**Time Series Modelling and Forecasting with Applications in R**

**Prof. Sudeep Bapat**

**Shailesh J. Mehta School of Management**

**Indian Institute of Technology Bombay**

**Week 10**

**Lecture 47: ARCH Models**

Hello all, welcome to this course on time series modeling and forecasting using R. Now, again, if you remember where we stopped in the last lecture, the focus area of this entire week will be on modeling volatile markets or how to capture or present models to capture the changing variance aspect of a time series or how to model volatility. Now, again, in the same spirit as the last lecture, if you remember, we discussed kinds of variances, ,tional variance, unconditional variance, and then we talked about different kinds of volatilities. So, before we start with anything new in today's session, we will try to elaborate a bit more on the very first kind of volatility we discussed, which is called HV or historical volatility. So, we will not go into detail on each and every kind of volatility.

So, again, if you remember, we talked about four such kinds. So, historical, realized, and then you had two more kinds of volatility seen in the markets: implied volatility and volatility clustering. So, these are the four kinds of volatility we discussed. But the very first kind of volatility is called historical volatility. We will try to elaborate on that a bit more.

So, historical volatility, or HV in short, measures the variability of an asset's returns over a specific period in the past. Rather, historical volatility quantifies the dispersion of an asset's past returns and is typically expressed as the annualized standard deviation of the daily returns. So, as discussed even in the last session. How do you find out historical volatility, or what do you mean by HV? HV is nothing but a very pure measure of the standard deviation of past asset returns. So, again, historical volatility quantifies the dispersion of an asset's past returns and is typically expressed as the annualized standard deviation of the daily returns.

So, it reflects the uncertainty or risk associated with the assets price movements during a given historical period. So, let us say if you are sitting today and then you want to focus or you want to find out or get an idea about how the variance has been behaving or variance has been changing in the recent past. So, for that you can actually go ahead and then try to find out the value of the HV for that. So, now what exactly is the interpretation of HV or interpretation of historical volatility? So, the first is high HV.

So, high historical volatility indicates large price swings and a higher level of risk. So, there would be large price swings and there would also be higher levels of risk and this is a very very typical for speculative stocks or cryptocurrencies for anybody who wants to speculate as to what would happen in the future in a particular stock or a cryptocurrency, then you have to also encounter a focus on such large price swings or the higher level of risk which is associated. At the same time what do you mean by low HV? So, low HV suggests smaller price fluctuations and a lower risk often seen in stable assets for example, like government bonds etcetera.

So, whenever if you have any stable asset like a government bond or something like that then people do not assume let us say really highly volatile structures then you can actually assume some lower volatility. So, suggest some smaller price fluctuations and lower risk ok. And lastly, changes in HV. So, changes in the historical volatility. So sudden spikes or drops can signal some market uncertainty, upcoming events or shifts in the investor sentiment.

So, the interpretation of HV is that one can actually encounter high HV, low HV, or how you can sort of capture some changes in HV. So, changes in HV could be captured using sudden spikes or drops in the signal, which can be due to some market uncertainty, upcoming news, upcoming events, or simply some shifts in investor sentiment. Now, thirdly, we will talk about what exactly the limitations of historical volatility are. So, firstly, HV is a backward-looking measure. Now, what do you mean by backward-looking?

So, backward-looking means that HV only reflects past price behavior and does not account for any future events that may impact volatility. Isn't it? So, since you are only focusing on capturing the standard deviation of past returns or past data, HV is always a backward-looking measure. So, it won't have the tendency to, let's say, forecast into the future. Second, HV is highly sensitive to the time period.

So, the length of the historical window chosen—let's say 20 days or 100 days—can significantly impact the calculated HV. So, are you averaging over the last 20 days and then finding the variance of that, or are you averaging over the last 100 days and finding the variance of that, right? So, these two can be highly distinct. Third, outliers and noise. So, extreme price movements can skew HV, making it less representative of typical asset behavior.

So, extreme price movements can skew the historical volatility, making it less representative of the typical asset behavior. And lastly, it ignores some structural changes. So, HV assumes that past market conditions are similar to the current or future conditions, which obviously may not hold during regime changes or crises, etc. So, again, this is a clear limitation of HV because HV sort of ignores any structural changes. So, let us say if you have some regime change or there is a sudden crisis, for example, COVID.

So, all these ideas would kind of impact the HV calculation very heavily because HV only focuses on the past data. So, if a sudden event comes in, or a sudden regime change comes in, or a sudden crisis comes in, then HV does not account for or focus on any of these ideas. So, a couple of limitations that we have outlined about HV, and hence why one moves to a slightly different or rather improved measure of volatility. So, rather than simply finding out the historical volatility using HV, why not try to put forward a model to capture the underlying changing variance? This is exactly the essence of today's lecture, or maybe the next one or the one after that.

Okay, so lastly, a few practical examples where one can actually see some volatility happening. So, the first one is risk management in finance. So, a financial institution might model the volatility of a stock price to estimate the potential risk exposure in its trading portfolio and adjust accordingly. So, adjust the capital reserves accordingly. Okay.

Second could be in portfolio optimization. So, a portfolio manager might use historical volatility estimates to balance a portfolio of stocks, bonds, and commodities, ensuring that the overall portfolio volatility is within some acceptable limits. Third is forecasting future asset prices. Hedge funds might use volatility models to predict the likelihood of price jumps in the underlying asset, such as during earnings reports or macroeconomic announcements. Lastly, market sentiment and behavioral finance.

Investors use the VIX to adjust their portfolios based on perceived market risk. For instance, if the VIX is rising, they may reduce exposure to riskier assets like stocks and increase holdings in safe-haven assets like bonds and gold. So, all these adjustments could be easily made if one puts forward and tries to analyze and study the underlying volatility in the market. So, I think now we are in a position to move ahead from the HV because we have seen some limitations of HV, and now the idea is to put forward some actual models, right? And these are the exact two models we discussed in the last lecture.

First one is ARCH, second one is G-ARCH or generalized ARCH, okay. So, again, ARCH stands for autoregressive conditional heteroskedasticity, and G-ARCH stands for generalized ARCH. Basically, these two are very famous models, and of course, their extensions that we study later on to capture the volatility in the market. So, a small primer about volatility models. So, let us say YT is the log return of an asset at time t. So, YT is the actual time series we have, and what does YT contain?

So, YT measures nothing but the log return of the asset at time t. So, one can say that the series $y_t$ is either serially uncorrelated or one can actually assume some minor lower-order correlation. So, one can actually assume that $y_t$ is entirely serially uncorrelated or you actually have some minor lower-order correlations, but it is assumed to be a dependent series. So, $y_t$ is assumed to be a dependent series. Now, let capital $F_{t-1}$ be the information set up to time t-1.

$$E(y_t|F_{t-1}) = \mu_t$$

$$V(y_t|F_{t-1}) = \sigma_t^2$$

Then, in that case, the conditional mean and variance of $y_t$ given $F_{t-1}$ is given by the expected value of $y_t$ given $F_{t-1}$, which is nothing but $\mu_t$, and the variance of $y_t$ given $F_{t-1}$, which is nothing but $\sigma_t^2$. So, the conditional mean, which is the expectation of $y_t$ given $F_{t-1}$, is assumed to be $\mu_t$, and the conditional variance, which is the variance of $y_t$ given $F_{t-1}$, is assumed to be $\sigma_t^2$. And again, just to repeat, what exactly is this capital $F_{t-1}$? So, this capital $F_{t-1}$ is nothing but the information set, which is kind of exactly similar to the omega we saw in the last lecture. So, capital $F_{t-1}$ is nothing but the information set up to the time point t-1.

And using that information set, I can immediately write down the conditional mean and the conditional variance of the log return, which is $y_t$. Make sense? Now, what we will do—or rather, what we will assume—is that let $y_t$ follow a simple ARMA (P, Q) model. So, let $y_t$ follow a simple ARMA (P, Q) model, and we can write it down in a slightly alternative manner like this. So, $y_t$ is assumed to be exactly equal to $\mu_t$ plus $e_t$. So, what exactly is $e_t$? $e_t$ is nothing but the random error.

$$E(y_t|F_{t-1}) = \mu_t = \phi_0 + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j e_{t-j}$$

$$V(y_t|F_{t-1}) = \sigma_t^2$$

So, I will write down the error here, and what is mu t? So, mu t is nothing but the conditional mean we saw in the last slide. So, mu t is nothing but the expected value of y t given f t minus 1. So, this is mu t. So, if you add this conditional mean to a small random error, we actually get hold of an ARMA model and then rather an ARMA-PQ model.

So, we will start with this guy. So, this is nothing but the expected value of y t given f t minus 1, which is nothing but the conditional mean, which is exactly equal to mu t, and now this mu t can be expressed as an ARMA PQ model. So, what we are assuming is that mu t equals phi naught plus summation i going from 1 to p phi i into y t minus i plus a different summation where j goes from 1 to q theta j into e t minus j. So, we have this underlying ARMA structure here with orders of PQ. So, the first, the AR order is P, the MA order is Q, all the phi i's and all the theta j's are the coefficients.

And secondly, on top of that, we can actually assume that you have this conditional variance also. So, the variance of y t given f t minus 1 happens to be sigma t square. Thus, the conditional mean of y t given f t minus 1 evolves through the interaction of past observations and shocks. So, in general, we are trying to put forward a model on this conditional mean, and we are sort of assuming that the conditional mean follows an ARMA process. Similarly, rather than focusing only on the conditional mean, can't we also analyze and study the conditional variance?

The answer is yes. So, similarly, the conditional variances may also evolve through the magnitude of their past shocks and some of the past variability. So, the sigma t square component could also be analyzed in a similar manner. So, the whole idea boils down to: can we put forward a suitable model to capture the tendencies or the behavior of the sigma t square component? Thus, the conditional heteroscedastic models are concerned with the evolution of the sigma t square component.

So, how does the sigma t square component behave? How does the sigma t square component evolve over time? So, we require some conditional heteroscedastic models to answer such questions. So, the conditional heteroscedastic models are concerned with the evolution of sigma t square, the manner in which sigma t square evolves over time. And by the way, this is the exact definition of volatility.

So, volatility is nothing but the under root or rather the positive under root of sigma t square, which is nothing but similar to a standard deviation. So, sigma t, which is nothing but the positive square root of sigma t square, is called the volatility. And lastly, volatility measures the size of the errors made in modeling returns and other financial variables. So, we have been discussing the impact and the definition of volatility even from the last lecture, and this is just a small repetition here. But again, this might be a small motivation to all of you as to why we should actually put forward some conditional kind of changing variance models, right?

Because again, sigma t square evolves over time. So, would it not be a good idea to put forward a model on its behavior? The answer is yes. So, rather than only focusing on conditional means that we have been doing so far, why not also focus on the changing variance aspect of the time series, right? Now, the next thing is that conditional heteroscedastic models can be classified into two general categories.

So, this is a small classification when it comes to capturing conditional heteroscedastic models. So, conditional heteroscedastic models can be classified into two general categories. The first one uses a stochastic equation to describe sigma t square. So, one can put forward a simple stochastic equation to describe the behavior of sigma t square. For example, let us say stochastic volatility models, or rather SV models.

So, what one can do is create or have an equation where sigma t square is one of the components and some stochastic equation on the other side. Then, the second kind of method is to use an exact function to govern the evolution of sigma t square. And this would be entirely the focus of this week here. So, the first kind of model is ARCH, which is Autoregressive Conditional Heteroscedastic Model, or ARCH, or the Generalized Autoregressive Conditional Heteroscedastic Model, or GARCH. So, again, just to summarize, the first kind of technique that one can follow is to create a simple stochastic equation to describe sigma t square.

And such equations are called SV models. On the other hand, one can actually use an exact function to govern the evolution of sigma t squared. And in the same spirit, we can study the ARCH models or GARCH models, etc. Now, the actual model-building process. So, how does one go ahead and try to build either the ARCH model, the GARCH model, or some of their extensions?

So, building a volatility model for a time series consists of four important steps. So, the first step is obviously to specify a mean equation by testing for serial dependence in the data and, if necessary, building a time series model—let us say ARMA or ARIMA—for the time series to remove any linear dependence. So, obviously, before one can focus or account for the variance aspect, one has to tackle the mean aspect first. So, the first point is to specify a mean equation, and one can actually put forward a mean model—let us say ARMA, ARIMA, SARIMA, whatever. So, all these are examples of capturing the mean underlying the time series.

The second point is to use the residuals of this mean equation that we have found in the first step to test for ARCH effects. So, then we have to actually test whether the underlying series or the underlying residuals of the mean equation contain some ARCH effects or not. Now, again, what do you mean by that? So, we are more concerned with whether the residuals show some changing variance tendencies or not. So, ARCH effects mean that there has to be some changing variance or there has to be some heteroscedastic nature in the underlying residuals, OK.

So, one can actually go ahead and test that. So, try to put forward a mean equation and gather the residuals of that model. So, let us say you put forward an ARMA model or ARIMA model, right? So, once you fit the ARMA model, you will obviously get the residuals.

Use the residuals of the mean equation to test for the ARCH effects. So, we want to make sure whether the residuals of the mean equation contain some changing variance aspects or not. And then we have a bunch of ARCH tests for that. So, we will discuss a couple probably soon in the next lecture, but we have some formal tests to show that, okay? And now, once you show that you do have some ARCH effects, then specify a volatility model if ARCH effects are statistically significant and perform a joint estimation of the mean and volatility equations.

So, once you are through with the second point and once you confirm that the residuals do show some ARCH effects, then you can actually put forward a model and try to jointly estimate the mean and variance using a couple of equations. And lastly, as always, check that the fitted model carefully performs well and refine it if necessary. So, this is the diagnostic checking aspect. So, of course, the diagnostic checking is not something one can ignore here.

So, once you put forward any model, be it ARMA, ARCH, or whatever, after the modeling exercise has been done, one has to ensure that the assumptions are being met or not. So, for that, this diagnostic checking step is kind of very, very important. So, hopefully, this is clear: first, try to model the mean, gather the residuals from the mean, and test for the ARCH effect in the residuals. If yes, then put forward a stochastic model on the variance or put forward a volatility model if the ARCH effects are there. And once you model using either ARCH or GARCH or one of its extensions, then perform the final diagnostic checks.

So, now we will give a very brief introduction to the ARCH model. So, ARCH is called autoregressive conditional heteroscedasticity models, a class of statistical models which are widely used to model time series data with varying volatility over time. So, if you have to capture changing variance, then one can actually put forward an ARCH model. These models are particularly useful in financial markets where volatility is time-varying and often exhibits clustering. Now, again, we have mentioned a number of times what you mean by clustering: that is, large changes tend to be followed by large changes, and small changes tend to be followed by small changes.

So, the changing variance also occurs in clusters. By the way, this entire idea about the ARCH model was introduced by Robert Engle in 1982 to model the time-varying volatility in financial time series. The primary feature of ARCH models is that the variance of the error term, or rather the variance of the residuals, is modeled as a function of past error terms, allowing for some heteroscedasticity, which is nothing but non-constant variance. So, in a way, if you allow the errors, or rather let us say E_t, to kind of follow or depend on its own lags, let us say E_t-1, E_t-2. So, this is exactly, or rather one way of slightly bringing the volatility models into play, and this is the whole idea about ARCH models.

So, initially we will study very briefly what the ARCH model structure is all about. So, in the ARCH model, the shocks, which are ET, are serially uncorrelated but dependent. So, we require some assumptions here that ET is serially uncorrelated but dependent. Now, this dependence of ET can be described by a simple quadratic function of its own lagged values, and exactly how. So, in general, if you talk about the ARCH-M model.

So, the ARCH-M model means the order is M. So, the ARCH-M model is a two-fold kind of model, or you have multiple levels there. So, the first level is ET equals sigma T multiplied by epsilon T. Now, again, remember what ET is. ET are the shocks, or ET are the random errors, right? And we are basically saying that ET also depends on the variance. So, ET is nothing but sigma T multiplied by some other IID random variables.

So, these are completely random errors. So, epsilon T's are entirely random errors, and ET's are the shocks of the time series. So, this is the first level of the equation: ET equals sigma T multiplied by epsilon T. And then we also show how sigma T squared, or rather how the variance changes, okay. And then again, if you remember, you can pause the video here and try to answer my question: this kind of model structure is similar to what model structure?

So, again, my question to all of you is: if you observe this equation for a second. So, which underlying time series model can be described using this equation? And then again, the answer is AR, isn't it? So, this is nothing but the AR model. Can you see that?

So, you have the dependent variable or the actual series, which is being modeled on the left-hand side. The alphas are the coefficients. So, alpha naught, alpha 1, alpha 2 up to alpha m, and all these are the residuals, right? So, E t minus 1 square, E t minus 2 square up to E t minus m square, and hence the order of this AR model is m. Because how many terms in the past do we have?

We have M terms. Again, just to reiterate the level. So, the first level is concerned with E t. So, we are kind of assuming, or rather stating, that the error behaves as if it is equal to sigma t multiplied by some other random error, which is epsilon t. And the second level is we are introducing, or trying to model, the variance using an AR structure or an autoregressive structure in this manner. So, sigma t square is alpha 0 plus alpha 1 E t minus 1 square all the way up to alpha m E t minus m square.

$$e_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \cdots + \alpha_m e_{t-m}^2$$

So, of course, again, where epsilon t is a sequence of IID random variables with mean 0 and, let us say, variance is 1. Alpha 0 is positive, and each of the alpha i's are non-negative for every i. So, this is all about the ARCHM model, ok. This is all about the ARCHM model, and again, some properties or some assumptions we require. So, alpha i should satisfy some regularity conditions to ensure that the unconditional variance of E t is finite. So, again, we do not want the

variance of the errors to go to infinity. So, we require some restrictions on each of the alpha i's, ok. Secondly, epsilon t is assumed to follow the standard normal or, rather, a standardized Student's t-distribution or a generalized error distribution, etcetera. So, the final random error that you see in the equation, which is epsilon t, one can, as per the regular assumptions, assume a normal distribution or one can put forward a slightly different distribution also, let us say Student's T or GED, etc.

And lastly, under the ARCH framework, large shocks tend to be followed by another large shock. So, you can see some leverage effect even here. So, now again, if you boil down to the ARCH 1 model. So, again, if you remember the ARCH M equation, one can immediately write down the ARCH 1 model, right, because M is 1. So, basically, M equals 1 here, and then immediately we get this two-level model again.

$$e_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2$$

Here, $E(e_t) = E[E(e_t|F_{t-1})] = E[\sigma_t E(\epsilon_t)] = 0$

So, E_t is nothing but sigma_t epsilon_t, where my sigma_t square is nothing but alpha_0 plus alpha_1 into E_t minus 1 square, and then you stop here because, since you are fixing the order to be 1, you would not go beyond the first lag. And, of course, here these are some other assumptions or other conclusions or other corollaries. So, the expected value of E_t is the expectation of the expectation of E_t given F_t minus 1. So, this is using the tower property. So, I am not sure how many of you are aware of the tower property.

So, if you want to find out the unconditional mean of something where you already have access to the conditional structure, you can actually write it down this way. So, you can have two expectations. So, the expectation of the conditional mean. This is nothing but equal to the expectation of sigma_t into the expectation of epsilon_t, which happens to be 0. So, in a way, one can actually prove all these properties about the ARCH(1) model as well.

So, if you talk about the unconditional mean, the expectation of E_t has to be exactly equal to 0. And now, the second thing is the variance function. So, further, the variance of E_t is nothing but the expected value of E_t square, and again, I can write it down in a tower property kind of statement here. So, the expected value of the conditional variance, or rather E of E of E_t square given F_t minus 1, which is the expectation of this guy, which is nothing but alpha_naught plus alpha_1 expectation of E_t minus 1 square. But again, just for a second, if you observe here.

$$Var(e_t) = E(e_t^2) = E[E(e_t^2|F_{t-1})] = E(\alpha_0 + \alpha_1 e_{t-1}^2)$$

$$= \alpha_0 + \alpha_1 E(e_{t-1}^2) = \alpha_0 + \alpha_1 Var(e_{t-1})$$

So, the expectation of E t minus square is nothing but can be written down as variance and why is that? Because we are assuming E t to be a stationary process. So, each of these expectations—expectation of E t square, expectation of E t minus 1 square, or for that matter at any lag—have to be exactly equal. So, eventually we get this equation.

$$Var(e_t) = \frac{\alpha_0}{1 - \alpha_1} \geq 0, \;\; 0 \leq \alpha \leq 1$$

So, the variance of E t equals alpha naught plus alpha 1 into the variance of E t minus 1, and one can actually combine a few entities here and then conclude that the unconditional variance is nothing but alpha 0 divided by 1 minus alpha 1. So, this is just an attempt to prove two things. So, unconditional mean under the ARCH1 setting or ARCH1 model and then the unconditional variance. So, expectation of E T and then the variance of E T. And lastly, we will discuss very briefly some limitations of the ARCH model.

So, the model depends on the square of the previous shocks, and almost all of you must have seen this tendency in the earlier slide. So, the model actually depends on the square of the previous shocks. Hence, the positive and negative shocks have the same effects on volatility. In practice, the price of a financial asset responds differently to a positive or a negative shock. Secondly, it often overpredicts the volatility as it responds slowly to large isolated shocks to the return series.

So, this is again one limitation of the ARCH model—that the ARCH model often overpredicts the volatility, as it has a tendency of responding slowly to large isolated shocks to the return series. And thirdly, parameter explosion. So, what do you mean by that? So, if the number of parameters grows with the order of Q, right? So, the number of parameters grows with the order of Q, right?

Because, let us say, if you are putting forward an ARMA (P, Q) model initially to capture the mean, right? So, the number of parameters grows with the order of Q. So, if Q is really large, right? Or, for large Q, the model becomes complex and computationally expensive. So, again, these are some limitations as to why people move from an ARCH model to a generalized ARCH model. So, obviously, in the next lecture and the lecture after that, we will try to study the GARCH model and maybe some of the very famous extensions of ARCH and GARCH, which are again very, very, very applied.

But again, one final point before you close this session is that almost all the examples where one can actually put forward an ARCH model or a GARCH model can be seen abundantly in financial literature. So, if you want to capture financial volatility—say, in underlying markets like cryptocurrencies, options, or stocks— ARCH and GARCH are just tools to capture the volatility, and abundant examples are found in finance-related contexts.

Thank you.