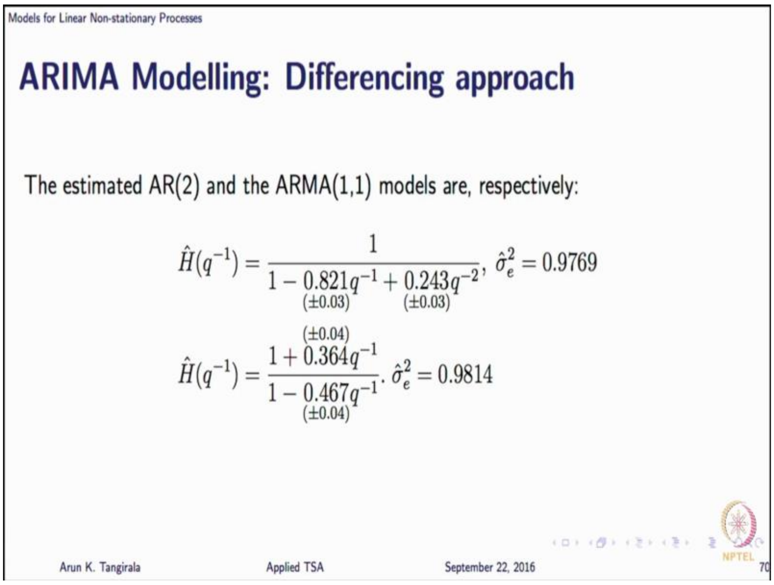
So, we fit an AR 3 model and here you have they AR three model, let us look at the residual first and then back. The residuals of their three models suggest that the residual survive. So, we can go back and look at the errors in parameter estimates. Here are the errors in parameter estimates. And there all fairly low, low to the extent that all the parameters that I have estimates that I have or significant statistically. Sigma square e is given to you. I am not told you what the underline process is here straight away I started with the series, so we will keep that as suspense. That because that is what is close to reality. Now I fit an AR model although as ARMA model I fit an AR model. So, this model is satisfactory in almost all respects I have not cross validated, but this is one model that I have.

(Refer Slide Time: 14:43)



The other option is to difference series because I saw let us say the unit the formal test gave me let us say that it say is that there is evidence and I went ahead and a difference series and you can see now the slowly decaying ACF has been replaced by an ACF of a typical stationary process and PACF suggest that I can fit what kind of a model to the different series second order I have or I can fit an MA also, either way is fine.

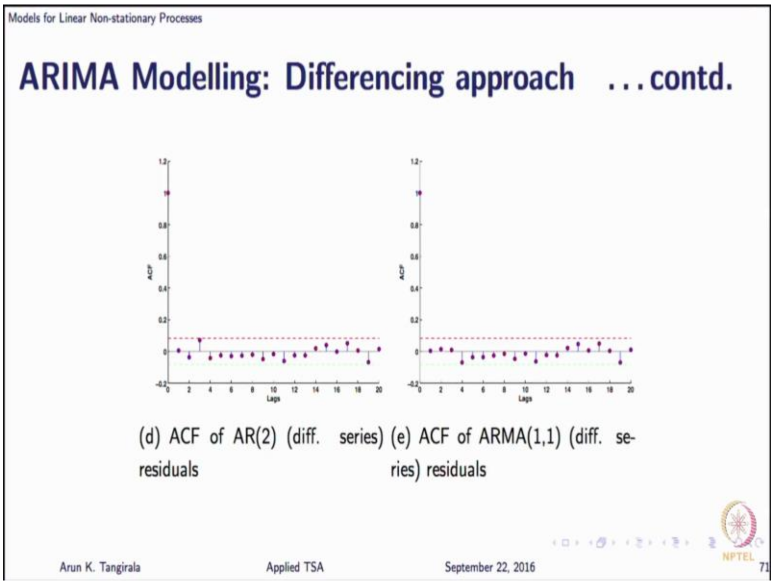
(Refer Slide Time: 15:17)



So, let us we fit an AR 2 model and in the end I have also what I have done is a fit an AR 2 model I have also fit an ARMA model both as the same number of parameters. Now which one do I choose? These are the models for the differenced series, please remember that. Which one do I choose? Now you can apply AIC to these two models you can look at one such criteria, but we will pick we I am not applying in AIC just for the sake of illustration we will select AR 2.

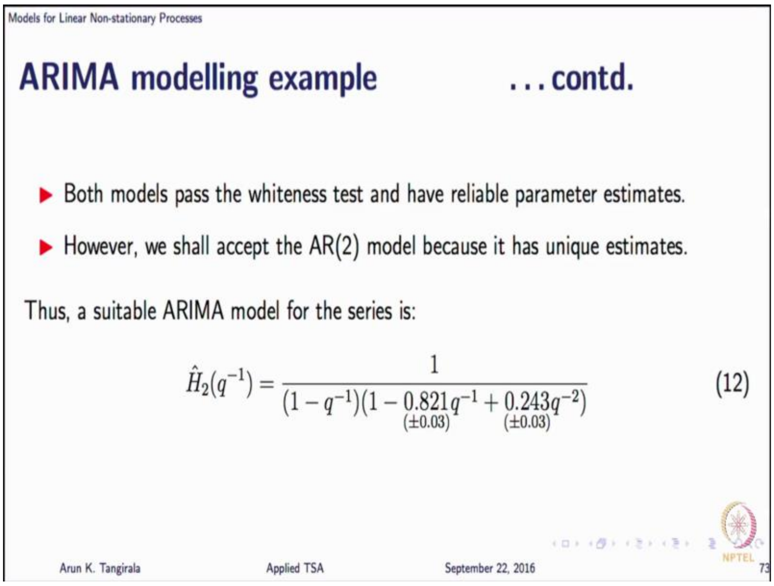
Primarily, because you obtain unique estimates that are one reason the other reason is you can see sigma square is lower. Always your estimate of sigma square e is a reflection of how well you have managed to predict or capture the predictable portion. Why, because sigma square e is the variance of the unexplained part always, because its variance of what you cannot predict. And with AR 2 I have sigma square e lower then ARMA 1 1 which means AR 2 has done a better job of explaining the give series. So, we stick to AR 2.

(Refer Slide Time: 16:27)



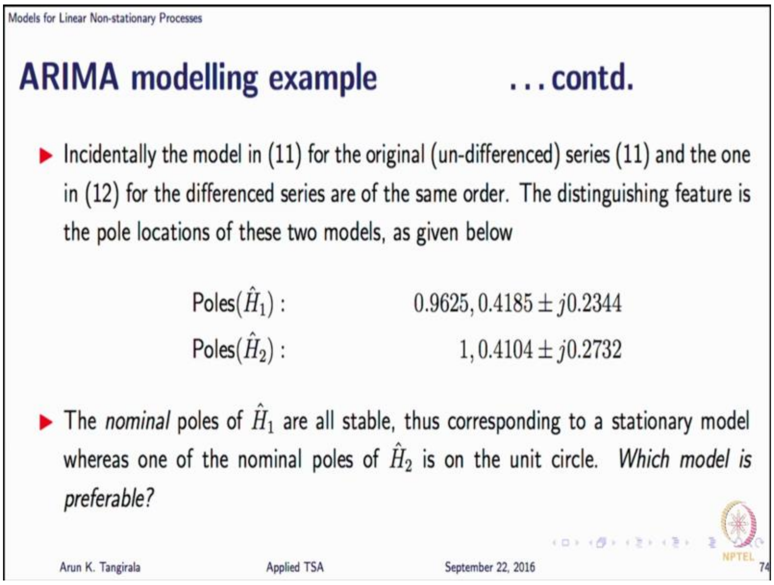
And I am just showing you the ACF and PCF of the residual from the AR 2 series.

(Refer Slide Time: 16:32)



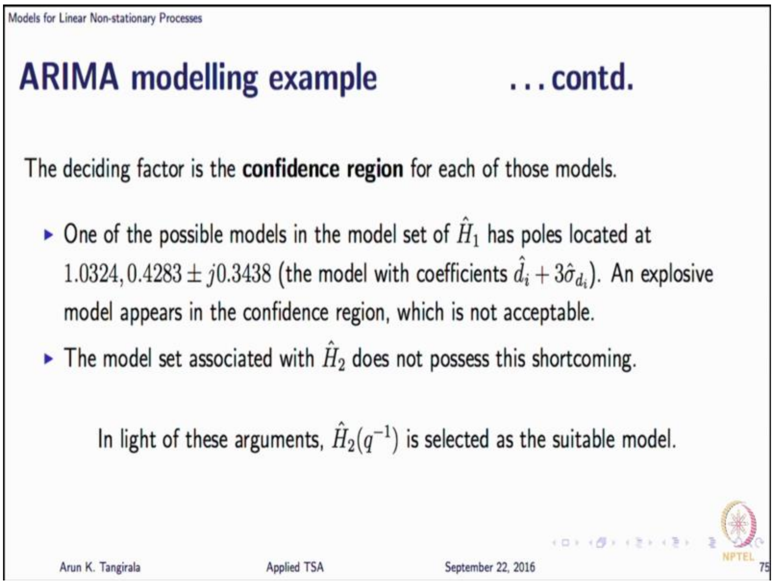
Therefore, the final model that we have with the differencing approach is an AR 3, but one of the poles fixed to unit circle other poles are here. So, now the question is I have two models; I had an AR 3 for the original series and AR 2 for the different series which one is better. And that is the final thing.

(Refer Slide Time: 16:53)



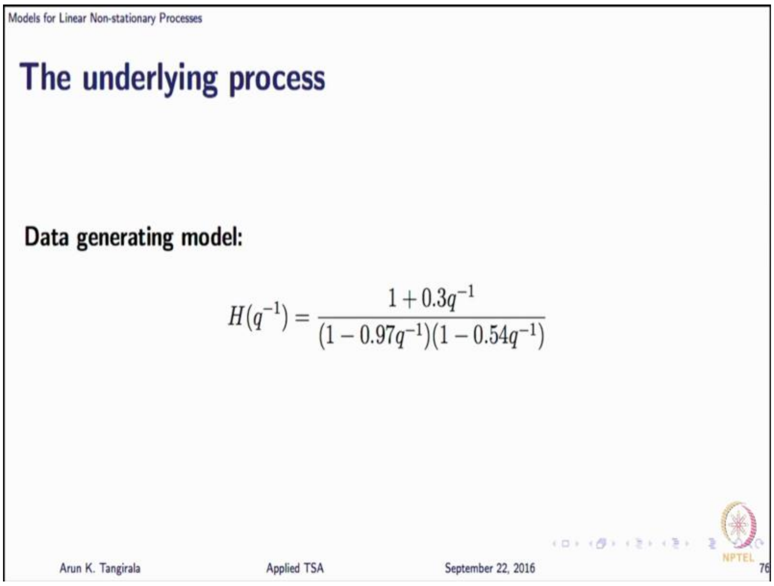
Now it turns out that when I look at the poles of AR 3 model you can see that all the poles are well within the unit circle no problem stationary, whereas obviously with the poles of the integrating this AR I model one pole is at the unit circle. Now, which one do I pick? We will pick the first one because the point estimate is stationary.

(Refer Slide Time: 17:20)



What about the confidence regions? Now it turns out that if you work out, there are many possibilities now. See what you have to see there are three poles for the AR 3 model and likewise you will have confidence region for each of those three poles. It turns out that if you pick one of the combinations, it turns out to be non stationary. And once again using the arguments that we used earlier we reject this AR 3 model in favour of the integrated process that integrated approach that is the AR 2 and integrated model. So, the final model that we choose is this one here; the second model H 2.

(Refer Slide Time: 18:02)

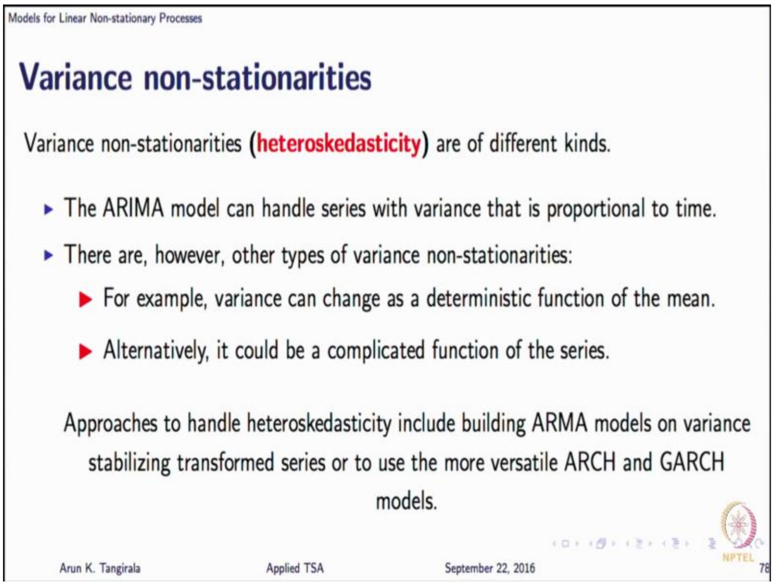


Although, the data generating process does not have an integrating effect. It has very close to integrating effect, but estimation errors have forced us; estimation aspects of forced us to live with an integrated model.

But again it depends on the method that you use to estimate these parameters. If your method is very good at estimating poles very close to unit circle, they you may not have to difference series. If it is able to estimate poles with really high precision then you are ok, but in this case I have use the method that perhaps as not done a great job. So, you should remember this thing; I will just take couple of minutes on non stationarities and then we will adjourn.

So, that kind of concludes the integrating effects thing. Now the last non stationarity that I wanted to very briefly discuss is that of heteroskedasticity.

(Refer Slide Time: 18:51)



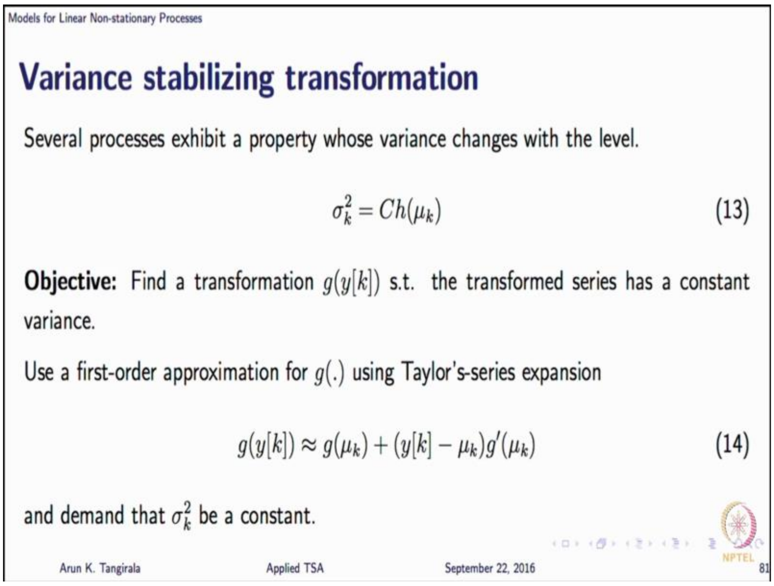
It is a big world it takes some time to pronounce it, but that means it also tells you it is a complicated process. So, the heteroskedasticity term is used for processes that have changing variance that whose variance changes with time.

You may say well the integrating process also has that; we have also already shown that the variance of the random of process changes with time. And yes that is also heteroskedastic process, but there is something else about the integrating process that required special attention. On the other hand you do not have to have integrating a fix you the mean may be changing in a different way. And the variance therefore may be changing in a different way.

So, the class of process whose variance changes with time are called Heteroskedastic processes and you will find numerous such processes and econometrics, and to a certain extend an engineering and in nature for sure. So, there are two possibilities: variance can be a deterministic function of the mean, that in many processes the variance may be coupled with the mean. For example, if I take a Gaussian process is variance coupled with the mean; no, what about the pause on process or what about the pause on process do you thing the variance and mean are coupled? Yes.

So, non Gaussian processes can have that kind of a characteristic or it could be a very complicated function of the series. Now generally one has this two approaches: one is the variance stabilizing transformation and the other is to a build what are known as ARCH models. Either you work with transformed the series to get rid of the variance changing with time or you work with generalized set of model another class of models known as the ARCH or GARCH models. ARCH models have said earlier auto regressive conditionally heteroskedastic models and that is a lot; that means, we will not discuss this in the course. But you should be aware and there very very popular in econometrics and GARCH is a generalization of that. We will not go in to that; I just want to quickly go over the variance stabilizing transformations.

(Refer Slide Time: 21:07)

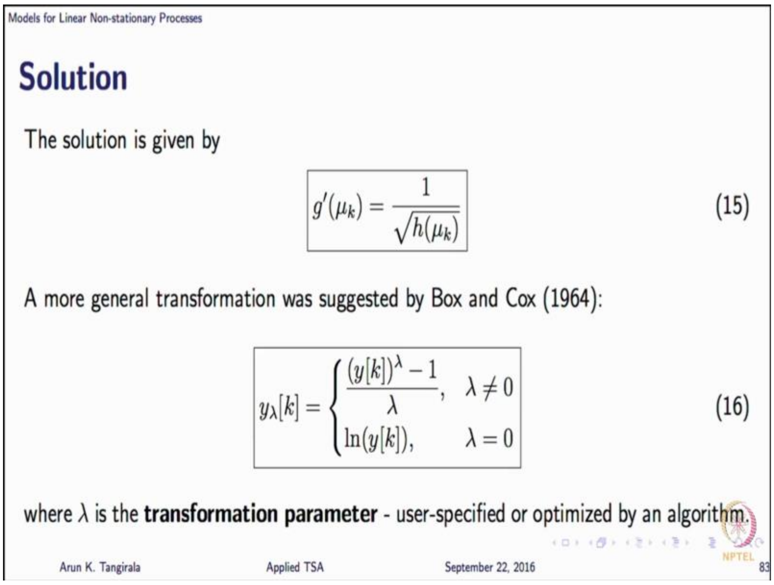


So, assume that your variance is some function h of the mean at any time and what you want to do is you want to find the transformation that such that the transform series as a constant variance. So, that is a goal here. So let us say the transformation is g, what we do is to figure out which transformation will get read of the way dependence of variance on the mean. We construct a first order approximation using Taylor's series expansion. And essentially right g of y k as construct the first order approximation of g of y k around the local mean. And that is what you have an equation 14. And what we want is the variance of g of y k to be independent of the mean or independent of time it should be constant.

Now, you can evaluate the variance of g of y k. On the right hand side you have variance of two terms: some of two terms the first term is a constant g of mu k at k it is constant.

So, you are left with only the variance of the second one. Remember adding constants do not change the variance. So, when you work out the math it is a very very simple math.

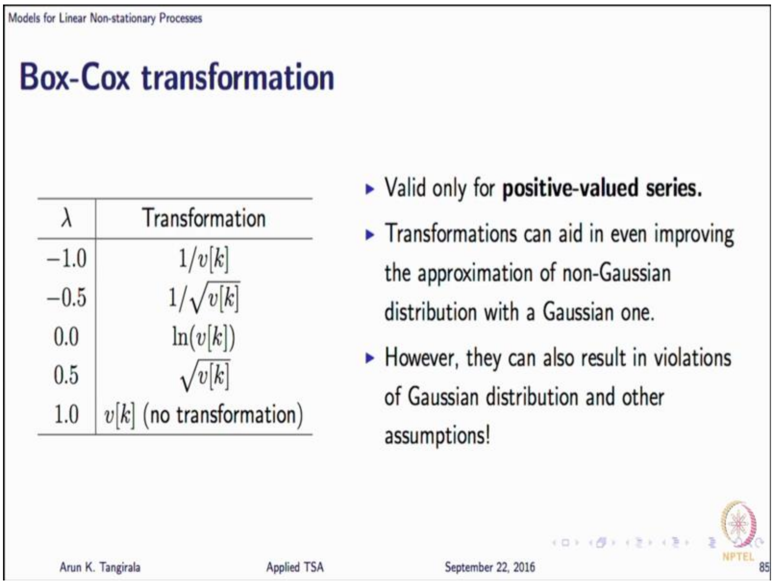
(Refer Slide Time: 22:12)



You find that the resulting transformation whatever transformation you are planning to use should satisfy this result. If you no h which is the dependence of various on mean it says you can figure out what the transformation is, and that is given in equation 14 here.

And from this came about the more general class of transformations by again Box and Cox you can see box appears everywhere and also the equation is a Box there.

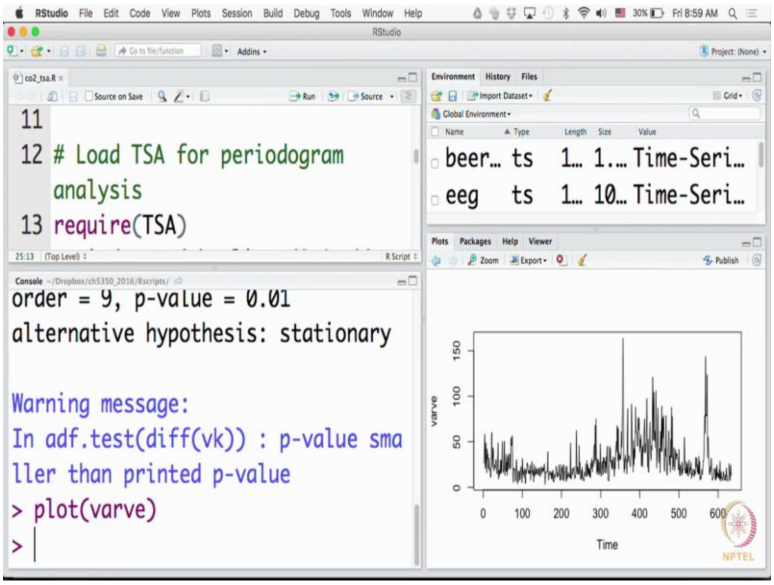
(Refer Slide Time: 22:55)



So the general set of transformations that where suggested by Box and Cox they are known as Box-Cox transformations. And there is a parameter lambda to control the nature of transformation that you do. For example, if lambda is 1 you are simply inverting the series 1 over v k or and if it is minus 0.5 your actually inverting the square root of the series. And if lambda is 0 then you will be taking a logarithmic my transformation.

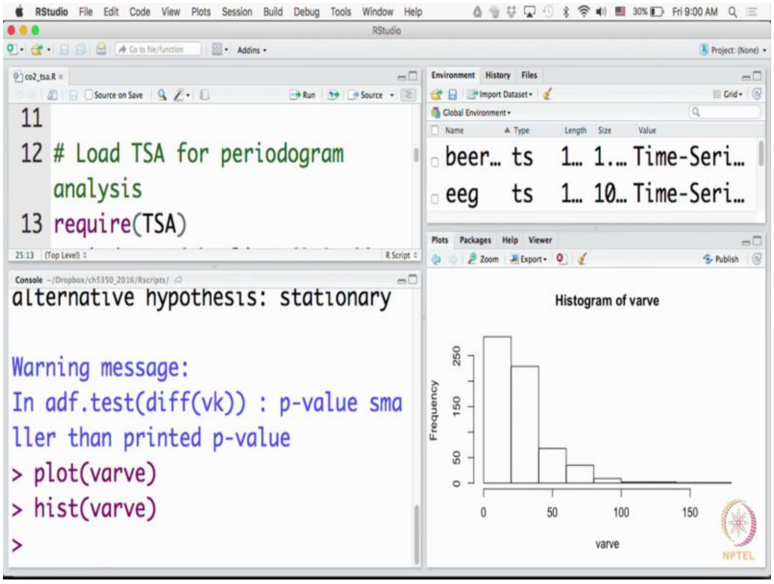
Remember logarithmic transformations are use for example, which have a growth type of characteristic to the process. So, generally you should remember that these kinds of transformations one are valid only for positive valued series. And this transformation can aid even improving this Gaussian approximations or Gaussianity that we are working with. Remember we said pause on process can have or even in a exponential process can have mean and variance coupled. The movement you would take this transformation your improving the distributional characteristics.

(Refer Slide Time: 23:55)



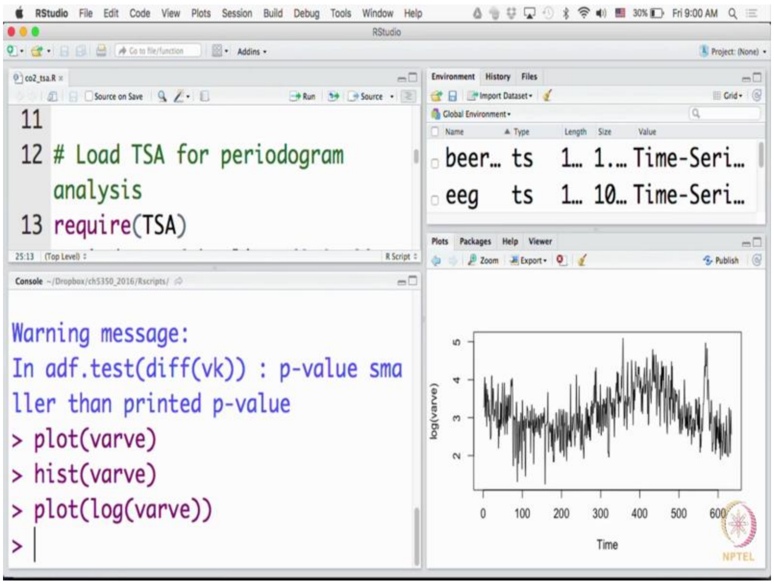
So, I will just conclude with an example from (Refer Time: 24:53), there is a series called varve which is essentially the deposition of the sand and silt and so on in a certain location (Refer Time: 25:03) it is an US not in UK. Over a certain period of time if you were to plot the varve series over so many years, you can see that there are locally variances exploding; the variances are actually changing with time, but these not an explosive series as such. So, its stationary process in that sense, but the variance is changing with time.

(Refer Slide Time: 24:32)



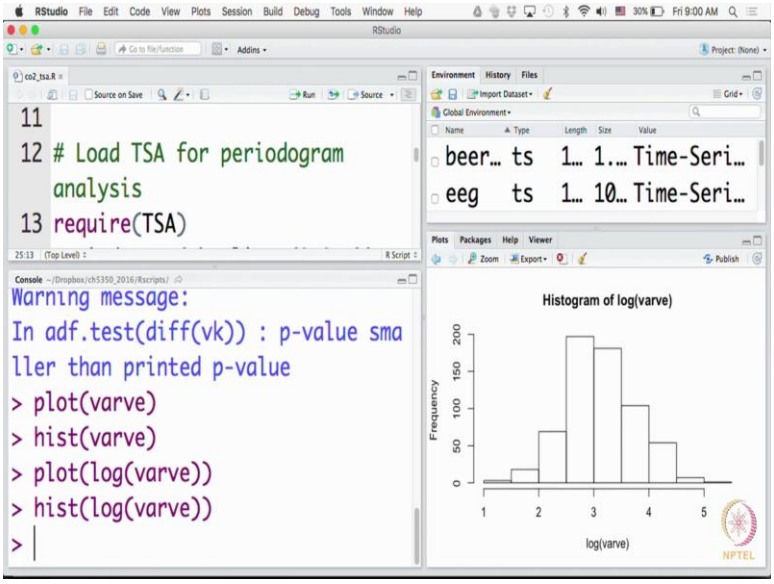
On the other hand, if you take, in fact if you look at the hist histogram of varve it has a Chi square kind of distribution it does not have a Gaussian distribution.

(Refer Slide Time: 24:40)



On the other hand if you take the log then you see that the variance as stabilized more or less over difference over difference segments of the series the variance is kind of uniform.

(Refer Slide Time: 24:50)



And finally, if you look at the histogram of logarithm it is Gaussian. So, you have to be careful also this means that if your series is Gaussian somehow and the variance is changing with time it is possible or whatever it is when you are transforming the series you may end up changing the distribution and therefore you have to be careful in working with theses transformations.

Anyway, so we will not go beyond this concludes the discussion on non stationary series. We will in the next class start off with Fourier series.