STATISTICAL METHODS WITH APPLICATION TO FINANCE

a.y. 2023-2024

Models for Changing Variance

R. Pappadà
rpappada@units.it

Department of Economics, Business, Mathematics and Statistics "B. de Finetti"
University of Trieste

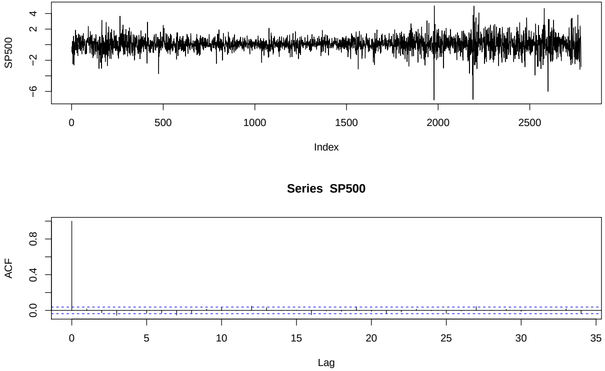
May 15, 2024

Outline

Table of Contents

S&P500 series

Consider the time plot and correlogram of the daily returns of the S&P500 Index (January 2, 1990 to December 31, 1999):



S&P500 series

Although many financial time series appear to be stationary, they often exhibit periods of increased variability (*volatility*)

If a series exhibits a changing variance, so that the variance is correlated in time, the series has a non-constant volatility that is called *conditional heteroscedastic*

The correlogram of a volatile series does not differ significantly from white noise but if the variance is non-constant the *correlogram of the squared values* (provided the series is adjusted to have zero mean) will do

S&P500 series

The mean of the S&P500 returns between January 2, 1990 and December 31, 1999 is 0.0458. The correlogram of the squared mean-adjusted values of the S&P500 index is given below:

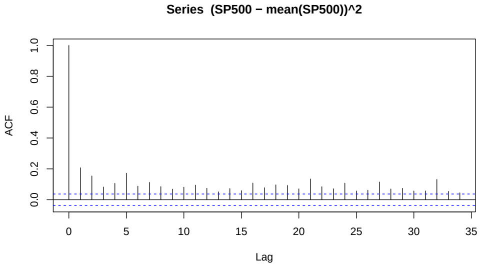


Figure 1: Returns of the Standard and Poors S&P500 Index: correlogram of the squared mean-adjusted values $(r_t - \bar{r})^2$

Conditional volatility

ARMA models were used to model the conditional mean of a process when the conditional variance was constant.

For instance, an AR(1) model for the log returns r_t implies that, conditional on the past return r_{t-1}, we have

$$\mathsf{E}(r_t|r_{t-1}) = \phi_0 + \phi_1 r_{t-1}$$
$$\mathsf{Var}(r_t|r_{t-1}) = \mathsf{Var}(a_t) = \sigma_a^2 \text{ (constant)}$$

where the error series $\{a_t\}$ is assumed to be a white noise series with zero mean and variance σ_a^2.

We focus on modelling the conditional variance of an asset return series by using models in the (Generalized) Autoregressive Conditionally Heteroscedastic–(G)ARCH–class

Volatility models

Volatility is not directly observable, although very often we observe that

- the volatility is high for certain time periods and low for other periods (volatility clusters)
- volatility varies within some fixed range
- volatility seems to react differently to a big price increase and a big price drop with the latter having a greater impact (*leverage effect*)

Such properties are important in the development of **volatility models**: ARCH models (Engle, 1982), later extended to generalized ARCH, or GARCH models (Bollerslev, 1986), and further varieties of ARCH models.

The manner under which the conditional variance

$$\sigma_t^2 = \text{Var}(r_t | F_{t-1})$$

evolves over time distinguishes one volatility model from another.

Table of Contents

The ARCH(1) model

Let $\{x_t\}$ be an observed series. Let $\{y_t\}$ be a series derived from $\{x_t\}$, by removing any trend and seasonal effects, or linear (short-term correlation) effects. Thus $\{y_t\}$ could, for example, be

- the series of residuals from a regression, an AR, or ARMA model
- the first differences of a financial time series such as the log of a share price (returns) for which a random walk model has been adopted

We may represent all such derived series having mean zero in the form

$$Y_t = \sigma_t Z_t$$

where $\{Z_t\}$ denotes a sequence of **iid random variables with zero mean and unit variance**, i.e., an iid WN (or SWN). We will further assume that the square of σ_t depends on the most recent value of $\{y_t\}$.

The ARCH(1) model

We start considering an autoregressive model for the variance process.

The first-order autoregressive conditionally heteroscedastic model, **ARCH(1)**, for Y_t is

$$Y_t = \sigma_t Z_t \quad (1)$$

$$\sigma_t^2 = \omega + \alpha Y_{t-1}^2 \quad (2)$$

where we assume that

- σ_t^2 is the conditional variance of Y_t given past values
- $\{Y_t\}$ has zero mean
- $Z_t \sim iid\, \mathrm{WN}(0, 1)$ (zero mean and unit variance)
- $\omega > 0$, $0 \leq \alpha < 1$ are model parameters.

The ARCH(1) model: Remarks

From Eq.(1) and Eq.(2) we see that

- If y_{t-1} has an unusually large absolute value, then σ_t is larger than usual and so y_t is also expected to have an unusually large magnitude.
- Because of this behaviour, unusual volatility in y_t tends to persist, though not forever.
- the ARCH(1) models returns as a white noise process with nonconstant conditional variance σ_t^2:
  - ACF of Y_t is that of a (weak) white noise
  - if Y_t is ARCH(1), then it can be shown that $\{Y_t^2\}$ has the same form of ACF as an AR(1) model

The ARCH(1) model: Properties

To see how the ARCH(1) model introduces volatility, square Eq.(1) to calculate the unconditional variance

$$\begin{aligned} \mathsf{Var}(Y_t) &= \mathsf{E}(Y_t^2) = \mathsf{E}[(\omega + \alpha\, Y_{t-1}^2) Z_t^2] \\ &= \mathsf{E}(Z_t^2)\mathsf{E}(\omega + \alpha\, Y_{t-1}^2) \\ &= \mathsf{E}(\omega + \alpha\, Y_{t-1}^2) \\ &= \omega + \alpha \mathsf{E}(Y_{t-1}^2) \\ &= \omega + \alpha \mathsf{Var}(Y_{t-1}) \end{aligned} \tag{3}$$

where we used the fact that Since Z_t is independent of Y_{t-1}, $\{Z_t\}$ has unit variance ($\mathsf{E}(Z_t^2) = 1$) and $\{Y_t\}$ has zero mean ($\mathsf{E}(Y_t^2) = \mathsf{Var}(Y_t)$).

- The variance of an ARCH(1) process behaves just like an AR(1) model. Hence, a decay in the autocorrelations of the squared residuals $\{a_t^2\}$ should indicate whether an ARCH model is appropriate or not for modeling $\{a_t\}$

The ARCH(1) model: Properties

To see how the ARCH(1) model introduces volatility, square Eq.(1) to calculate the unconditional variance

$$\begin{aligned}
\mathsf{Var}(Y_t) &= \mathsf{E}(Y_t^2) = \mathsf{E}[(\omega + \alpha\, Y_{t-1}^2) Z_t^2] \\
&= \mathsf{E}(Z_t^2)\mathsf{E}(\omega + \alpha\, Y_{t-1}^2) \\
&= \mathsf{E}(\omega + \alpha\, Y_{t-1}^2) \\
&= \omega + \alpha \mathsf{E}(Y_{t-1}^2) \\
&= \omega + \alpha \mathsf{Var}(Y_{t-1})
\end{aligned} \tag{3}$$

where we used the fact that Since Z_t is independent of Y_{t-1}, $\{Z_t\}$ has unit variance ($\mathsf{E}(Z_t^2) = 1$) and $\{Y_t\}$ has zero mean ($\mathsf{E}(Y_t^2) = \mathsf{Var}(Y_t)$).

- The variance of an ARCH(1) process behaves just like an AR(1) model. Hence, a decay in the autocorrelations of the squared residuals $\{a_t^2\}$ should indicate whether an ARCH model is appropriate or not for modeling $\{a_t\}$

The ARCH model: Properties

- From Eq.(3) the (unconditional) variance can be obtained by assuming Y_t stationary (Var$(Y_{t-1}) = $ Var$(Y_t) = \sigma^2$)

$$\sigma^2 = \frac{\omega}{1-\alpha}, \quad 0 < \alpha < 1$$

- the ARCH(1) model has a constant mean (both conditional and unconditional)

$$\mathrm{E}(Y_t | Y_{t-1}, \dots) = 0$$

- and a time-varying conditional variance

$$\mathrm{Var}(Y_t | Y_{t-1}, \dots) = \sigma_t^2$$

Simulated ARCH(1) model

The simulated series (Y_t) is generated from the ARCH(1) model

$$Y_t = \sigma_t Z_t, \quad \sigma_t^2 = \omega + \alpha Y_{t-1}^2$$

with $Z_t \sim N(0,1), \omega = 0.1, \alpha = 0.4$. This is equivalent to Eq.(1)-(2), where σ_t^2 denotes the conditional variance.

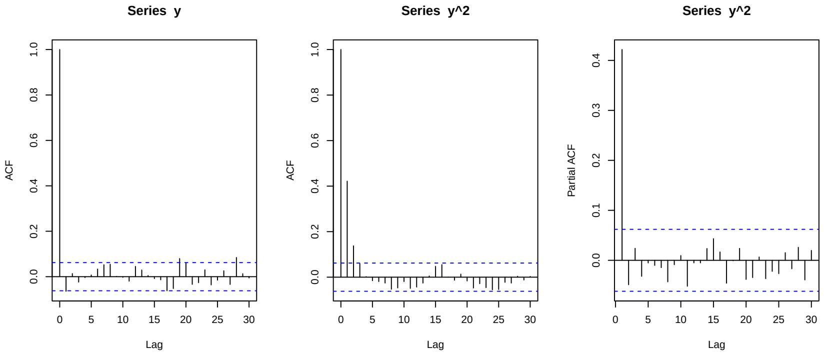


Figure 2: *Left:* ACF of simulated series; *Middle and Right:* ACF/PACF of squared values of simulated series from the ARCH(1) model.

Table of Contents

The ARCH(m) model

The first-order ARCH model can be extended to a mth-order process by including higher lags. An ARCH(m) process is given by

$$Y_t = \sigma_t Z_t \tag{4}$$

where

$$\sigma_t^2 = \omega + \alpha_1 Y_{t-1}^2 + \cdots + \alpha_m Y_{t-m}^2 \tag{5}$$

- σ_t^2 is the conditional variance of Y_t given the past values
- $\{Z_t\}$ iid process with mean zero and variance 1
- $\omega > 0$, $\alpha_1, \ldots, \alpha_m \geq 0$.

Note that Eq.(4) is the same as Eq.(1), while Eq.(5) now contains the past values $Y_{t-1}^2, \ldots, Y_{t-m}^2$.

ARCH Models: pros and cons

ARCH models have some main advantages in analyzing asset returns:

- the dependence of Y_t can be described by a simple quadratic function of its lagged values
- they can produce volatility clusters
- they allow for heavy tails

ARCH models also have some weaknesses:

- they assume positive and negative shocks have the same effects on volatility because it depends on the square of the previous shocks
- the conditional standard deviation can exhibit more persistent periods of high or low volatility than seen in an ARCH process

ARCH Models: pros and cons

ARCH models have some main advantages in analyzing asset returns:

- the dependence of Y_t can be described by a simple quadratic function of its lagged values
- they can produce volatility clusters
- they allow for heavy tails

ARCH models also have some weaknesses:

- they assume positive and negative shocks have the same effects on volatility because it depends on the square of the previous shocks
- the conditional standard deviation can exhibit more persistent periods of high or low volatility than seen in an ARCH process

Table of Contents

GARCH(m, s), models

A generalization of the ARCH model that allows the variance to depend on past values of both the series and the volatility in squared form is the **generalized ARCH (or GARCH) model**.

Y_t is said to follows a GARCH model of order (m, s) when

$$Y_t = \sigma_t Z_t, \quad \{Z_t\} \sim SWN(0,1)$$

and the local conditional variance is given by

$$\sigma_t^2 = \omega + \sum_{i=1}^{m} \alpha_i Y_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2 \tag{6}$$

where $\omega \geq 0$, $\alpha_i, \beta_j \geq 0$, and the sum $\sum \alpha_i + \sum \beta_j < 1$ in order for the process to be stationary.
The GARCH(m, s) model has the ARCH(m) model as the special case GARCH(m, 0).

The GARCH(1,1) model

The GARCH(1,1) model is

$$Y_t = \sigma_t Z_t, \quad \sigma_t^2 = \omega + \alpha_1 Y_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{7}$$

with $Z_t \sim iid\text{WN}(0,1)$, $\omega, \alpha_1, \beta_1 \geq 0$, and $\alpha_1 + \beta_1 < 1$ to ensure stability.

- $|Y_t|$ has a chance of being large if either $|Y_{t-1}|$ is large or σ_{t-1} is large (volatility clustering)
- Similar to ARCH models, the tail distribution of a GARCH(1,1) process is heavier than that of a normal distribution
- GARCH(1,1) unconditional variance is

$$\sigma^2 = \frac{\omega}{1 - \alpha_1 - \beta_1}$$

- Similarly to ARCH models, one can establish parallels with the ARMA(1,1) process

Simulated GARCH model

The simulated series (Y_t) is generated from the GARCH(1,1) model

$$Y_t = \sigma_t Z_t, \quad \sigma_t^2 = \omega + \alpha_1 Y_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

with $Z_t \sim N(0,1)$, $\omega = 0.1$, $\alpha_1 = 0.4$, $\beta_1 = 0.2$. This is equivalent to Eq.(7), where σ_t^2 denotes the conditional variance.

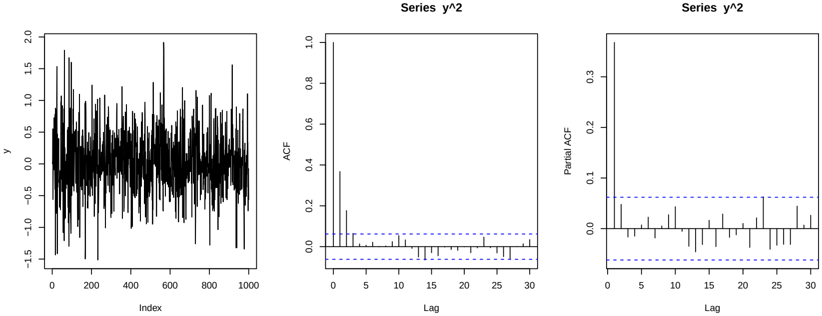


Figure 3: *Left:* ACF of simulated series; *Middle and Right:* ACF/PACF of squared values of simulated series from the GARCH(1,1) model ($n = 1000$)

Table of Contents

t-GARCH

In the previous example with simulated data, the $\{Z_t\}$ are standard normal innovations.

If we want to account for asymmetry or fat tails, we can consider alternative distributions for the Z_t process, depending on additional parameters that modify the skewness and kurtosis

Two models employing alternative distributions for the innovations are

- t-GARCH model: the process Z_t follows a (scaled) Student's t distribution with ν dof (to be estimated)
- skew t-GARCH model: the return distribution can be asymmetric

Skew t-GARCH

- For the Student-t distribution, as the degrees of freedom increase the tails become shorter and the peak becomes lower.
- For the *skew* t distribution with 5 df, a skew parameter η equal to 0.75, 1, and 1.5, produces left-skew, symmetric, and right-skew density, respectively.

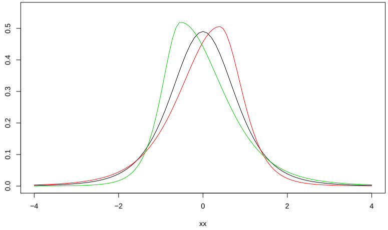


Figure 4: The skew standardized Student-t with 5 df and degrees of skewness $\eta = 0.75$ (red), 1 (black), and 1.5 (green).

Table of Contents

GARCH-based volatility prediction

Suppose that the return data $Y_1, \ldots, Y_n$ follow a particular model in the GARCH family

- We want to forecast future volatility, i.e, to predict the value of σ_{n+h} for $h \geq 1$
- We again assume that we have access to the infinite history of the process up to time $t = n$ ($\mathcal{F}_n$) and adapt our prediction formula to take account of the finiteness of the sample
- Assume that the GARCH model has been fitted and its parameters estimated

We consider the case of simple GARCH(1,1) models that can be easily generalized.

Prediction in the GARCH(1, 1) model

For a GARCH(1,1) model the conditional variance is

$$\sigma_t^2 = \omega + \alpha_1 Y_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

Predictions of Y_{n+1}^2 based on $\mathcal{F}_n$ are given by

$$E(Y_{n+1}^2 | \mathcal{F}_n) = \text{Var}(Y_{n+1} | \mathcal{F}_n) = \sigma_{n+1}^2$$

and

$$\sigma_{n+1}^2 = \omega + \alpha_1 Y_n^2 + \beta_1 \sigma_n^2$$

We approximate σ_n^2 by an estimate of squared volatility $\hat{\sigma}_n^2$, hence we obtain a recursive scheme for estimating volatility one step ahead:

$$\sigma_{n+1}^2 = \hat{\omega} + \hat{\alpha}_1 y_n^2 + \hat{\beta}_1 \hat{\sigma}_n^2 \tag{8}$$

ARMA-GARCH Model Specification

A common approach is to fit an **ARMA model with GARCH errors** to the series of daily log returns:

$$X_t = \mu_t + a_t \tag{9}$$

where

- *mean equation*

$$\mu_t = \phi_0 + \sum_{i=1}^{p} \phi_i X_{t-i} + \sum_{i=1}^{q} \theta_i a_{t-i}$$

- *variance equation*

$$a_t = \sigma_t Z_t, \quad \{Z_t\} \sim \text{SWN}(0,1)$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2$$

Z_t can have a non-normal distribution (e.g., Student-t or *skew* Student-t distribution); $\omega > 0$, $\alpha_i, \beta_j \geq 0$, $\sum_i \alpha_i + \sum_j \beta_j < 1$.

Residuals for ARMA-GARCH

We consider a general ARMA-GARCH model of the form $X_t - \mu_t = a_t = \sigma_t Z_t$. We distinguish between

- the ordinary residuals $\hat{a}_1, \ldots, \hat{a}_n$ from the ARMA model

$$\hat{a}_t = x_t - \hat{x}_t$$

  (under the hypothesized model they should behave like a realization of a pure GARCH process)

- the standardized residuals that are calculated from the former by

$$\hat{z}_t = \hat{a}_t / \hat{\sigma}_t \qquad \hat{\sigma}_t^2 = \hat{\omega} + \sum_{i=1}^{m} \hat{\alpha}_i \hat{a}_{t-i}^2 + \sum_{j=1}^{s} \hat{\beta}_j \hat{\sigma}_{t-j}^2$$

(Starting values of $\hat{a}_t$ can be set equal to zero and starting values of the volatility $\hat{\sigma}_t$ equal to either the sample variance or zero)

Model Checking

The standardized residuals $\hat{z}_t = \hat{a}_t / \hat{\sigma}_t$ where $\hat{\sigma}_t$ expresses the volatility, should behave like an **SWN** ($\hat{z}_t$ and $\hat{z}_t^2$ should be uncorrelated); this can be investigated by

- performing Ljung-Box Tests with various lags
- constructing correlograms of raw and absolute values

The null hypothesis for these tests should be accepted in order to consider the fitted model as a good one; normality tests can be used if the Z_t are assumed to be $N(0, 1)$

Prediction in an ARMA(1,1)-GARCH(1,1) model

Assume a model of the form (9) $X_t - \mu_t = a_t$ where

- μ_t describes an ARMA(1,1) model
- $a_t = \sigma_t Z_t$ follows a GARCH(1,1) model

We have a sample $x_1, \dots, x_n$ and we fit an ARMA(1,1) model; the forecast of X_{n+1} is

$$E(X_{n+1}|\mathcal{F}_n) = E(\mu_{n+1}|\mathcal{F}_n) = \hat{x}_n(1) = \hat{\mu} + \hat{\phi}_1(x_n - \hat{\mu}) + \hat{\theta}_1 a_n;$$

the following yields prediction of σ^2_{n+1}

$$\mathrm{Var}(X_{n+1}|\mathcal{F}_n) = E(a^2_{n+1}|\mathcal{F}_n) = \hat{\omega} + \hat{\alpha}_1 a_n^2 + \hat{\beta}_1 \sigma_n^2$$

and these are approximated by substituting inferred values for a_t and σ_t obtained from the residual equations.