Time Series Modelling and Forecasting with Applications in R

Prof. Sudeep Bapat

Shailesh J. Mehta School of Management

Indian Institute of Technology Bombay

Week 06

Lecture 26: Persistent and Long-Memory Processes: Examples and Implications

Hello all, welcome to this course on time series modeling and forecasting using R. Now we have entered a new week, and there will be some very exciting and interesting topics to follow. Now, the whole idea we will discuss this week, or the focus of this week, will be another sort of model. So, we are transitioning from, let us say, simple ARMA models, ARIMA models, or SARIMA models. And then we are trying to capture some aspect of a time series called persistence. So, persistence or long memory.

So, as you can see very clearly in front of you, the title of today's lecture, especially for the first half of this week, will be persistent and long memory processes. And then, in short, we'll discuss one particular kind of model called the ARFIMA process. So, ARFIMA, right. So, again, right at the start, I will tell you the full form of this. Of course, we will discuss a lot of details about the ARFIMA process, the functionalities of this ARFIMA process, and use cases of ARFIMA in the subsequent lectures, right. Now, ARFIMA's full form is autoregressive fractionally integrated moving average.

So far, if you remember, we have discussed ARIMA models or, for that matter, SARIMA, etc. So, what was the ARIMA model? So, the ARIMA model was autoregressive integrated moving average. So, if you have a trend aspect in the dataset or if you have a trend aspect in the underlying process, then one can actually use an ARIMA model. Or, on top of that, if you have both trend and seasonality, then obviously the ARIMA model won't do the job.

Then you would actually shift to a SARIMA model. SARIMA is seasonal ARIMA. So, seasonal autoregressive integrated moving average. What would happen if you want to perceive some persistence in the underlying time series process, which you also call long memory? Now, again, one small point to note here is that both ARIMA and SARIMA do not preserve the long memory process or the long memory nature of the underlying process.

So, first, you can understand what do you mean by persistence? What do you mean by long memory? And then, eventually, probably in the subsequent sessions this week, we will talk more deeply about what exactly the ARFIMA process is and the properties of the ARFIMA process. Now, the first session this week that we are starting today will be more of a theoretical kind of session. So, by theoretical, I mean there will be a lot of definitions, some use cases, and some examples, right?

So, we will try to discuss or find out exactly where you can use this ARFIMA process and what exactly are the advantages of applying the ARFIMA process as compared to, let us say, ARIMA or SARIMA, etc., okay? Alright, so what exactly is the motivation? So, the first motivation is: what do you mean by persistence in time series? So, this term persistence is kind of important, and we come across it in disciplines. So, let us say finance, ecology, or psychology, etc.

So persistence in time series refers to the tendency of the past values to have a long lasting influence on the future values. Now again if you want you can pause the video and then reread this definition number of times as to what exactly persistence stands for. So persistence in time series refers to the tendency of the past values to have a long lasting influence on the future values. Okay, or in other words, it reflects the degree of memory a time series has, right? It reflects the degree of memory a time series has, that is, how long shocks or disturbances to the system affect the future observations, right?

Now, a couple of things, so probably I will explain this in a bit more detail. So, firstly, persistence is what? So, persistence is nothing but the tendency of the past values to have some effect or some long-lasting influence on the future value. So in a way you are saying that the future values are sort of dependent on the past values very very very significantly also. And the other term we are coining here is the degree of memory.

So it reflects the degree of memory imposed on the time series which also stands for that how long a particular shock or a particular disturbance to the system would affect the future observations. So a small example in this regard is that you have a stock price data over several days. So, let us say you have a daily stock price data and suddenly in between due to some black swan events, there might be some really high jumps, really sharp moves, which are also called as shocks. So, what are shocks? So, shocks are nothing but some external factors which are sort of driving the stock price to move very erratically.

So, let us say you have a lot of jumps suddenly and so on. So, once you get past that shock, if the influence of that shock is sort of persistent, Again, if you go by the English definition, persistent means something which stays for some time or quite long. So, if the influence of that particular sharp behavior in the stock price or the shock, in other terms, is sort of seen in the future also, then such tendencies are called persistent time series. Or this tendency is called persistence in the underlying time series.

So, this whole idea about why we want a different sort of model called ARFIMA is to sort of preserve this persistence in the time series, ok. Alright, now we actually have several characteristics of persistence. So, what do we mean by that? So, the first one is long-term impact, right. So, if a time series exhibits persistence, any shock or a shock to the system will have an influence that decays slowly over time rather than suddenly disappearing, ok.

So, probably this would be better explained through a kind of graph. So, let us say you have the stock price behavior like this, and then suddenly you see a shock. So, you have a very sharp behavior. And once the shock happens, then you can see that the effect of the shock does not disappear suddenly. You have a persistence kind of thing.

So this is exactly where the shock happened. And then the effect of the shock is such that even after the shock has happened, you can still see some fluctuations. So this is what the long-term impact is. So rather than the effect of the shock dying suddenly, it takes a while for the influence of the shock to decay. Now the second characteristic is autocorrelation decay.

So if you draw a simple ACF plot in such a situation, how would the ACF behave? So here, if there is persistence in the time series, the autocorrelations decay slowly as the lag increases. So the ACF values, or the underlying correlations, decay slowly as opposed to suddenly or quickly. As opposed to, let us say, some other model called an ARMA model or ARIMA model. So this contrasts with some of the short-memory processes.

So, by the way, ARMA, ARIMA, or SARIMA. So all these are short-memory processes. So these models are not capable of preserving the long-term impacts or the persistence in the time series. So if you have ACF plots of such models—ARMA or ARIMA, etc. There, the autocorrelations drop off more rapidly.

But if you are creating the ACF of, let us say, an ARFIMA model, then there will be some slow decay in the autocorrelations. And lastly, long versus short memory. So, what exactly do you mean by that? So, you have a thing called long memory, which is also referred to as persistent processes, right, such as those modeled by ARFIMA, which exhibit a slow hyperbolic decay in the ACF or the autocorrelation function. And exactly opposite to that, as we discussed a short while back.

So, you have some processes which are short-memory processes like ARMA, ARIMA, etc., which typically have autocorrelations that decay exponentially and quickly approach zero. So, again, probably this might be explained better by some plots. So, let us say I will draw two plots, and then both these would be ACF plots, right? So, let us say initially, and then these would be the hypothetical bands, let us say, right, or the confidence bands. So, let us say if you are modeling an ARMA process, and then this is the ACF for the ARMA process, then here one can actually see that the autocorrelations are decaying sort of exponentially.

So, after a while or after a very short while, you can see that all the correlations are sort of not significant anymore, and they are sort of approaching zero. But as opposed to that, if you have an ARFIMA ACF, then the ARFIMA ACF takes time to decay. So, you have a very slow decay here. So, the autocorrelations would decay, but then, since you have a persistent nature in the time series, it will take a while for the autocorrelations or the ACFs to sort of decay. So, the top ACF plot is, let us say, for ARMA, and then the bottom one could be for a particular ARFIMA.

So, this is the difference between why you should move from, let us say, something like ARMA or ARIMA to another model, which is ARFIMA. All right. Now, a few examples of persistence. So, can you give some practical use cases where persistence is important? Of course.

For example, the first one is the financial market. So, let us say in finance, persistence is observed in volatility or returns, right? So, if you talk about, let us say, returns of a stock or, you know, returns of any asset for that matter. So, how volatile are the returns? So, volatility means nothing but how fluctuating the returns are, right?

So, if the returns are fluctuating very highly, then we will call that the series is highly volatile. So, volatility is a measure of how fluctuating the underlying asset price or the underlying asset return is. So, if a stock price rises sharply, let us say, persistent trends might suggest that the price will continue to be influenced by the rise for a very long time, as opposed to reverting to a previous level quickly. And I think this is exactly what we talked about a while back, where we drew the last graph also. So, again, if you go back a slide, then this is exactly the example that we are talking about in the first situation.

So, let us say if the price suddenly jumps and then you have a shock here, it will take a while for the effect of the shock to subside rather than quickly returning to the actual levels. So, this is the nature of persistence in the underlying time series. A second example could be environmental data. So, let us say climate data like temperature and rainfall often exhibit persistence. Why? We will take up an example.

So, let us say a period of above-average rainfall might be followed by more frequent such periods or prolonged rains. Isn't it? So, this is something we actually observe on a day-to-day basis. Especially in the monsoon season in Mumbai, if you have a day with really high rainfall, then probably the next day you will again see a similar kind of rainfall.

It will not happen that one day you see really high rainfall and the next day you see none. This is not what we call persistence. Persistence means that if you have periods of above-average rainfall activity, such activity would take some time to subside. A third example could be in economics. So, let us say economic indicators like inflation or GDP growth rates might exhibit persistence, where economic shocks have long-lasting effects on growth trends.

So again the common thread in all the three examples is that if you have some really sharp activity then it will take a while to sort of come back to the actual levels. So that activity would have some persistence in the future also. Now importance of persistence. So why should we study the idea of persistence and then why is studying persistence sort of important? So this is a twofold kind of idea.

So the first one is forecasting. So in time series forecasting, identifying whether a process is persistent or helps to determine the model to use, right? Isn't it? So, if you know that there is some persistence in the underlying data, let us say stock price behavior or temperature data, rainfall data, whatever, then accordingly we can pick the correct model, right?

I will give you a very simple example in this regard. So, let us say if you want to forecast, let us say, above average rainfall for some reason. So, you have data observations where each observation gives you the amount of rainfall, but then you know for some reason that all these observations are above bridge. Okay. Then which model should you apply?

So, should you apply the ARMA technique, the ARIMA technique, or the ARFIMA technique? Right. So, this is the question. So, the correct model to apply would be ARFIMA because there will surely be some persistence in the underlying rainfall data. All right.

Hence, persistent series require models that account for long memory, like ARFIMA. Right. As they cannot be effectively modeled by short-memory models like ARMA. Okay. So, in this situation, you would rather go with an ARFIMA model as compared to something like ARMA or ARIMA.

So, the first idea is forecasting. So, why exactly should we study persistence to sort of pinpoint a particular model that could be used to preserve the underlying persistence in the time series? And the second angle is risk management. So, let us say often in financial markets, persistent volatility indicates that large fluctuations are likely to persist. So, if you have really high volatility, then there would be some persistence effect in the future, which can be critical for risk assessment and hedging strategies.

So, these are some financial terminologies. So, if anyone is familiar with these terms, they might have a better connection to this example regarding the importance of persistence in risk management. So, if you know for a fact that the underlying time series or the underlying data has some persistence, then accordingly you can fit a model, try to forecast, and mitigate the effect of risks. Alright, so now probably what we will do is delve deep into some of the use cases. So, persistence in financial markets.

So, the first one is stock market returns. So, consider a stock that experiences a sudden surge in price due to, let us say, a positive earnings report. So now, if the stock's price follows a persistent process, the effects of this information or the effects of this surge will last for a significant period, and future prices will remain elevated for a longer time compared to a stock with no persistence. So hopefully, by now, the idea of persistence would be clear. So again, if you want to graph this situation, how would the graph look?

So, let us say it would basically be similar to the graph we drew earlier, right? So, let us say you have a stock price that sort of behaves like that, and then assuming you have a trend also. Now, suddenly some information comes in, right? So, let us say a positive earnings report or some really positive sort of information. So, then what would happen?

So, there will be a jump, right? So, let us say the stock price moved there. Now, once the stock price moves there, the very next day one cannot assume that the stock would again come back to this level, isn't it? Right. So once the positive earnings report is out, then there's a huge jump in the stock price, and such a thing may not happen immediately.

So what might happen is that after this huge jump, there would be some persistence. So the stock price might revolve around this range and eventually decide its future course. OK, so this is the idea of persistence. So once you see some shocks or observe some really high fluctuations in the data set, it will take a while to come out of that influence and then return to the actual levels, okay? So the first example is stock market returns, right? And now we will talk about implications.

So what are the implications of the previous example? So, if returns are persistent—if the returns of a stock price are persistent—a positive or negative return is likely to be followed by another positive or negative return for a while. And this is exactly what we saw. So, let us say if you want to draw the graph of returns, and let us say this is the zero line, then how do the returns behave? So, returns could be either positive or negative.

Let's say, due to the fact that the positive earnings report has come out in the market, there would be a prolonged period of positive returns, and there could be some persistence in the series. So, all the following returns may be positive for a while. So, this is the idea, or this is the implication of the persistent nature. So, this can suggest a momentum effect where trends continue for some time. Conversely, in a market that does not show persistence—so, let us say if you have some dataset or some market situation where persistence is not there, for example, short memory—then price movements may quickly reverse, as seen in any mean-reverting process.

So, can you sort of plot this tendency now? Of course. So, let us say if there is no persistence, what would happen? Now again, if you want to chart the returns data, let us say this is the behavior of the returns. Now, let us say due to some positive earnings report, the returns are positive for a while, but again, it will not take long for the returns to, let us say, come back to 0 and maybe oscillate around 0.

So, the first example is when persistence is there because here you can see that you have a group where positive returns are together showing persistence. But here, even if the positive earnings report was there when the price fluctuated or the returns became positive or highly positive and then sort of continued for a very, very short while, again, it sort of reverted back to the mean, right? So, such processes are called non-persistent processes or mean-reverting processes because, let us say, the mean is 0 (assuming initially the mean was 0), but suddenly the mean shifted. But it did not take long for the series to again come back to its mean, which is 0. Now, persistence in financial markets—another example is volatility clustering.

So, in finance, you have this term called volatility clustering. So, what do you mean by that? So, it means that a sudden increase in market volatility, such as, let us say, during a financial crisis, tends to be followed by a prolonged period of high volatility. So, let us say if you have really high volatility, then there will be some period in the future where one can observe high volatility as well. So, this sort of tendency is called volatility clustering.

So, again, if you want to draw the graph of that, you get a similar kind of picture. So, let us say initially the volatility is like that. Now, again, let us say the volatility shifts for some reason, and then in this period, you observe the same volatility. So, you have two clusters, right? So, one cluster is here, and then the other volatile cluster is right here.

So, such tendencies are called volatility clustering. Or, in other words, large price movements—let us say increases or decreases—are followed by more large movements, while small movements tend to follow small ones. So, all these ideas point to just one argument or one idea, which is called persistence, basically. Now, again, the implication of the prior example is what? So, let us say volatility persistence is crucial for risk management.

So, if you know that there is some persistence in the underlying volatility, then can you sort of curb some of the risk measures there? The answer is yes. So, a market with persistent volatility implies that periods of high risk—or high risk means what? So, high risk means the market is really volatile. So, large price swings are not isolated events but are likely to last for quite some time in the future.

And then this might require some adjustments in risk models. So, if somebody knows that you have some persistence in the underlying volatility and you have really volatile groups in the future as well, and there is some persistence in the volatility nature in the future also, then I can accordingly act as per how my risk management should be, right, and then I can accordingly adjust that. And volatility clustering also has an impact on option pricing, right. So, this behavior introduces several complexities in accurately pricing options, right.

So, again, the idea is: can you sort of capture or can you propose an accurate kind of model, such as ARFIMA or some extension of ARFIMA, to sort of capture the volatility persistence? Now, one more example could be seen in interest rates, right? So, let us say interest rates often exhibit persistence, particularly in long-term yields. What do you mean by that? So, a small example again.

So, let us say if the central bank raises the interest rates, right? If the central bank raises the interest rates, the effects of this policy change may persist for a long time, influencing borrowing costs, savings, and investment decisions over extended periods, right? So again, a similar kind of example as the prior two that we saw. So, what is going on here? So, let us say if the central bank raises the interest rates, its effect won't die down immediately.

This is the point. So, the effect of raised interest rates would linger for some time in the future also, which is exactly called persistence. And what could be its implications? So, let's say persistent interest rates might suggest that current levels will remain influential for future rates also. So, if you have, let's say, the central bank raises the interest rates at some point, then that particular level to wherever the raise has been would remain influential for future rates also.

So, all the decision-making, all the risk management would kind of depend on that particular heightened level from now on. And then, this is critical for bond pricing, and let us say bonds are very sensitive to changes in interest rates. So, if you know that there will be some news of increasing the interest rates suddenly, then accordingly, I can adjust my risk management, and accordingly, I can put forward some models to, let us say, price bonds, etc. So, long-memory behavioral interest rates, especially long-term rates like, let's say, the 10-year Government of India bond, right? So, all these are long-term rates, right?

I mean, once the government decides on the interest rate for, let's say, a 10-year Government of India bond, then it's very rare that there will be any changes in the interest rates, right? So, long-memory behavior in interest rates can complicate forecasting and require more sophisticated models that account for persistence, right. So, if somebody knows that there will be some persistence in any of the practical aspects of whatever underlying data you are studying in the future, then this can actually complicate forecasting, and one would want a very accurate sort of model to capture the persistence, okay. Now, how about persistence in environmental data? So, if you shift from, let us say, a financial framework to, let us say, a different area, which is environmental data.

So, one very easy example here would be temperature persistence. So, temperature persistence is called global warming, right. So, the other name for temperature persistence is nothing but global warming, of which all of us are sort of aware now, right. What do you mean by that? That the global average temperature has shown persistent increases over the past century due to, let us say, climate change, right.

So, this persistence is caused by the continuous accumulation of greenhouse gases like, let us say, CO_2, carbon dioxide, in the atmosphere, leading to a warming effect. And this is exactly what we mean by persistence. So, once you see a very significant increase in temperatures, let us say, in a particular year, then that effect won't die down very quickly. This is the whole point. So, if you see a heightened sort of temperature on one particular day or in any particular month, then its effect would sort of go deep into the future also due to persistence.

What could be its implications? So, let us say temperature anomalies from year to year exhibit persistence. Where higher than average temperatures in one year are often followed by higher temperatures in subsequent years. And you have some underlying effects also. So, you have some temperature effects where if you have higher temperatures in a particular year, which are sort of above average, let us say, then the very next year you also sort of observe the similar tendency.

So, again like I say it would not happen that in one particular year you are seeing higher than average temperatures and then suddenly in the next year let us say starting from Jan onwards you are seeing below average temperature. So, such tendencies are not possible. And this is an important implication of persistence in let us say temperature data. Okay. So, persistent temperature rises can affect ecosystems, weather patterns and human health over longer periods.

Okay. So, persistence be it in finance or let us say environmental data requires some attention. So, this is the whole point. So, we cannot simply try to model that using let us say some ARMA processes or you know ARIMA processes, SARIMA processes. We require to capture the underlying persistence as well.

All right. So persistence in environmental data—one more example could be, let us say, precipitation and drought persistence. So droughts often exhibit persistence, where below-average precipitation or below-average rainfall continues for extended periods. And this can be due to some interaction between atmospheric circulation patterns and land surface conditions, whatever. Now, again, the whole idea in both the environmental aspects or environmental examples is that if you are seeing some heightened temperatures and, let us say, drought-like situations, then such situations might take a while to sort of come back to their actual state or come back to their standard state.

Now, what would be drought implications? So, regions like California or Australia frequently experience drought persistence. So, where you see dry years followed by more dry years, leading to water shortages and environmental degradation. So, long-lasting droughts can severely impact agriculture, water resources, and ecosystems, with long recovery periods required once the drought ends. Now, the last slide for today would be what we call anti-persistent time series.

So, anti-persistent time series is any time series where you do not observe any persistence. Or the other name for this is mean-reverting. So we have seen one example before also. So anti-persistent or mean-reverting time series means that an increase in the time series will be followed by a decrease and vice versa. So, let us say in one year you are seeing increased temperatures.

The immediate next year, you will see decreased temperatures. So, this is what anti-persistent behavior is. So, sequences in which high values are often followed by low values and vice versa. So, sequences where low values are followed by high values, demonstrating a tendency to revert or oscillate between the extremes. So, once the series reaches one of the extremes, it would have a tendency to revert back and then attain the other extreme.

So, maximum to minimum or minimum to maximum, etc. So, such behavior is characterized by a negative autocorrelation structure, meaning that periods of high or low values tend to be followed by periods of low or high values, respectively. So, of course, there will be some negative sort of a So, again, the last thing we will do for today is I will try to graph this tendency. So, let us say again, if you want to draw the ACF plot for some reason, right?

So, this is the ACF plot of an anti-persistent time series, and again, let us say these are the bands. So, in the first state, if you observe positive correlation, then in the next state, the correlation might be, let us say, something like negative. So, something like that, then again positive correlation, right, and again negative correlation. So, here you might want to extend the axis to the negative side also. So, this is the positive side, and this is the negative side.

So, here, can you see a negative correlation? Once you have high temperatures, then the subsequent year or the subsequent time point, you will see low temperatures, right? So, the correlation between the high observation and the low would be kind of oscillating, okay? So, again, in the next session, we will talk more deeply about the ARFIMA process. So, how exactly can you sort of make use of the ARFIMA process to capture the persistence in any underlying time series, okay?

Thank you.