Abstract: Well-known Box-Jenkins Autoregressive Integrated Moving Average (ARIMA) methodology has virtually dominated analysis of time-series data, particularly during the period 1930-‘80. However, one limitation of this methodology is that it is not capable of modelling those data sets that depict volatility. Fortunately, to this end, Autoregressive Conditional Heteroscedastic (ARCH) families of parametric nonlinear time-series models have been proposed during last two decades or so. Various aspects of these models are thoroughly discussed. Estimation procedure for fitting ARCH models is also described. Modelling and forecasting of volatile black pepper price time-series data is carried out. An extension of ARCH family to incorporate the effect of regressor in mean and variance equations is also considered. Details for application of AR(p)-ARCH-in-Mean model to describe volatility in onion price time-series data are provided. In the end, some extensions, like Generalized ARCH (GARCH), and Threshold ARCH (TARCH), are briefly mentioned.

Key words: Time-series, ARIMA, Heteroscedasticity, Volatility, Nonlinear Time-Series, ARCH.

1. Introduction
A data set containing observations on a single phenomenon observed over multiple time periods is called time-series. In time-series data, both the values and the ordering of the data points have meaning. For many agricultural products, data are usually collected over time. Well-known Box-Jenkins Autoregressive Integrated Moving Average (ARIMA) methodology has virtually dominated analysis of time-series data for last five decades or so, viz. 1930-80. The least squares model is most important in applied econometrics. This is a natural choice, because applied econometricians are typically called upon to determine how much one variable will change in response to a change in some other variable. Increasingly however, econometricians are being asked to forecast and analyze the size of the errors of the model. In this case, the questions are about volatility, and the standard models have become the Autoregressive Conditional Heteroscedastic (ARCH) models. Data in which the variances of the error terms are not equal, in which the error terms may reasonably be expected to be larger for some points or ranges of the data than for others, are said to suffer from heteroscedasticity. The standard warning is that in the presence of heteroscedasticity, the regression coefficients for an ordinary least squares regression is still unbiased, but the standard errors and confidence intervals estimated by conventional procedures will be too narrow, giving a false sense of precision. Instead of considering this as a problem to be corrected, ARCH and GARCH models treat heteroscedasticity as a variance to be modeled. As a result, not only are the deficiencies of least squares corrected,
but also a prediction is computed for the variance of each error term. This prediction turns out often to be of interest, particularly in applications in finance.

2. Linear Time-Series Models
When the function of the regressor variables is linear then it is called linear time-series models. In other word, if the function relating the observed time-series \( \{X_t\}_{t=0}^{\infty} \) and the underlying shocks \( \{e_t\}_{t=0}^{\infty} \) is linear then this is called linear time-series model. There are different kinds of linear time-series models such as:

2.1 Autoregressive Models
A stochastic model that can be extremely useful in the representation of certain practically occurring series is the autoregressive model. In this model, the current value of the process is expressed as a finite, linear aggregate of previous values of the process and a shock \( \varepsilon_t \).
Let us denote the values of a process at equally spaced time \( t, t-1, t-2, \ldots \) by \( y_t, y_{t-1}, y_{t-2}, \ldots \), then \( y_t \) can be described by the following expression:

\[
y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \ldots + \varphi_p y_{t-p} + \varepsilon_t.
\]

If we define an autoregressive operator of order \( p \) by \( \phi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \ldots - \varphi_p B^p \), where \( B \) is the back shift operator such that \( B y_t = y_{t-1} \), the autoregressive model can be written as \( \phi(B) y_t = \varepsilon_t \).

2.2 Moving Average Models
Another kind of model of great practical importance in the representation of observed time-series is the finite moving average process. MA(\( q \)) model is defined as

\[
y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q}.
\]

If we define an moving average operator of order \( q \) by \( \theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q \), where \( B \) is the back shift operator such that \( B y_t = y_{t-1} \), the autoregressive model can be written as \( y_t = \theta(B) \varepsilon_t \).

2.3 ARMA(\( p, q \)) Model
To achieve greater flexibility in fitting of actual time-series, it is sometimes advantageous to include both autoregressive and moving average model. This leads to the mixed autoregressive–moving average model

\[
y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \ldots + \varphi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q}
\]
or
\[
\phi(B) y_t = \theta(B) \varepsilon_t.
\]

This is written as ARMA(\( p, q \)) model. In practice it is frequently true that adequate representation of actually occurring stationary time series can be obtained with
autoregressive, moving average, or mixed models, in which \( p \) and \( q \) are not greater than 2 and often less than 2.

2.4 ARIMA(\( p, d, q \)) Model

A generalization of ARMA models which incorporates a wide class of non stationary time-series is obtained by introducing the differencing into the model. The simplest example of a non stationary process which reduces to a stationary one after differencing is Random Walk. A process \( \{ y_t \} \) is said to follow an Integrated ARMA model, denoted by ARIMA\((p, d, q)\), if

\[
\nabla^d y_t = (1 - B)^d \varepsilon_t
\]

is ARMA\((p, q)\). The model is written as \( \phi(B)(1-B)^d y_t = \theta(B)\varepsilon_t \) where \( \varepsilon_t \sim WN(0, \sigma^2) \).

The integration parameter \( d \) is a nonnegative integer. When \( d = 0 \) we have the usual ARMA model, that is ARIMA\((p, d, q)\) \( \equiv \) ARMA\((p, q)\).

2.5 Structural Time Series Modeling (STSM)

One limitation of ARIMA approach is the assumption of stationarity. It is not always possible to make a time-series stationary by differencing or by some other means. So ARIMA approach can be applied only to a limited set of data. A quite promising, mechanistic approach, which does not suffer from this drawback, is “Structural time series modeling (STSM)”.

The distinguishing feature of this methodology is that observations are made up of distinct components, such as trend and cyclical fluctuations and each of which is modeled separately. The simplest Structural time series models consist of a trend component plus a random disturbance term. The random disturbance term may be interpreted as an irregular component in the time-series or as a measurement error. Either way the model may be written as \( y_t = \mu_t + \varepsilon_t \), where \( \mu_t \) is the trend and \( \varepsilon_t \) is the disturbance term, which is assumed to be uncorrelated with any stochastic elements in \( \mu_t \). The trend may take variety of forms. For example the deterministic linear trend is \( \mu_t = \alpha + \beta t \). For time-series data having prominent cyclical fluctuations, the possible STSM models are: Cycle plus noise model (CNM), trend plus cycle model (TCM) and cyclical trend model (CTM). An excellent discussion of various aspects of STSM is given in Harvey (1996). The techniques that emerge from STSM approach are extremely flexible and are capable of handling a much wider range of problems than is possible through ARIMA approach.

3. Limitation of ARIMA and STSM Models

One limitation of both ARIMA and STSM is that these yield “linear models”. However, in reality, underlying relationships among variables is highly complex and can not be described satisfactorily through a linear modeling approach. There are many features, like existence of threshold value, which can be described through a nonlinear approach. During the last two decades or so, a new area of “Nonlinear time-series modeling” is fast coming up. Here, there are basically two possibilities, viz. Parametric or Nonparametric approaches. Evidently, if in a particular situation, we are quite sure about the functional form, we should use the former; otherwise the latter may be employed. However, in this seminar, we shall confine our attention only to “Parametric approach”.

3
4. Nonlinear Time Series Modelling

When dealing with nonlinearity, Campbell et al. (1997) made the distinction between:

• Linear Time-series: shocks are assumed to be uncorrelated but not necessarily identically independent distributed (iid).
• Nonlinear Time-series: shocks are assumed to be iid, but there is a nonlinear function relating the observed time-series \( \{X_t\}_{t=0}^\alpha \) and the underlying shocks, \( \{\varepsilon_t\}_{t=0}^\alpha \).

They suggested the following structure to describe a nonlinear process:

\[
X_t = g(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) + \varepsilon_t h(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots). \quad \mathbb{E}[X_t / \psi_{t-1}] = g(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) \\
\text{Var}[X_t / \psi_{t-1}] = \mathbb{E}[\{(X_t - \mathbb{E}(X_t)) / \psi_{t-1}\}^2] = \mathbb{E}\left[\left(\varepsilon_t h(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) / \psi_{t-1}\right)^2\right] \\
= \text{Var}\left[\left(\varepsilon_t h(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) / \psi_{t-1}\right)\right] \\
= \{h(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) / \psi_{t-1}\}^2,
\]

where the function \( g(\cdot) \) corresponds to the conditional mean of \( X_t \), and the function \( h(\cdot) \) is the coefficient of proportionality between the innovation in \( X_t \) and the shock \( \varepsilon_t \).

The general form above leads to a natural division in Nonlinear Time-series literature in two branches:

• Models Nonlinear in Mean: \( g(\cdot) \) is nonlinear;
• Models Nonlinear in Variance: \( h(\cdot) \) is nonlinear.

Examples

• Nonlinear Moving Average Model: \( X_t = \varepsilon_t + \alpha \varepsilon_{t-1}^2 \). Here the function \( g = \alpha \varepsilon_{t-1}^2 \) and the function \( h = 1 \). Thus, it is nonlinear in mean but linear in variance.
• Engle’s (1982) ARCH Model: \( X_t = \varepsilon_t \sqrt{\alpha \varepsilon_{t-1}^2} \). The process is nonlinear in variance but linear in mean. The function \( g(\cdot) = 0 \) and the function \( h = \sqrt{\alpha \varepsilon_{t-1}^2} \).

4.1 Volatility

Volatility refers to the spread of all likely outcomes of an uncertain variable. Typically, in financial markets, we are often concerned with the spread of asset returns. Statistically, volatility is often measured as the sample standard deviation. Sometimes, variance, \( \sigma^2 \), is used also as a volatility measure. Standard deviation has the same unit of measure as the mean, \textit{i.e.} if the mean is in dollar, then standard deviation is also expressed in dollar whereas variance will be expressed in dollar square. For this reason, standard deviation is more convenient and intuitive when we think about volatility.

Volatility is related to, but not exactly the same as, risk. Risk is associated with undesirable outcome, whereas volatility as a measure strictly for uncertainty could be due to a positive outcome. This important difference is often overlooked. Volatility (or standard deviation) is only a measure for the spread of a distribution and has no information on its shape.

5. Autoregressive Conditional Heteroscedastic (ARCH) Models
Definition: The ARCH(q) model for the series \{\varepsilon_t\} is defined by specifying the conditional distribution of \varepsilon_t given the information available up to time \(t-1\). Let \(\psi_{t-1}\) denote this information. It consists of the knowledge of all available values of the series, and anything which can be computed from these values, e.g., innovations, squared observations, etc. In principle, it may even include the knowledge of the values of other related time-series, and anything else which might be useful for forecasting and is available by time \(t-1\).

We say that the process \{\varepsilon_t\} is ARCH(q) if the conditional distribution of \{\varepsilon_t\} given the available information \(\psi_{t-1}\) is

\[\varepsilon_t \mid \psi_{t-1} \sim N(0, h_t) \quad \text{...(5.1)}\]

and

\[h_t = a_0 + \sum_{i=1}^{q} a_i \varepsilon_{t-i}^2 \quad \text{...(5.2)}\]

where \(a_0 > 0\), \(a_i \geq 0\) for all \(i\) and \(\sum_{i=1}^{q} a_i < 1\).

Equation (5.1) says that the conditional distribution of \{\varepsilon_t\} given \(\psi_{t-1}\) is normal, \(N(0, h_t)\). In other words, given the available information \(\psi_{t-1}\) the next observation \{\varepsilon_t\} has a normal distribution with a (conditional) mean of \(E[\varepsilon_t / \psi_{t-1}] = 0\), and a (conditional) variance of \(\text{var}[\varepsilon_t / \psi_{t-1}] = h_t\).

Equation (5.2) specifies the way in which the conditional variance \(h_t\) is determined by the available information. Note that \(h_t\) is defined in terms of squares of past innovations. This, together with the assumptions that \(a_0 > 0\) and \(a_i \geq 0\), guarantees that \(h_t\) is positive, as it must be since it is a conditional variance.

5.1 Properties of the ARCH Model

The ARCH (5.1) model is defined as

\[y_t = \varepsilon_t h_t^{1/2} \quad \text{...(5.3)}\]

\[h_t = a_0 + a_1 y_{t-1}^2,\]

where \(a_0 > 0\), \(a_1 \geq 0\) and \{\varepsilon_t\} is a white noise process that means \{\varepsilon_t\} is a sequence of independent and identically distributed (i.i.d) random variables with mean zero and variance 1.

First the unconditional mean of \(y_t\) remains zero because,

\[E(y_t) = E(E(y_t | \psi_{t-1})) = E(\sqrt{h_t} E(\varepsilon_t)) = 0\]

Second the unconditional variance of \(y_t\) can be defined as
\[ \text{Var}(y_t) = E(y_t^2) = E[E(y_t^2 | \psi_{t-1})] = E(a_0 + a_1 y_{t-1}^2) = a_0 + a_1 E(y_{t-1}^2). \]

If \( y_t \) is a stationary process with \( E(y_t) = 0, \) \( \text{var}(y_t) = \text{var}(y_{t-1}) = E(y_{t-1}^2). \) Therefore, we have \( \text{var}(y_t) = a_0 + a_1 \text{var}(y_t) \) and \( \text{var}(y_t) = \frac{a_0}{1 - a_1}. \) Since the variance of \( y_t \) must be positive, we require \( 0 \leq a_1 < 1. \)

Third, in some applications, we need higher order moments of \( y_t \) to exist and, hence, \( a_i \) must satisfy some additional constraints. For instance, to study its tail behavior, we require that the fourth moment of \( y_t \) is finite. Under the normality assumption of \( \{ \varepsilon_t \} : \)

\[ E(y_t^4 | \psi_{t-1}) = 3 \left[ E(y_t^2 | \psi_{t-1}) \right]^2 = 3(a_0 + a_1 y_{t-1}^2)^2. \]

Therefore, \( E(y_t^4) = E[E(y_t^4 | \psi_{t-1})] = 3E(a_0 + a_1 y_{t-1}^2)^2 = 3E(a_0^2 + 2a_0 a_1 y_{t-1}^2 + a_1^2 y_{t-1}^4). \)

If \( y_t \) is fourth order stationary with \( m_4 = E(y_t^4), \) then we have

\[ m_4 = 3(a_0^2 + 2a_0 a_1 \text{var}(y_t) + a_1^2 m_4) = 3a_0^2 \left( 1 + 2 \frac{a_1}{1 - a_1} \right) + 3a_1^2 m_4 \]

Consequently, \( m_4 = 3a_0^2 \frac{(1 + a_1)}{(1 - a_1)(1 - 3a_1^2)}. \)

This result has two important implications (Tsay, 2005):

(i) since the fourth moment is positive, we see that \( a_1 \) must also satisfy the condition \( 1 - 3a_1^2 > 0; \) that is \( 0 \leq a_1^2 < \frac{1}{3}; \) and

(ii) the unconditional kurtosis of \( y_t \) is

\[ \frac{E(y_t^4)}{[\text{var}(y_t)]^2} = 3 \frac{a_0^2 (1 + a_1)}{(1 - a_1)(1 - 3a_1^2)} \times \frac{(1 - a_1)^2}{a_0^2} = 3 \frac{(1 - a_1^2)}{(1 - 3a_1^2)} > 3. \]

Thus the excess kurtosis of \( y_t \) is positive and the tail distribution of \( y_t \) is heavier than that of a normal distribution. These properties continue to hold for general ARCH models, but the formula become more complicated for higher order ARCH models.

Heavy tails are a common aspect of financial data, and hence the ARCH models are so popular in this field. Besides that, Bera and Higgins (1993) mention the following reasons for the ARCH success:

- ARCH models are simple and easy to handle
- ARCH models take care of clustered errors
- ARCH models take care of nonlinearities
- ARCH models take care of changes in the econometrician’s ability to forecast.
In spite of its name, the ARCH model is not autoregressive. However, if we add \( \eta_t = \varepsilon_t^2 - h_t \) to both sides of Equation (5.2), we get \( \varepsilon_t^2 = a_0 + \sum_{i=1}^{q} a_i \varepsilon_{t-i}^2 + \eta_t \). It can be shown that \( \eta_t \) is zero mean white noise. Therefore, the squared process \( \{ \varepsilon_t^2 \} \) is an autoregression \( \text{AR} (q) \) with nonzero mean, and AR parameters \( a_1, a_2, \ldots, a_q \). The ARCH \( (q) \) model is nonlinear, since if the \( \{ \varepsilon_t \} \) could be expressed as \( \varepsilon_t = \sum_{k=0}^{a} a_k \varepsilon_{t-k} \), and then we would have \( \text{Var}[\varepsilon_t | \psi_{t-1}] = \text{Var}[\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots] = \text{Var}[\varepsilon_t] \), a constant. This contradicts Equations (5.1) and (5.2), so \( \{ \varepsilon_t \} \) must not be a linear process.

5.2 Testing for ARCH Effects

Let \( \varepsilon_t = y_t - \mu_t \), be the residuals of the mean equation. The squared series \( \{ \varepsilon_t^2 \} \) is then used to check for conditional heteroscedasticity, which is also known as the ARCH effects. The test for conditional heteroscedasticity is the Lagrange multiplier test of Engle (1982). This test is equivalent to usual F statistic for testing \( a_i = 0 \) (i=1,2,...,q) in the linear regression \( \varepsilon_t^2 = a_0 + a_1 \varepsilon_{t-1}^2 + \ldots + a_q \varepsilon_{t-q}^2 + \varepsilon_t \), t = q+1,..., T, where \( \varepsilon_t \) denotes the error term, q is the prespecified positive integer, and T is the sample size. Specifically, the null hypothesis is \( H_0: a_1 = a_2 = \ldots = a_q = 0 \). Let \( SSR_0 = \sum_{t=q+1}^{T} (\varepsilon_t^2 - \bar{\varepsilon})^2 \), where \( \bar{\varepsilon} = (1/T) \sum_{t=q+1}^{T} (\varepsilon_t^2) \) is the sample mean of \( \{ \varepsilon_t^2 \} \), and \( SSR_1 = \sum_{t=q+1}^{T} (\hat{\varepsilon}_t^2) \) where \( \hat{\varepsilon}_t \) is the least square residual of the prior regression. Then we have \( F = \frac{(SSR_0 - SSR_1)/q}{SSR_1 (T - 2q - 1)} \), which is asymptotically distributed as chi squared distribution with q degrees of freedom under the null hypothesis. The decision rule is to reject the null hypothesis if \( F > \chi^2_q (\alpha) \) where \( \chi^2_q (\alpha) \) is the upper 100(1-\( \alpha \))th percentile of \( \chi^2_q \) or the p-value of F is less than \( \alpha \).

5.3 Estimation

Three likelihood functions (Fan and Yao, 2003) are commonly used in ARCH estimation. Under the normality assumption, the likelihood function of an ARCH \( (q) \) model is

\[
f(y_1, y_2, \ldots, y_T | \alpha) = f(y_T | y_{T-1}) f(y_{T-1} | y_{T-2}) \ldots f(y_{q+1} | y_q) f(y_1, y_2, \ldots, y_q | \alpha),
\]

where \( \alpha = (a_0, a_1, \ldots, a_q)' \) and \( f(y_1, y_2, \ldots, y_q | \alpha) \) is the joint probability density function of \( y_1, y_2, \ldots, y_q \). Since the exact form of \( f(y_1, y_2, \ldots, y_q | \alpha) \) is complicated, it is commonly dropped from the prior likelihood function, especially when the sample size is sufficiently large. This results in using the conditional likelihood function

\[
f(y_{q+1}, y_{q+2}, \ldots, y_T | \alpha, y_1, y_2, \ldots, y_q) = \prod_{t=q+1}^{T} \frac{1}{\sqrt{2\pi h_t}} \exp\left( -\frac{y_t^2}{2h_t} \right),
\]
Autoregressive Conditional Heteroscedastic (ARCH) Family of Models for Describing Volatility

where \( h_t \) can be evaluated recursively. We refer to estimates obtained by maximizing likelihood estimates (MLEs) under normality. Maximizing the conditional likelihood function is equivalent to maximizing its logarithm, which is easier to handle. The conditional log likelihood function is

\[
\ell(y_{q+1}, y_{q+2}, \ldots, y_T \mid a, y_1, y_2, \ldots, y_q) = \sum_{t=q+1}^{T} \left( -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(h_t) - \frac{1}{2} \frac{y_t^2}{h_t} \right).
\]

Since the first term \( \ln(2\pi) \) does not involve any parameters, the log likelihood function becomes

\[
\ell(y_{q+1}, y_{q+2}, \ldots, y_T \mid a, y_1, y_2, \ldots, y_q) = -\sum_{t=q+1}^{T} \left( \frac{1}{2} \ln(h_t) + \frac{1}{2} \frac{y_t^2}{h_t} \right),
\]

where \( h_t = a_0 + a_1 y_{t-1}^2 + \ldots + a_q y_{t-q}^2 \) can be evaluated recursively.

In some applications, it is more appropriate to assume that \( \varepsilon_t \) follows a heavy tailed distribution such as a standardized Student-t distribution. Let \( x_\nu \) be a Student-t distribution with \( \nu \) degrees of freedom. Then \( \text{Var}(x_\nu) = \nu/(\nu-2) \) for \( \nu > 2 \), and use \( \varepsilon_t = x_\nu / \sqrt{\nu(\nu-2)} \) the probability density function of \( \varepsilon_t \) is

\[
f(\varepsilon_t \mid \nu) = \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\nu(\nu-2)\pi}} \left( 1 + \frac{\varepsilon_t^2}{\nu-2} \right)^{-(\nu+1)/2}, \quad \nu > 2,
\]

where \( \Gamma(x) \) is the usual gamma function (\( i.e., \Gamma(x) = \int_0^\infty y^{x-1}e^{-y} \, dy \)). Using \( y_t = \varepsilon_t h_t^{1/2} \) we obtain the conditional likelihood function of \( y_t \) as

\[
f(y_{q+1}, y_{q+2}, \ldots, y_T \mid a, y_1, y_2, \ldots, y_q) = \prod_{t=q+1}^{T} \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\nu(\nu-2)\pi}} \frac{1}{\sqrt{h_t}} \left( 1 + \frac{y_t^2}{(\nu-2)h_t} \right)^{-(\nu+1)/2}
\]

where \( \nu > 2 \). We refer to the estimates that maximize the prior likelihood function as the conditional MLEs under t-distribution. The degrees of freedom of the t-distribution can be specified a priori or estimated jointly with other parameters. A value between 3 and 6 is often used if it is prespecified, then the conditional log likelihood function is

\[
\ell(y_{q+1}, y_{q+2}, \ldots, y_T \mid a, y_1, y_2, \ldots, y_q) = -\sum_{t=q+1}^{T} \left[ \frac{v+1}{2} \ln \left( 1 + \frac{y_t^2}{(\nu-2)h_t} \right) + \frac{1}{2} \ln(h_t) \right]
\]

Finally, it may assume a generalized error distribution (GED) with probability density function

\[
f(x) = \frac{\nu \exp \left( -\frac{1}{2} \left[ \frac{x}{\lambda} \right]^\gamma \right)}{\lambda 2^{(\nu+1)/\gamma} \Gamma(1/\nu)}, \quad -\alpha < x \leq \alpha, \quad 0 < \nu < \alpha.
\]
where $\Gamma()$ is the gamma function and $\lambda = \left[2^{(1-2/v)}\Gamma(1/v)/\Gamma(3/v)\right]^{1/2}$

This distribution reduces to Gaussian distribution if $v=2$ and it has heavy tails when $v<2$.

5.4 Forecasting

Forecasts of the ARCH model can be obtained recursively as those of an AR model. Consider an ARCH (q) model. At the forecast origin $t$, the one step ahead forecast of $h_{t+1}$

$$h_t(1) = a_0 + a_1 y_t^2 + \ldots + a_q y_{t+1-q}^2.$$ 

The two step ahead forecast is $h_t(2) = a_0 + a_1 h_t(1) + a_2 y_t^2 + \ldots + a_q y_{t+2-q}^2$, and $l$ step ahead forecast for $h_{t+l}$, $h_t(l) = a_0 + \sum_{i=1}^{q} a_i h_t(l-i)$ where $h_t(l-i) = y_{t+i}^2$ if $l-i \leq 0$.


![Black Pepper](image)

**Figure 5.1**: Time plots of first differences of the logs of black pepper prices.

From Figure 5.1, it can be observed that there are periods where the price fluctuates heavily. Furthermore, there seems to be larger price increases than there are large negative price changes. Hence the pepper price seems to have positive skewness.

**Table 5.1**: An estimated autocorrelation function is given in the following table:

<table>
<thead>
<tr>
<th>Series</th>
<th>$\hat{\rho}_1$</th>
<th>$\hat{\rho}_2$</th>
<th>$\hat{\rho}_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>EACF of $x_t$</td>
<td>0.338</td>
<td>0.024</td>
<td>0.012</td>
</tr>
<tr>
<td>EACF of $\hat{y}_t^2$</td>
<td>0.136</td>
<td>0.194</td>
<td>0.01.</td>
</tr>
</tbody>
</table>

The $x_t$ series is the return series, that is $x_t = \log(z_t) - \log(z_{t-1})$, where $z_t$ is the price levels of black pepper. By examining EACF of $x_t$, AR(1) model is fitted. The estimated
residuals are denoted as $\hat{y}_t$. For the pepper price the returns have substantial first order autocorrelation, which can be removed by fitting an AR(1) model to this series.

On examining the estimated residuals of the AR(1) model it is found that the series $\hat{y}_t$ is having the ARCH effect. So the current series may be described by the AR(1)-ARCH(2) model, as the first two estimated autocorrelations of $\hat{y}_t^2$ are significant. Let the model is

$$x_t - \phi_0 - \phi_1 x_{t-1} = y_t = \sqrt{h_t} \eta_t, \quad h_t = a_0 + a_1 y_{t-1}^2 + a_2 y_{t-2}^2, \quad \eta_t \sim \text{NID}(0,1)$$

**Table 5.2**: The parameter estimate of the above model is given in the following table:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi_0$</td>
<td>-0.000907</td>
<td>(0.002350)</td>
</tr>
<tr>
<td>$\phi_1$</td>
<td>0.321</td>
<td>(0.067)</td>
</tr>
<tr>
<td>$a_0$</td>
<td>0.000607</td>
<td>(5.73E-05)</td>
</tr>
<tr>
<td>$a_1$</td>
<td>0.180</td>
<td>(0.072)</td>
</tr>
<tr>
<td>$a_2$</td>
<td>0.139</td>
<td>(0.072)</td>
</tr>
</tbody>
</table>

**Figure 5.2**: Estimated conditional standard deviation from an ARCH(1) model.
6. ARCH-in-Mean (ARCH-M) Model

The ARCH-M model is often used in financial applications where the expected return on an asset is related to the expected asset risk. If we introduce the conditional variance into the mean equation, we get the ARCH-in-Mean (ARCH-M) model. ARCH-M may be represented as follows:

\[ y_t = \mu_t + \varepsilon_t, \quad h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2, \quad \mu_t = \beta + \delta \sqrt{h_t}. \]

Some times the mean equation and the variance equation may contain the regressor variables.

6.1 AR(p)-ARCH(q) in Mean Model

Here the time-series is autoregressive with order p and the errors are following the ARCH model with order q. The conditional mean is the linear function of conditional standard deviation. AR(p)-ARCH(q) in Mean model is defined as

\[ y_t = \mu_t + \sum_{i=1}^{p} \phi_i y_{t-i} + \varepsilon_t, \]

\[ \mu_t = \beta + \delta \sqrt{h_t}, \quad h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2, \quad \varepsilon_t = \sqrt{h_t} \eta_t, \quad \eta_t \sim \text{NID}(0,1). \]

Example 6.1: As exhibited in Figure 6.1, the wholesale monthly onion price per quintal fluctuated between Rs.100/- and Rs.300/- during April, 1996 to June, 1998 except for a sudden burst to a level of Rs.550/- in the month of February, 1998. The price increased steadily in the second half of 1998 and touched the maximum of Rs.1000/- during October, 1998 to December, 1998. The price subsequently decreased in the range of Rs.125/- to Rs.500/- during the later time period till October, 2001. Thus, the data clearly indicates presence of a quadratic trend.
Ghosh et al. (2005) found that the assumption that conditional variance of one-period ahead forecast error of detrended and deseasonalised residual series is constant does not hold. Better forecast results are expected if additional information allowing conditional variance to depend upon the past realization is taken into account. This can only be incorporated through Autoregressive conditional heteroscedastic (ARCH) process in nonlinear time-series modeling.

They pointed out that the current price of onion depends not only on its price in the past but also on risk and uncertainty involved in onion prices. Accordingly, AR(p)-ARCH(q)-in-Mean model, was developed.

**Estimation and Forecasting**

Generally, parameters in mean equation are estimated by ordinary least squares and residuals are then computed. From these residuals, efficient estimates in variance equation can be constructed. Based upon these estimates, efficient estimates of parameters in mean as well as variance equations are then obtained. For the present onion price data, estimates of parameters in AR(5)-ARCH(4) model are obtained as

\[ \hat{\phi}_1 = -0.11, \quad \hat{\phi}_2 = -0.05, \quad \hat{\phi}_3 = -0.03, \quad \hat{\phi}_4 = 0.31, \quad \hat{\phi}_5 = 0.38 \]

\[ \hat{\alpha}_0 = 3341.24, \quad \hat{\alpha}_1 = 0.06, \quad \hat{\alpha}_2 = 0.26, \quad \hat{\alpha}_3 = 0.10, \quad \hat{\alpha}_4 = 0.43 \]

However, a perusal of Q-statistic values of standardized residuals and squared standardized residuals reveal that there are misspecifications in both mean and variance equations. Accordingly, attempts were made towards incorporating ARCH–in-Mean with market availability data \{av_t\} as regressor in variance equation as it is well recognized that there is a high negative correlation between onion price and its availability in the market. The market availability data, as an exogenous variable, also has an impact on the one-period ahead forecast variance of onion price.
Therefore, incorporating market availability series \{av_t\} as regressor in variance equation and employing the Numerical derivative approach, as available in EViews (2000) package, final estimated AR(5) - ARCH(4) model was obtained. Using EViews (2000) package, the Q-statistic values in respect of standardized residuals as well as squared standardized residuals are now seen to be insignificant confirming thereby proper identifications of mean and variance equations. In order to get a visual idea, the graph of fitted ARCH model along with data is depicted in Figure 6.2. Evidently, the fit of ARCH model is good.

![Graph of fitted ARCH model](image)

**Figure 6.2**: Fitting AR(5)-ARCH(4)- in- Mean model to onion price along with datapoints

**Table 6.1**: Forecast values of wholesale monthly onion price data (Rs. per quintal) on the basis of fitted ARCH models

<table>
<thead>
<tr>
<th>Month</th>
<th>Actual</th>
<th>Forecast value by ARCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>November, 2001</td>
<td>500</td>
<td>538.16 (45.83)</td>
</tr>
<tr>
<td>December, 2001</td>
<td>300</td>
<td>380.00 (43.90)</td>
</tr>
<tr>
<td>January, 2002</td>
<td>250</td>
<td>315.81 (45.18)</td>
</tr>
<tr>
<td>February, 2002</td>
<td>250</td>
<td>270.34 (29.85)</td>
</tr>
</tbody>
</table>

### 7. Limitation of ARCH Models

The model assumes that positive and negative shocks have the same effects on volatility because it depends on the square of the previous shocks. In practice, it is well known that price of a financial asset responds differently to positive and negative shocks. Further, it is not a parsimonious representation of the model and requires estimation of a large number of parameters.

### 8. The Generalized ARCH Model

Bollerslev (1986) proposed Generalized ARCH (GARCH) model, in which conditional variance is also a linear function of its own lags and has the form

\[
h_t = a_0 + \sum_{i=1}^{q} a_i y_{t-i}^2 + \sum_{j=1}^{p} b_j h_{t-j}
\]

\[
... (8.1)
\]
The conditional variance defined by (8.1) has the property that the unconditional autocorrelation function of $y_t^2$; if it exists, can decay slowly. For the ARCH family, the decay rate is too rapid compared to what is typically observed in financial time-series, unless the maximum lag $q$ is long. As (8.1) is a more parsimonious model of the conditional variance than a high-order ARCH model, most users prefer it to the simpler ARCH alternative. The overwhelmingly most popular GARCH model in applications has been the GARCH $(1, 1)$ model, that is, $p = q = 1$ in (8.1).

A sufficient condition for the conditional variance to be positive with probability one is $a_0 > 0, a_i \geq 0, i = 1, 2, ..., q, b_j \geq 0, j = 1, 2, ..., p$. The GARCH $(p, q)$ process is weakly stationary if and only if $\sum_{i=1}^{q} a_{i} + \sum_{j=1}^{p} b_{j} < 1$.

To study the property of GARCH models, it is advantageous to use the following representation. Let $\eta_t = y_t^2 - h_t$ so that $h_t = y_t^2 - \eta_t$. By plugging $h_{t-i} = y_{t-i}^2 - \eta_{t-i}$ $(i=0,1,\ldots,q)$ into Equation (8.1) we can rewrite the GARCH model as

$$y_t^2 = a_0 + \sum_{i=1}^{\text{Max}(p,q)} (a_i + b_i) y_{t-i}^2 + \eta_t + \sum_{j=1}^{p} b_j \eta_{t-j} \quad \ldots(8.2)$$

It is easy to check that the $\{\eta_t\}$ is a martingale difference series (i.e., $E(\eta_t) = 0$ and $\text{cov}(\eta_t, \eta_{t-j}) = 0$) for $j \geq 1$). However, $\{\eta_t\}$ in general is not an i.i.d sequence. Equation (8.2) is an ARMA form for the squared series $y_t^2$. Thus a GARCH model can be regarded as an application of the ARMA idea to the squared series $y_t^2$. Using the unconditional mean of an ARMA model, we have $E(y_t^2) = \frac{a_0}{1 - \sum_{i=1}^{\text{Max}(p,q)} (a_i + b_i)}$ provided that the denominator of the prior fraction is positive.

9. Recent Advances

Several extensions of GARCH model have been proposed. For example, to allow the asymmetric effect of positive and negative shocks, Exponential GARCH (EGARCH) model may be used. The other important extension, viz. Threshold GARCH (TGARCH) may be employed to handle the leverage effects. The volatility seems to react differently to a big price increase or a big price drop, referred to as the leverage effect. However, application of these models is extremely cumbersome.

10. Conclusions

When volatility is present in a data set, we can use ARCH model for describing it. Sometimes, important regressor variables need to be introduced in the mean and variance equations, thereby resulting in AR-ARCH-in-Mean models. Although some isolated attempts have been made to apply above types of models to data from the field of agriculture, there is a need to make more concerted efforts in this direction. Further, extensions of these models, viz. GARCH, EGARCH and TGARCH models may also be applied in due course of time.
References