# STAT 5170: Applied Time Series

*Course notes for part B of learning unit 7*

## Section 7B.1: Volatility models.

To this point, our discussion of time series analysis and time series models has focused on relationships involving the mean or conditional mean of time series, which, with the introduction of ideas such as autocorrelation and seasonality, can become quite complex. Mathematical treatment of the variance of a time series has been extremely simple. For a time series, $(x_t)$, our focus is now on its *conditional variance*,

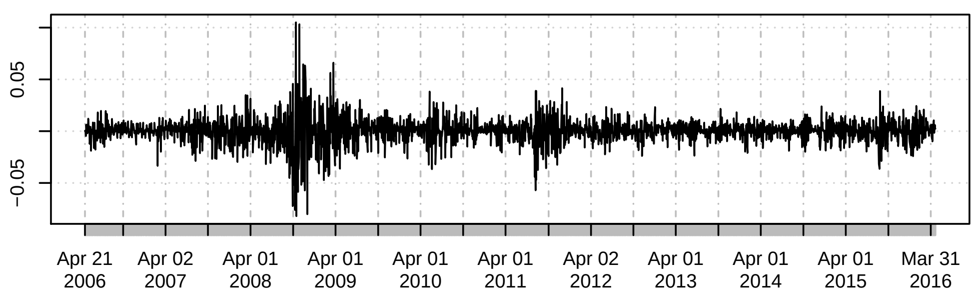
$$Var[x_t | x_{t-1}, x_{t-2}, \ldots].$$

Our aim is to cover a few classes of models for conditional variance that connect to ideas that we have used to model mean relationships, namely autoregression and moving-averaging.

An alternative terminology uses *volatility* in place of conditional variance. This arises from economic and financial time series analysis, where much of the development of volatility models has occurred. The following offers an example of the type of economic phenomenon the that volatility models attempt to address:

**Example: Dow Jones Industrial Average**

The figure below displays the time series of daily returns of the Dow Jones Industrial Average (DJIA) from April 20, 2006 to April 20, 2016. The period around the financial crisis of 2008 is one of high volatility.




□

In keeping with the focus on economic and financial context, our discussion of volatility models develops ideas on the *return values* of a time series. These are defined by the formula

$$r_t = \frac{x_t - x_{t-1}}{x_{t-1}},$$

which can be inverted to produce the representation of the original data in terms of the returns,

$$x_t = (1 + r_t)x_{t-1}.$$

In other words, if we are able to find a suitable model for the returns time series $(r_t)$, then it becomes possible to translate it to a suitable model for the original time series $(x_t)$, from which inferences and forecasts would be made. Oftentimes the returns time series is defined more simply as the first-difference of a logarithmic transformation, $r_t \approx \nabla \log x_t$. This makes sense given the mathematical approximation $\log(1+u) \approx u$ for $u \approx 0$, which suggests that "small" returns imply

$$\nabla \log x_t = \log x_t - \log x_{t-1} = \log \frac{x_t}{x_{t-1}} = \log(1+r_t) \approx r_t.$$

As a start to our discussion of volatility models, consider that ARMA models have constant volatility. This is quickly seen from the relationship

$$r_t = \phi_1 r_{t-1} + \cdots + \phi_p r_{t-p} + w_t + \theta_1 w_{t-1} + \cdots + \theta_q w_{t-q},$$

where $(r_t)$ is the returns time series and $(w_t)$ is white-noise with standard deviation $\sigma_w$. Assuming causality and invertibility, the volatility is

$$\begin{aligned} Var[r_t|r_{t-1}, r_{t-2}, \ldots] &= Var\left[w_t + \theta_1 w_{t-1} + \cdots + \theta_q w_{t-q}|r_{t-1}, r_{t-2}, \ldots\right] \\ &= Var\left[w_t|r_{t-1}, r_{t-2}, \ldots\right] = \sigma_w^2, \end{aligned}$$

wherein the first steps follows from the definition of the ARMA relationship, the second step follows from invertibility, which implies that each $w_{t-j}$ is a function of $r_{t-j}, r_{t-j-1}, \ldots$, and the third step follows from causality.

The first non-trivial class of volatility models to discuss is as follows.

**Definition:** The returns time series $(r_t)$ is an $ARCH(1)$ *autoregressive conditionally heteroscedastic* time series if

$$\begin{aligned} r_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 r_{t-1}^2, \end{aligned}$$

where $\sigma_t$, $\alpha_0$, and $\alpha_1$ are constant parameters and $(\epsilon_t)$ is a Gaussian white-noise time series with variance one, $Var[\epsilon_t] = 1$. □

The above definition implies that $\sigma_t^2$ is a conditional variance, and that an equivalent definition is

$$r_t|r_{t-1} \sim N(0, \alpha_0 + \alpha_1 r_{t-1}^2).$$

Simple variations of the $ARCH(1)$ model are commonly considered, such as relaxing the requirement of conditional Gaussianity; *e.g.*, the conditional value of $r_t$ given $r_{t-1}$ might instead be formulated from a scaled t distribution. The study of $ARCH(1)$ models also takes into account potential constraints on the parameters $\alpha_0$ and $\alpha_1$ to induce properties like stationarity and causality. This is discussed later. Relatedly, the restriction $\alpha_0 > 0$ and $\alpha_1 > 0$ is often imposed as sensible for modeling variance, which cannot be negative. The equation $r_t = \sigma_t \epsilon_t$ implies that the mean function of $(r_t)$ is zero.

An extension of the $ARCH(1)$ model is defined as follows.

**Definition:** The returns time series $(r_t)$ is an $ARCH(m)$ *autoregressive conditionally heteroscedastic* time series if

$$\begin{aligned} r_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 r_{t-1}^2 + \cdots + \alpha_m r_{t-m}^2, \end{aligned}$$

where $\sigma_t$, $\alpha_0, \alpha_1, \ldots, \alpha_m$ are a constant parameters that are possibly subject to constraints for stationarity, causality, and related properties, and $(\epsilon_t)$ is a Gaussian white-noise time series with variance one, $Var[\epsilon_t] = 1$. □

A related idea is expressed by the following.

**Definition:** The returns time series $(r_t)$ is an $GARCH(1,1)$ *generalized autoregressive conditionally heteroscedastic* time series if

$$\begin{aligned} r_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \end{aligned}$$

where $\sigma_t$, $\alpha_0$, $\alpha_1$, and $\beta_1$ are a constant parameters, possibly subject to constraints, and $(\epsilon_t)$ is a Gaussian white-noise time series with variance one, $Var[\epsilon_t] = 1$. □

Perhaps the most obvious extension of this idea is as follows.

**Definition:** The returns time series $(r_t)$ is an $GARCH(p,q)$ *generalized autoregressive conditionally heteroscedastic* time series if

$$\begin{aligned} r_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \sum_{j=1}^{p} \alpha_j r_{t-j}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, \end{aligned}$$

where $\sigma_t$, $\alpha_0, \alpha_1, \ldots, \alpha_p$ and $\beta_1, \ldots, \beta_q$ are a constant parameters, possibly subject to constraints, and $(\epsilon_t)$ is a Gaussian white-noise time series with variance one, $Var[\epsilon_t] = 1$. □

It would not be outrageous to have the sense that the above definitions are setting up a mathematical formulation of volatility models that parallel the development of autoregressive moving-average models. Such a parallel is readily observed by working with the time series of squared returns $(r_t^2)$. Consider, for example, that under an $ARCH(1)$ model for $(r_t)$, the squared-returns time series satisfies

$$\begin{aligned} r_t^2 &= \sigma_t^2 \epsilon_t^2 \\ \sigma_t^2 &= \alpha_0 + \alpha_1 r_{t-1}^2. \end{aligned}$$

By setting $v_t = \epsilon_t^2 - 1$ and re-writing $\sigma_t^2 \epsilon_t^2 = \sigma_t^2 + \sigma_t^2 v_t$, then substituting $\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2$ for the first term, we see that a squared-returns value is

$$r_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \sigma_t^2 v_t.$$

Because the $(\epsilon_t)$ time series is white noise, so too is $(v_t)$ an independent and identically distributed time series. Moreover, the property $\epsilon_t \sim N(0,1)$ implies that each $\epsilon_t^2 \sim \chi_1^2$, from which it follows that $E[\epsilon_t^2] = 1$ and $Var[\epsilon_t^2] = 2$, hence $E[v_t] = 0$ and $Var[v_t] = 2$. In other words, $(v_t)$ is a white-noise time series, although it is not a Gaussian white-noise time series. This sets up a sufficiently strong parallel between the mathematical form of the above relationship satisfied by the the squared-returns time series, $(r_t)$, and that of an $AR(1)$ time series $(x_t)$, where $x_t = \alpha + \phi_1 x_{t-1} + w_t$, for which $(w_t)$ is white noise.

The squared-returns time series of a $GARCH(1,1)$ model may be understood in a similar way. Here, $(r_t)$ satisfies

$$\begin{aligned} r_t^2 &= \sigma_t^2 \epsilon_t^2 \\ \sigma_t^2 &= \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \end{aligned}$$

which implies

$$r_t^2 = \sigma_t^2 + \sigma_t^2 v_t = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \sigma_t^2 v_t,$$

having written $v_t = \epsilon_t^2 - 1$, as before. From this we have the sense that the squared-returns time series satisfies relationships similar to an $ARMA(1,1)$ time series.

These parallels continue in the extensions to $ARCH(m)$ and $GARCH(p,q)$ models: the squared-returns time series of an $ARCH(m)$ model satisfies a relationship resembling that of an $AR(m)$ time series; the squared-returns time series of an $GARCH(p,q)$ model satisfies a relationship resembling that of an $ARMA(p,q)$ time series.

The returns time series $(r_t)$ generated under an $ARCH(m)$ or $GARCH(p,q)$ model would exhibit short periods of high volatility, but the time series is stable overall. Exploration of the simple $AR(1)$ model suggest some of the subtle mathematical properties that suggest the usefulness of this class of time series models for describing patterns of volatility.

**Section 7B.2: Mathematical properties and implications for modeling.**

Leaving details omitted, two relevant conditional calculations, which apply to both $ARCH(m)$ and $GARCH(p,q)$ models, characterize the time series of returns $(r_t)$ as falling into a special class of time series models called *martingale difference* time series. One defining characteristic of a martingale difference time series is $E[r_t | r_{t-1}, r_{t-2}, \ldots] = 0$. This property implies, for example,

$$E[r_t] = E\left[E[r_t | r_{t-1}, r_{t-2}, \ldots]\right] = E[0] = 0,$$

and, since $Cov[r_{t+h}, r_t] = E[r_{t+h} r_t]$ when $E[r_t] = 0$,

$$\begin{aligned} Cov[r_{t+h}, r_t] &= E\left[E[r_{t+h} r_t | r_{t+h-1}, r_{t+h-2}, \ldots]\right] \\ &= E\left[r_t E[r_{t+h} | r_{t+h-1}, r_{t+h-2}, \ldots]\right] = E[0] = 0. \end{aligned}$$

These deductions show that $(r_t)$ is a mean-zero uncorrelated time series, suggesting that $ARCH(m)$ and $GARCH(p,q)$ models are a kind of extension of white noise. Accordingly, these models may serve as the starting point to more elaborate models such as an ARMA model with ARCH or GARCH errors, say, or a de-trending model with ARCH or GARCH residual deviations.

**<u>Example:</u> Dow Jones Industrial Average (continued)**

A sophisticated model for the Dow Jones Industrial Average time series (of returns) combines and $AR(1)$ model for the returns with a $GARCH(1,1)$ model for the errors, together with a modified distributional assumption that replaces standard normal distributions with t distributions. The relationship is

$$r_t = \phi_1 r_{t-1} + y_t,$$

which identifies the $AR(1)$ piece, where $(y_t)$ is a $GARCH(1,1)$ time series such that

$$\begin{aligned} y_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \end{aligned}$$

where $\epsilon_t$ follows a t distribution. □

The $ARCH(1)$ setup offers a concept for understanding how these general classes of models get a handle on volatility. Suppose $(r_t)$ is an $ARCH(1)$ time series of returns, such that $r_t = \sigma_t \epsilon_t$ and $\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2$, for a Gaussian white-noise time series, $(\epsilon_t)$, with variance one, $Var[\epsilon_t] = 1$. As we have already discussed, the squared-returns time series satisfies the $AR(1)$-type relationship $r_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \sigma_t^2 v_t$. From this it is straightforward to deduce

$$E[r_t^2] = \frac{\alpha_0}{1-\alpha_1},$$

which follows from the observation

$$E[r_t^2] = \alpha_0 + \alpha_1 E[r_{t-1}^2] + E[v_t^*] = \alpha_0 + \alpha_1 E[r_t^2].$$

Manipulating the formula $r_t^2 = \sigma_t^2 \epsilon_t^2 = (\alpha_0 + \alpha_1 r_{t-1}^2)\epsilon_t^2$ will provide that, under conditions of stationarity and causality,

$$E[r_t^4] = \frac{3\alpha_0^2}{(1-\alpha_1)^2} \frac{1-\alpha_1^2}{1-3\alpha_1^2}.$$

To see this, first expand the square of $r_t^2 = (\alpha_0 + \alpha_1 r_{t-1}^2)\epsilon_t^2$ as

$$r_t^4 = (\alpha_0 + \alpha_1 r_{t-1}^2)^2 \epsilon_t^4 = (\alpha_0^2 + \alpha_1^2 r_{t-1}^4 + 2\alpha_0 \alpha_1 r_{t-1}^2)\epsilon_t^4;$$

then make use of causality to take expectations:

$$E[r_t^4] = (\alpha_0^2 + \alpha_1^2 E[r_{t-1}^4] + 2\alpha_0\alpha_1 E[r_{t-1}^2])E[\epsilon_t^4].$$

Stationarity implies that $E[r_t^4] = E[r_{t-1}^4]$, and we have already worked out the formula for $E[r_{t-1}^2]$; additionally noting the property $E[\epsilon_t^4] = 3$ yields

$$E[r_t^4] = 3\alpha_0^2 + 3\alpha_1^2 E[r_t^4] + 6\alpha_0^2 \alpha_1/(1-\alpha_1),$$

hence

$$(1-3\alpha_1^2)E[r_t^4] = 3\alpha_0^2 + 6\alpha_0^2 \frac{\alpha_1}{1-\alpha_1} = 3\alpha_0^2 \frac{1+\alpha_1}{1-\alpha_1} = 3\alpha_0^2 \frac{1-\alpha_1^2}{(1-\alpha_1)^2},$$

from which the above formula for $E[r_t^4]$ follows.

Since $E[r_t^4]$ cannot be negative, the above formula only makes sense if $0 \leq \alpha_1 < 1$ and $3\alpha_1^2 < 1$. These settings are, in fact, required for the returns time series $(r_t)$ to be stationary and to have finite variance.

It follows from the formulas for $E[r_t^2]$ and $E[r_t^4]$ that the kurtosis of $(r_t)$ is

$$\kappa \quad = \quad \frac{E(r_t^4)}{\{E(r_t^2)\}^2} = 3\frac{1-\alpha_1^2}{1-3\alpha_1^2}.$$

It is not difficult to deduce that a lower bound for for the kurtosis is $\kappa \geq 3$. (To see this, consider, *e.g.*, that $\kappa = 3$ when $\alpha_1 = 0$, and that the formula's denominator is smaller than the numerator for $\alpha_1 > 0$). A kurtosis exceeding 3 (which is the kurtosis of a normal distribution) characterizes the distribution of $r_t$ as having "fat tails." In other words, the time series $r_t$ tends to exhibit outliers, which are expressed as periods of high volatility.