Statistical Modelling of Volatility in Financial Time Series

Balakrishna, N.

Department of Statistics
Cochin University of Science and Technology,
Kochi, 682 022,

July 3, 2017
Outline

- Introduction - Time Series
- Box Jenkins Methods
- Financial Time Series
- Observation-driven models - ARCH/GARCH
- Parameter-driven models
- Examples and Illustrations
Assume that the observed time series (data) \( \{z_t\} \) is a realization of a discrete time stochastic process.

The sequence of errors \( \{Z_t\} \) is not iid, but stationary.

Then apply the results from the theory of stationary stochastic process to analyze the model.

Examine the dependence structure and identify an appropriate model.

Estimate the parameters of the model and check for adequacy.

Use for forecasting if it is found good for the data.
Characteristics of a Time Series

Mean function, \( m(t) = E(Z_t) \)

Variance function, \( V(t) = E(Z_t - m(t))^2 \)

Covariance function, \( \gamma_{t,s} = Cov(Z_t, Z_s) = E(Z_t - m(t))(Z_s - m(s)) \)

Correlation function, \( \rho_{t,s} = \frac{\gamma_{t,s}}{\sqrt{V(t)V(s)}}. \)

Autocovariance of order \( k \) for the series \( \{Z_t\} \) is
\( \gamma_k = Cov(Z_t, Z_{t-k}) = E(Z_t - m(t))(Z_{t-k} - m(t - k)) \) and the autocorrelation function, ACF of order \( k \)
\( \rho_k = \frac{\gamma_k}{\sqrt{V(t)V(t-k)}}. \)

PACF = \( Corr(Z_t, Z_{t-k}) \), by fixing the effect of \( Z_{t-1}, Z_{t-2}, \ldots, Z_{t-k+1} \)
Stationary Time series

A **Stochastic process** \( \{ Z_t \} \) is

- **strictly stationary** if the joint distribution of \( (Z_{t_1}, Z_{t_2}, \ldots, Z_{t_n}) \) and \( (Z_{t_1+h}, Z_{t_2+h}, \ldots, Z_{t_n+h}) \) remain same for every \( t_1, t_2, \ldots, t_n \) and \( h \) in the parameter(time) space \( T \).

- **weakly stationary** if \( E(Z_t) = m, \) a constant, \( Var(Z_t) = \sigma^2, \) finite and the \( Cov(Z_t, Z_s) \) is a function of \( |s - t| \) only.

- A sequence of uncorrelated rvs \( \{ a_t \} \) with mean zero and constant variance is referred to as a **white noise.**
Correlogram

- The SACF is taken as an estimate of the ACF
- Plot of SACF as a function of lag is the correlogram
- Used for identifying the dependence structure
- For a stationary time series, SACF either geometrically decreases or oscillates between negative and positive values in a geometrically decreasing envelope
- Helps identifying the order of dependence
- Similarly sample PACF also useful to analyze the dependence structure
ARMA($p,q$) Model

\[ Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} - \ldots - \phi_p Z_{t-p} = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q} \]

\[ \Rightarrow \]

- The model is stationary if AR component is stationary and invertible if its MA component is so
- More detailed analysis may be carried out for specified values of $p$ and $q$
Given time series \( \{Z_t\} \) is not stationary

- It becomes stationary \( \text{ARMA} \) process after a finite number \((d)\) of differences

**Notation:** \( Z_{t-1} = BZ_t \), \( B \) is the back shift operator
\[
Z_t - Z_{t-1} = Z_t - BZ_t = (1 - B)Z_t = \nabla Z_t,
\]
\( \nabla = 1 - B \) is the difference operator, \( \nabla^2 Z_t = \nabla(\nabla Z_t) \)

- If \( d = 1 \) then \( Z_t - Z_{t-1} = \nabla Z_t \) becomes stationary \( \text{ARMA} \)

- \( \{Z_t\} \) is an \( \text{ARIMA}(p, d, q) \) process if \( \{\nabla^d Z_t\} \) is a stationary \( \text{ARMA}(p, q) \) process
Test for Stationarity

The stationarity of a given time series may be examined by plotting the graphs of

- The series as a function of time
- Sample autocorrelation function
Introduction to Financial time series

Let \( \{P_t = S(lt)\} \) be the price of an asset at time \( lt \),
\( t = 1, 2, \ldots, \) where \( l \) is the interval between the observations. Then the sequence of return is defined by
\[
Y_t = \ln\left(\frac{P_t}{P_{t-1}}\right), \quad t = 1, 2, \ldots.
\]

At any given point of time, \( t \), \( Y_t \) is a random variable and hence we have a time series \( \{Y_t, t > 0\} \) of returns.
Example: $S&P500$ price index series

- The data is the daily closing values of the $S&P500$ index on trading days from 4th April 2010 to 4th October 2013
- Following are the relevant plots of the time series
  - Price series
  - Log returns and squared log returns
  - Autocorrelation function (ACF) of the returns and squared returns
Sample ACF of S&P 500 Log return series
Time series of squared log returns (S&P500)
Sample ACF of the squared log returns
Density histogram of the S&P500 log returns

Log returns vs Frequency

Frequency distribution of log returns for the S&P500 index.
Statistical Modelling of Volatility in Financial Time Series

Descriptive statistics of S&P500 log return series

\( y_t = \) Log return series

- Mean \((y_t) = 0.00042345\)
- Variance \((y_t) = 0.00011913\)
- Standard deviation = 0.0109
- Min = -0.0690, Max = 0.0463
- Kurtosis = 7.1840
- Excess kurtosis = 4.1840
Empirical studies on financial time series show that $\{Y_t, t > 0\}$ is characterized by

- Absence of autocorrelation (serial correlation)
- Squared returns exhibit significant serial correlation
- The marginal variables follow heavy-tailed distribution, kurtosis $> 3$
- Conditional variance of $Y_t$ given the past observations is not constant, (heteroskedasticity)
- Evidence of volatility clustering: ie, Low values of volatility followed by low values and high values of volatility followed by high values.
The linear time series models available in the literature such as ARIMA/ARMA are not suitable to model the financial time series with the above features. So new classes of models are introduced to deal with such time series.

Before trying such models for a given financial data we need to check if it possesses some of the above properties.

Some of the properties can be visualized through graphical methods.

An authentic statement requires a statistical test.
Engle (1982) introduced ARCH models to capture volatility clustering in time series data.

An ARCH($p$) model assumes that

$$Y_t - \mu_t(y_{t-1}, y_{t-2}, ...) = \sigma_t \varepsilon_t, \text{ with } \sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i Y_{t-i}^2,$$

where $\mu_t(y_{t-1}, y_{t-2}, ...) = E(Y_t|y_{t-1}, y_{t-2}, ...)$, and $\{\varepsilon_t\}$ is a sequence of independent and identically distributed (iid) random variables with mean zero and variance 1, $\omega > 0$, and $\alpha_i \geq 0$ for $i > 0$, $\alpha_1 + \alpha_2 + \ldots + \alpha_p < 1$.

The conditional mean $\mu_t(y_{t-1}, y_{t-2}, ...)$ can be estimated by Box-Jenkin’s method.

Our focus is to model the conditional variance, $\sigma_t^2$. So we set $\mu_t(y_{t-1}, y_{t-2}, ...) = 0$. 
The ARCH(1) model is defined by

\[ Y_t = \sigma_t \varepsilon_t , \quad \sigma_t^2 = \omega + \alpha_1 Y_{t-1}^2 \]  

where \( \omega > 0 \) and \( 0 \leq \alpha_1 < 1 \).

The properties of the model are

1. The unconditional mean of \( Y_t \) is zero, since

\[ E(Y_t) = E(E(Y_t|Y_{t-1})) = E(\sigma_t E(\varepsilon_t)) = 0 \]
2. The conditional variance of $Y_t$ is

$$E (Y_t^2|Y_{t-1}) = \omega + \alpha_1 Y_{t-1}^2 = \sigma_t^2$$

3. The unconditional variance of $Y_t$ is

$$Var (Y_t) = \frac{\omega}{1 - \alpha_1}$$
4. Assuming that the fourth moment of $Y_t$ are finite, the Kurtosis $K_Y$ of $Y_t$, is given by

$$K_Y = \frac{E(Y_t^4)}{E(Y_t^2)^2} = 3 \frac{1 - \alpha_1^2}{1 - 3\alpha_1^2} > 3$$

provided $\alpha^2 < 1/3$.

5. The autocorrelation function (ACF) of $Y_t$ is zero. The ACF of $\{Y_{t}^2\}$ is $\rho_{Y_{t}^2}(k) = \alpha_1^k$ and notice that $\rho_{Y_{t}^2}(k) \geq 0$ for all $k$, a result which is common to all linear ARCH models.
Properties of $ARCH(p)$ models

- Representation of ARCH in the form of an AR model helps in utilizing the established properties the latter to analyze the former.

- An $ARCH(p)$ model may be expressed as an $AR(p)$ in $Y_t^2$:
  $Y_t^2 = \omega + \alpha_1 Y_{t-1}^2 + \alpha_2 Y_{t-2}^2 + \ldots + \alpha_p Y_{t-p}^2 + u_t$

- From this we have
  $E(Y_t) = 0$, $Var(Y_t|y_{t-1}, y_{t-2}, \ldots) = \sigma_t^2$
  $Var(Y_t) = \frac{\omega}{1-\alpha_1-\alpha_2-\ldots-\alpha_p}$

- ACF of $\{Y_t\}$ vanishes.

- ACF of $\{Y_t^2\}$ can be obtained by solving the Yule-Walker’s equations in terms of the AR coefficients.
GARCH Models

- GARCH model is an extension of ARCH by Bollerslev (1986).
- The model allows the conditional variance to depend on the past conditional variance and the squares of past returns.
- The GARCH \((p, q)\) is defined by

\[ Y_t = \sigma_t \varepsilon_t , \quad \sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i Y_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \]

where \(\{\varepsilon_t\}\) is a sequence of iid random variables with mean 0 and variance 1, \(\omega > 0, \alpha_i \geq 0, \beta_j \geq 0\), and \(\sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) < 1\).
GARCH(1,1)

\[ Y_t = \sigma_t \varepsilon_t \]
\[ \sigma_t^2 = \omega + \alpha_1 Y_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \]
where \(0 \leq \alpha_1, \beta_1 \leq 1, \alpha_1 + \beta_1 < 1\).

Literature on financial time series analysis stress that a GARCH(1,1) model with only three parameters in the conditional variance equation is adequate to obtain a good model fit.

So for all practical purposes, GARCH(1,1) model is good enough, see Hansen and Lunde (2004).
Properties of the GARCH(1,1) model are

- The unconditional mean of $Y_t$ is zero, since

$$E(Y_t) = E(E(Y_t|Y_{t-1})) = E(\sigma_tE(\varepsilon_t)) = 0$$

- The conditional variance of $Y_t$ is

$$E(Y_t^2|Y_{t-1}) = \omega + \alpha_1 Y_{t-1}^2 + \beta_1 \sigma_{t-1}^2 = \sigma_t^2$$

- The unconditional variance of $Y_t$ is

$$Var(Y_t) = \frac{\omega}{1 - (\alpha_1 + \beta_1)}$$
The Kurtosis of $Y_t$, $K_Y$, is given by

$$K_Y = \frac{3 \left[ 1 - (\alpha_1 + \beta_1)^2 \right]}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2} > 3$$

Consequently, similar to ARCH models, the tail distribution of GARCH(1,1) process is heavier than that of a normal distribution if $1 - 2\alpha_1^2 - (\alpha_1 + \beta_1)^2 > 0$.

The ACF of $\{Y_t\}$ is zero and the ACF of $\{Y_t^2\}$ is given by

$$\rho_{Y_t^2}(k) = (\alpha_1 + \beta_1)^{k-1} \frac{\alpha_1 \left( 1 - \alpha_1 \beta_1 - \beta_1^2 \right)}{1 - 2\alpha_1 \beta_1 - \beta_1^2}$$

$k = 1, 2, ...$
ACF of the Squared GARCH(1,1) series

![ACF of the Squared GARCH(1,1) series graph]

- a1=.5, b1=.3
- a1=.6, b1=.3
- a1=.65, b1=.3
ACF of the Squared GARCH(1,1) series
If the errors are standard normal, then the likelihood function of an GARCH\((p, q)\) model is

\[
L(\theta|y_1, y_2, \ldots, y_n) = \prod_{t=p+1}^{n} \frac{1}{\sqrt{2\pi \sigma_t^2}} \exp \left( -\frac{y_t^2}{2\sigma_t^2} \right) f(y_1, y_2, \ldots, y_p|\theta)
\]

where \(\theta = (\omega, \alpha_1, \ldots, \alpha_p, \beta_1, \beta_2, \ldots, \beta_q)'\)

\(f(y_1, y_2, \ldots, y_p|\theta)\) is the joint probability density function of \(y_1, y_2, \ldots, y_p\).
The conditional-likelihood function becomes
\[ L(\theta; y_{p+1}, y_{p+2}, \ldots, y_n | y_1, y_2, \ldots, y_p) = \prod_{t=p+1}^{n} \frac{1}{\sqrt{2\pi \sigma_t^2}} \exp \left( -\frac{y_t^2}{2 \sigma_t^2} \right). \]  

(3)

The conditional log-likelihood function is
\[ l(\theta; y_{p+1}, y_{p+2}, \ldots, y_n | y_1, y_2, \ldots, y_p) = \sum_{t=p+1}^{n} \left( -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(\sigma_t^2) - \frac{y_t^2}{2 \sigma_t^2} \right). \]  

(4)

Maximizing the log-likelihood function with respect to the parameters we get the MLEs.

Errors may also be assumed to follow other symmetric distributions such as Laplace, student-t, GED, etc.
Drawbacks of GRACH models

- The GARCH models are unable to represent volatility asymmetry.
- Due to the presence of lagged $Y_t^2$ in the variance equation, the positive and negative values of the lagged innovations have the same effect on the conditional variance.
- In the finance literature, however, it has been recognized that volatility often responds to positive and negative shocks in different ways.
- For example, the volatility of stock returns tends to increase (decrease) when there is “bad news” ( “good news”).
- To ensure positiveness of $h_t$ in the GARCH model, non-negative constraints are imposed on the coefficients in the variance equation.
Several generalizations of GARCH models are available in the literature. One may find the details in the following references

The parameter driven models assume that the volatilities are generated by some latent models, in terms of unobservable variables. The log-normal stochastic volatility (SV) model by Taylor (1986) is the simplest and the best known example:

\[ Y_t|h_t \sim N(0, \exp(h_t)) , \quad h_t = \alpha + \beta h_{t-1} + \eta_t \]

where \( h_t \) represents the log-volatility, which is unobserved but can be estimated using the observations. Several parametrization of this model are available in the literature.
The parameter driven model, may be assumed to have a structure, where the returns $Y_t$ are generated by the model of the form:

$$Y_t = \epsilon_t \sqrt{X_t}$$

with

$$X_t = e^{h_t}, \ h_t = \alpha_0 + \alpha_1 h_{t-1} + \eta_t,$$

where $\{\epsilon_t\}$ is iid symmetric with $D(0,1)$, $\{\eta_t\}$ is iid independent of $\{\epsilon_t\}$. $\{Y_t\}$ so defined possesses the special features of a financial time series and is referred to as **Stochastic Volatility Model (SVM)**.
Normal-lognormal SVM

This one of the widely studied SVM, where

\[ Y_t = \beta \varepsilon_t \exp \left( \frac{h_t}{2} \right), \quad h_t = \mu + \phi(h_{t-1} - \mu) + \eta_t, \] (5)

where \( \beta = e^{\mu/2} \), \( \varepsilon_t \) and \( \eta_t \) are two independent Gaussian white noises, with variances 1 and \( \sigma_{\eta}^2 \), respectively.

Due to the Gaussianity of \( \eta_t \), this model is called a log-normal SV model.
Second order properties

As $\eta_t$ is Gaussian, $\{h_t\}$ is a standard Gaussian AR(1) process, it will be (strictly and covariance) stationary if $|\phi| < 1$ with:

$$\mu_h = E(h_t) = \mu, \sigma_h^2 = Var(h_t) = \frac{\sigma_{\eta}^2}{1 - \phi^2}$$

As $\{\varepsilon_t\}$ is always stationary, $\{Y_t\}$ will be stationary if and only if $\{h_t\}$ is stationary, $Y_t$ being the product of two stationary process.

The odd moments of $Y_t$ of all order vanish. Using the properties of lognormal distribution, if $r$ is even, all the moments exist if $h_t$ is stationary and are given by

$$E(Y_t^r) = E(\varepsilon_t^r)E\left(\exp\left(\frac{r}{2}h_t\right)\right) = \frac{r!e^{\frac{r\mu}{2}}}{2^{\frac{r}{2}}(\frac{r}{2})!} \exp\left(\frac{r}{2}\mu_h + \frac{r^2\sigma_h^2}{8}\right)$$
Hence the kurtosis is

\[
\frac{E(Y_t^4)}{E(Y_t^2)^2} = 3 \exp(\sigma_h^2) \geq 3
\]

which shows that the SV model has fatter tails than the corresponding normal distribution.

The ACF of \(\{Y_t^2\}\) is

\[
\rho_{Y_t^2}(k) = \frac{Cov(Y_t^2, Y_{t-k}^2)}{Var(Y_t^2)} = \frac{\exp(\sigma_h^2 \phi^k) - 1}{3 \exp(\sigma_h^2) - 1} \approx \frac{\exp(\sigma_h^2) - 1}{3 \exp(\sigma_h^2) - 1} \phi^k.
\]
Some Observations

- The dependence structure of $\{Y_t\}$ is decided by that of $\{h_t\}$
- If $\phi < 0$, $\rho_{Y^2}(k)$ can be negative, unlike ARCH models.
- There is no need for non-negativity constraints or for bounded kurtosis constraints on the coefficients.
- This is an advantage over ARCH/GARCH models
- The problem of estimation is more complicated in SVM compared to that in ARCH/GARCH models.
- Likelihood based inference works well in ARCH/GARCH set up, but not in SVM as we see next.
Estimation

- A major problem here is to estimate $\theta = (\mu, \phi, \sigma^2_\eta)$.
- The likelihood function of $\theta$ based on $y = (y_1, y_2, \ldots, y_T)$ is the pdf of $y$ given $\theta$.
- But, from the model structure, it is clear that the likelihood function will depend on the unknown vector $h = (h_1, h_2, \ldots, h_T)$.
- We may get an explicit form of the likelihood function by integrating out the latent variables. Thus

$$L(\theta) = f(y|\theta)$$

$$= \int f(y, h|\theta) dh$$

$$= \int_{h_T} \int_{h_{T-1}} \ldots \int_{h_1} f(y_1, y_2, \ldots, y_T, h_1, h_2, \ldots, h_T|\theta) dh_1 dh_2 \ldots dh_T$$
The multiple integral $L(\theta)$ cannot be factored into a product of $T$ one-dimensional integrals due to the dependence of $h_t$ on the past.

The difficulties in obtaining explicit forms of MLEs, lead to several numerical methods.

Markov Chain Monte Carlo (MCMC) is commonly used method for numerical estimation.

Several algorithms for this method are also available in the literature.

Bayesian analysis of the lognormal SV model is studied by Jacquier, Polson and Rossi (1994) (JPR).

Generalized Method of Moments (GMM) is applied whenever the moments of several order have analytical expressions.
Remark:

- The structure of an SVM allows us to interpret it as specifying a prior over the sequence of \{h_t\}.
- The prior is that the volatilities evolve according to an AR(1) process.
- Note that the likelihood function \( L(\theta) = \int f(y|h)f(h|\theta)dh \) may be treated as the expectation of \( f(y|h) \) with respect to the prior distribution of \( h|\theta \).
- Recall that an integral can be approximated by a sum over a proper partition of the range of the integral.
- This motivates us to approximate the likelihood function of SVM by an appropriate sum and then maximize by numerical methods.
The SV model described above insists that the log-volatility should have normal distribution, but need not be true in reality.

Abraham, Balakrishna and Sivakumar (2006) studied an SV model $Y_t = \epsilon_t \sqrt{X_t}$ where $\{X_t\}$ is a stationary gamma sequence defined by first order gamma AR (GAR(1)).
Let $Y_t$ be the return on an asset at time $t$, $t = \pm 1, \pm 2, \ldots$. Define

$$Y_t = \varepsilon_t \sqrt{X_t},$$

(6)

We discuss the properties of $\{Y_t\}$ when $\{X_t\}$ is a Markov sequence with stationary gamma marginal density function:

$$f(x; \lambda, p) = \frac{e^{-\lambda x} x^{p-1} \lambda^p}{\Gamma(p)}, \quad x \geq 0, \lambda > 0, p > 0$$
If we assume a $G(\lambda, p)$ distribution for $X_t$ in (13) then the characteristic function (cf) of $Y_t$ is given by

$$
\phi(s) = E \left[ e^{isY_t} \right] = \left( \frac{2\lambda}{2\lambda + s^2} \right)^p.
$$

(7)

This is the cf of the difference of two iid gamma r.v.'s with parameters $\sqrt{2\lambda}$ and $p$.

The distribution of such r.v.'s are referred to as generalized Laplace distributions.

At $p = 1$, the density corresponding to (14) becomes

$$
f_1(y) = \sqrt{\frac{\lambda}{2}} \exp \left\{ -|y| \sqrt{2\lambda} \right\}, \quad -\infty < y < \infty, \quad \lambda > 0,
$$

which is the Laplace pdf.
The odd moments of $Y_t$ are zero and its even moments are given by

$$E [Y_t^{2r}] = (2r - 1)(2r - 3) \cdots 3 \cdot 1 \cdot \frac{\Gamma(p + r)}{\Gamma(p)} \lambda^{-r}, \quad r = 1, 2, \ldots.$$ 

The kurtosis of $y_t$ becomes

$$\gamma = \frac{E(Y_t^4) - [E(Y_t)]^4}{[var(Y_t)]^2} = 3 + \frac{3}{p} > 3.$$ 

Note that if $p = 1$, then $\gamma = 6$ which is the kurtosis of a Laplace distribution and as $p \to \infty$, $\gamma \to 3$, the one corresponding to a normal distribution.
Advantages of Gamma SV

- By choosing smaller $p$, one can get a distribution with larger kurtosis. So $Y_t$ has a leptokurtic marginal distribution.
- The literature on financial time series indicates that the return series always shows the tendency to follow leptokurtic distributions.
- Hence the generalized Laplace distribution is a good candidate for modeling such data.
- To generate such sequences, we need models to generate the dependent sequences of gamma rvs.
GAR(1) Model

- GAR(1), Gaver and Lewis (1980)

\[ X_t = \rho X_{t-1} + \eta_t, \quad 0 \leq \rho < 1, \ t = 1, 2, \ldots, \]

where \( \eta_t = \sum_{j=1}^{N} \rho^{U_j} E_j, \ E_j \sim \exp(\lambda) \) and \( U_j \sim U(0, 1) \).

- \( N \) is Poisson with mean \( p \log (1/\rho) \).

- The r.v’s \( U_j, E_j \) and \( N \) are mutually independent for every \( j \).

- If \( N \equiv 0 \) then we take \( \eta_j \equiv 0 \).

- Then, \( \{X_t\} \) defined above is stationary and has \( G(\lambda, \rho) \) as the marginal distribution.
2. Parameter Estimation

- The likelihood function involves the unobservable Markov dependent latent variables.
- These variables have to be integrated out using multiple integrals and this complicates the parameter estimation by the method of maximum likelihood.
- Moreover, in the present case, the probability density function of $\eta_t$ in expression (11) does not have a closed form.
- In view of this
The model parameters are estimated using the method of moments.

The resulting estimators perform well for larger sample sizes, confirmed by simulation. Theoretically supported by the generalized method of moments introduced by Hansen (1982).

The Gamma SV model is applied to analyze the stock price index returns of Canada (TSE300), Japan (TOPIX), the UK (FTSE100) and the USA (S&P500).

It is observed that Gamma Stochastic Volatility model captures the kurtosis of the marginal distribution better than other standard models.

See the following Table. Details in Abraham, Balakrishna and Sivakumar (2006).
Gamma SV Model for stock price index

Kurtosis of stock returns and the estimated models

<table>
<thead>
<tr>
<th>Country</th>
<th>Data</th>
<th>GARCH(1,1)</th>
<th>Gamma SV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada</td>
<td>7.4606</td>
<td>5.4639</td>
<td>7.3789</td>
</tr>
<tr>
<td>Germany</td>
<td>5.1075</td>
<td>5.8098</td>
<td>5.0382</td>
</tr>
<tr>
<td>Japan</td>
<td>5.1656</td>
<td>4.9674</td>
<td>5.1481</td>
</tr>
<tr>
<td>U.S.</td>
<td>5.7057</td>
<td>4.4596</td>
<td>5.6650</td>
</tr>
</tbody>
</table>
Some Recent Work
Gumbel Extreme Value SV Model

Model and Properties

Let \( \{y_t\} \) be a sequence of returns on certain financial asset and the volatilities are generated by a Markov sequence \( \{\exp(h_t)\} \) of non-negative rvs. Define the SV model

\[
y_t = \exp\left(\frac{h_t}{2}\right) \varepsilon_t
\]

(8)

\[
h_t = \alpha h_{t-1} + \eta_t, \quad t = 1, 2, ..., \quad 0 < \alpha < 1
\]

(9)

where \( \{\varepsilon_t\} \) is a sequence of independent and identically distributed (iid) standard normal random variables (rvs).
The sequence \( \{\varepsilon_t\} \) is independent of \( h_t \) and \( \eta_t \) for every \( t \).

Here we assume that for every \( t \), the volatility, \( h_t \) is a GEV rv with probability density function (pdf)

\[
    f_{h_t}(x; \mu, \sigma) = \frac{1}{\sigma} \exp \left( \frac{x - \mu}{\sigma} \right) \exp \left( -\exp \left( \frac{x - \mu}{\sigma} \right) \right),
\]

\(-\infty < x < \infty, -\infty < \mu < \infty, \sigma > 0.\)

If \( \{h_t\} \) follows an Extreme value distribution, then \( \exp(h_t) \) follows a Weibull distribution.
In order to have this marginal distribution for \( \{ h_t \} \) defined by (9) we need to have the distribution of \( \eta_t \) expressed by

\[
\eta_t \overset{L}{=} (1 - \alpha) \mu - \sigma Z, \text{ and } Z \overset{L}{=} - \log (U^{-\alpha})
\]

where \( L \) denotes the equality in distribution, \( U \) denotes a positive stable rv with Laplace transform

\[
\varphi(s) = e^{-s^\alpha}, \ 0 < \alpha < 1.
\]

The mean and variance of \( \{ \eta_t \} \) are respectively given by

\[
E(\eta_t) = (1 - \alpha) (\mu - \sigma \gamma) = \mu^*; \ V(\eta_t) = (1 - \alpha^2) \frac{\pi^2 \sigma^2}{6} = \sigma^{2*},
\]

where \( \gamma \approx 0.5772 \) is the Euler’s constant.
Since the sequence \( \{ \varepsilon_t \} \) follows standard normal distribution, the odd moments of \( y_t \) are zero and its even moments are given by

\[
E \left( y_t^{2r} \right) = (2r - 1) (2r - 3) \ldots 3.1 \cdot e^{r \mu} \Gamma (r \sigma + 1), \quad r = 1, 2, \ldots.
\]

Then \( V (y_t) = e^{\mu} \Gamma (\sigma + 1) \) and the kurtosis of \( y_t \) becomes

\[
K_r = 3 \frac{\Gamma (2 \sigma + 1)}{[\Gamma (\sigma + 1)]^2}. \quad (13)
\]

By choosing different values for \( \sigma \), one can get a distribution with larger kurtosis which makes \( r_t \) suitable for modelling the financial returns.
Figure: The plot of kurtosis $K$ of $y_t$. 
The structure of the model (8) implies that autocorrelation function (ACF) of \( \{y_t\} \) is zero and that of \( \{y_t^2\} \) is significant.

The variance and covariance function of the squared return series are obtained as

\[
V(y_t^2) = E(y_t^4) - (E(y_t^2))^2 \\
= e^{2\mu} \left( 3 \Gamma(2\sigma + 1) - [\Gamma(\sigma + 1)]^2 \right).
\]

\[
\gamma_{y_t^2}(k) = e^{2\mu} \left\{ \frac{\Gamma(\sigma + 1) \Gamma[(\alpha^k + 1)\sigma + 1]}{\Gamma(\alpha^k\sigma + 1)} - [\Gamma(\sigma + 1)]^2 \right\}.
\]
Hence the lag $k$ autocorrelation of the squared sequence $\{y^2_t\}$ is

$$\rho_{y^2_t} (k) = \text{Corr} (y^2_t, y^2_{t-k}) = \frac{\Gamma (\sigma + 1) \Gamma [(\alpha^k + 1) \sigma + 1] - [\Gamma (\sigma + 1)]^2 \Gamma (\alpha^k \sigma + 1)}{\Gamma (\alpha^k \sigma + 1) \left\{ 3 \Gamma (2\sigma + 1) - [\Gamma (\sigma + 1)]^2 \right\}}$$

(14)

The ACF is an exponentially decreasing function of the lags for different values of the parameters, as can be seen in figure 2.
Figure: The ACF of squared return for different combinations of the parameters

- Sigma=3.0, Alpha=0.9
- Sigma=1.5, Alpha=0.7
- Sigma=2.0, Alpha=0.5
Let \( (y_1, y_2, \ldots, y_T) \) be a realization of length \( T \) from the GEV-SV model (8) and \( \theta = (\mu, \sigma, \alpha) \) be the parameter vector to be estimated.

We use the moments

\[
E(y_t^2) = e^{\mu} \Gamma(\sigma + 1),
\]

\[
E(y_t^4) = 3e^{2\mu} \Gamma(2\sigma + 1)
\]

\[
E(y_t^2 y_{t-1}^2) = e^{2\mu} \frac{\Gamma(\sigma + 1) \Gamma[(\alpha + 1)\sigma + 1]}{\Gamma(\alpha \sigma + 1)}
\]

to estimate the parameters.
The resulting moment equations for \( \mu, \sigma \) and \( \alpha \) are expressed as

\[
\hat{\mu} = \log \left( \frac{\bar{Y}_2}{\Gamma(\hat{\sigma} + 1)} \right); \quad \frac{\bar{Y}_2^2}{\bar{Y}_4} = \frac{\Gamma(\hat{\sigma} + 1)^2}{3 \Gamma(2\hat{\sigma} + 1)}; \\
\frac{\bar{Y}_{22}}{e^{2\mu} \Gamma(\sigma + 1)} = \frac{\Gamma[(\alpha + 1)\sigma + 1]}{\Gamma(\alpha \sigma + 1)},
\]

(15)

where \( \bar{Y}_2 = (1/T) \sum_{t=1}^{T} y_t^2 \), \( \bar{Y}_{22} = (1/T) \sum_{t=1}^{T} y_t^2 y_{t-1}^2 \) and \( \bar{Y}_4 = (1/T) \sum_{t=1}^{T} y_t^4 \).

As noted earlier, now using the results of Hansen (1982), we can show that the moment estimators are consistent and asymptotically normal (CAN).
Now we have the following result, proved by Hansen(1982)

**Result:** Suppose that the sequence \( \{ r_t : -\infty < t < \infty \} \) satisfies the assumptions stated by Hansen(1982). Then

\[
\left\{ \sqrt{T} \left( \hat{\theta} - \theta \right), \ T \geq 1 \right\}
\]

converges in distribution to a normal random vector with mean 0 and dispersion matrix

\[
\left[ DS^{-1} D' \right]^{-1}
\]

where

\[
D = E \left[ \frac{\partial}{\partial \theta} f (r_t, \theta_0) \right] \quad \text{and} \quad S = \sum_{k=-\infty}^{\infty} \Gamma(k), \ \Gamma(k) = E \left( f_t f'_{t-k} \right).
\]
Since the sequence \( \{h_t\} \) defined by (9) is stationary and ergodic, it follows that the sequence \( \{y_t\} \) given in (8) also possesses these properties.

Further all the moments of \( y_t \) and \( h_t \) are finite.

Hence the asymptotic dispersion matrix becomes \( \frac{1}{T} \Sigma \), where

\[
\Sigma = \left[D \ S^{-1} \ D' \right]^{-1}.
\] (16)
Simulation Study

- We carry out a simulation study to understand the performance of the estimators with sample sizes 1000 and 2000.
- First, we generate a sample of size $T$ from the GEV Markov sequence specified in (9) using the innovation random variable described in (11).
- Then simulate the sequence $\{y_t\}$ using GEV AR(1) model. We use this simulated sample to obtain the estimators of the parameters.
For each specified value of the parameter, we repeat the experiment 1000 times for computing the estimates and then averaged them over the repetitions.

From the Table, we observe that when the sample size is large, the estimates perform reasonably well and there is a significant reduction in asymptotic standard deviations and root mean square errors.

Hence we claim that the method of moment estimation yields good estimates for the parameters involved.
<table>
<thead>
<tr>
<th>True Values</th>
<th>( \hat{\mu} )</th>
<th>( \hat{\sigma} )</th>
<th>( \hat{\alpha} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu )</td>
<td>( \sigma )</td>
<td>( \alpha )</td>
<td>Mean</td>
</tr>
<tr>
<td>0.5</td>
<td>1.5</td>
<td>0.9</td>
<td>0.5334</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.8</td>
<td>0.5342</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.7</td>
<td>0.5287</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.5</td>
<td>0.5137</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.2</td>
<td>0.5237</td>
</tr>
<tr>
<td>0.7</td>
<td>1</td>
<td>0.9</td>
<td>0.7016</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.8</td>
<td>0.7022</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.7</td>
<td>0.7027</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.5</td>
<td>0.6950</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.2</td>
<td>0.6984</td>
</tr>
</tbody>
</table>
Data Analysis

To illustrate the application of the proposed model, we analysed the following three sets of data for the period: 01.01.2010 - 13. 06. 2017.

1. the closing index data of Bombay Stock Exchange (BSE),
2. daily high and
3. daily low index data of Standard and Poor’s 500 (S&P 500).

The time series plots of these data are given in the following figure. The left panels show the plots of actual data series and the log-return series are on the right panels.
Figure: Time series plot of the stock prices and the return
Denoting the daily price index by $p_t$, the returns are transformed into continuously compounded rates centered around their sample mean:

$$y_t = \ln \left( \frac{p_t}{p_{t-1}} \right) - \left( \frac{1}{T} \right) \sum_{t=1}^{T} \ln \left( \frac{p_t}{p_{t-1}} \right).$$

The summary statistics of the return series are reported in table, where $Q(20)$ and $Q^2(20)$ are the Ljung-Box statistic for return and squared return series with lag 20.

The test suggests that the return series is serially uncorrelated whereas the squared return series has significant serial correlation.

The kurtosis of the returns for all the series is greater than three which implies that the distribution of the returns is leptokurtic in nature.
Table: Summary Statistics of Log Return Series

<table>
<thead>
<tr>
<th>Statistics</th>
<th>BSE Closing Index</th>
<th>S&amp;P High Index</th>
<th>S&amp;P Low Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Size</td>
<td>1827</td>
<td>1872</td>
<td>1872</td>
</tr>
<tr>
<td>Mean</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.010435</td>
<td>0.7204921</td>
<td>0.8855629</td>
</tr>
<tr>
<td>Minimum</td>
<td>-6.938474</td>
<td>-3.563278</td>
<td>-8.35233</td>
</tr>
<tr>
<td>First Quartile</td>
<td>-0.644441</td>
<td>-0.317065</td>
<td>-0.40053</td>
</tr>
<tr>
<td>Median</td>
<td>0.011182</td>
<td>-0.001545</td>
<td>0.06957</td>
</tr>
<tr>
<td>Third Quartile</td>
<td>0.680049</td>
<td>0.361892</td>
<td>0.45031</td>
</tr>
<tr>
<td>Maximum</td>
<td>4.061033</td>
<td>3.612929</td>
<td>4.27316</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>4.525028</td>
<td>6.17833</td>
<td>10.60638</td>
</tr>
</tbody>
</table>
Figure: ACF of the returns (top panels) and the squared returns (bottom panels)
Table: Parameter Estimates using Method of Moments

<table>
<thead>
<tr>
<th>Parameters</th>
<th>BSE Closing Index</th>
<th>S&amp;P High Index</th>
<th>S&amp;P Low Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\mu}$</td>
<td>0.115052</td>
<td>-0.6688589</td>
<td>-0.5561562</td>
</tr>
<tr>
<td>$\hat{\sigma}$</td>
<td>0.7034</td>
<td>1.02925</td>
<td>1.537918</td>
</tr>
<tr>
<td>$\hat{\alpha}$</td>
<td>0.2101337</td>
<td>0.7393923</td>
<td>0.6374557</td>
</tr>
</tbody>
</table>
Once the estimates of parameters are obtained, the next stage is the model diagnostic checking.

That is, we need to check whether the assumptions on the model (8) are satisfied with respect to the data we have analysed.

The model (8) is in terms of the volatilities $h_t$, which are unobservable. This aspect makes the diagnosis problem difficult.

One of the methods suggested in such cases is to employ Kalman filtering by rewriting the model (8) in the state-space form.

For more details on Kalman filter method and associated theory, one can refer Jacquier et al (1994) and Tsay (2005).
Since the Kalman Filter method is developed under the normality assumptions, we approximate the distribution of $\eta_t$ specified in (9) by a normal distribution and then adopt Kalman filter method for estimating the volatilities.

Using these estimated volatilities, we can compute the residuals using the equation (8).

The state space representation of the SV model given in (8) can be written as

$$\log y_t^2 = -1.27 + h_t + \nu_t, \quad E(\nu_t) = 0, \quad V(\nu_t) = \frac{\pi^2}{2}$$

and

$$h_t = \alpha h_{t-1} + \eta_t$$

where $\eta_t$ is assumed to be normally distributed with mean $\mu^* = (1 - \alpha)(\mu - \sigma \gamma)$ and variance $\sigma^2* = (1 - \alpha^2)\left(\frac{\pi^2\sigma^2}{6}\right)$ which are given in (12).
If the distribution of $\nu_t$ is approximated by a normal distribution then the preceding system (17) becomes a standard dynamic linear model, to which the Kalman filter can be applied.

Let us define that $\bar{h}_{t|t-1}$ be the prediction of $h_t$ based on the information available at time $t - 1$ and $\Omega_{t|t-1}$ be the variance of the predictor.

Here we are making a presumption that the update that uses the information at time $t$ as $\bar{h}_{t|t}$ and the variance of the update as $\Omega_{t|t}$. 
The equations that recursively compute the predictions and updating are given by

\[
\bar{h}_{t|t-1} = \alpha \bar{h}_{t-1|t-1} + (1 - \alpha) (\mu - \sigma \gamma)
\]

\[
\Omega_{t|t-1} = \alpha^2 \Omega_{t-1|t-1} + (1 - \alpha^2) \frac{\pi^2 \sigma^2}{6}
\]

and

\[
\bar{h}_{t|t} = \bar{h}_{t|t-1} + \frac{\Omega_{t|t-1}}{f_t} \left[ \log r_t^2 + 1.27 - \bar{h}_{t|t-1} \right]
\]

\[
\Omega_{t|t} = \Omega_{t|t-1} \left(1 - \frac{\Omega_{t|t-1}}{f_t} \right)
\]

where \( f_t = \Omega_{t|t-1} + \frac{\pi^2}{2} \).
Then the residuals are calculated by the equation

$$\hat{\varepsilon}_t = y_t \exp \left( -\frac{\bar{h}_t}{2} \right)$$  \hspace{1cm} (18)

and use this sequence for the residual analysis.

The system is initialized at the unconditional values, 
$$\Omega_0 = \frac{\pi^2 \sigma^2}{6} \text{ and } h_0 = \mu - \sigma \gamma.$$  

The residual analysis is carried out using this prediction error.

The parameters $\mu$, $\sigma$ and $\alpha$ in the above system are replaced by their respective estimates.
Further, we also checked the significance of ACF in the residuals by computing the Ljung-Box statistic for the series \( \{\hat{\varepsilon}_t\} \) and \( \{\hat{\varepsilon}_t^2\} \), which are summarized in the table below.

All these values are less than the 5% chi-square critical value (10.117) at degrees of freedom 20.

Hence we conclude that there is no significant serial dependence among the residuals and the squared residuals.
Table: Ljung Box Statistic for the Residuals and Squared Residuals

<table>
<thead>
<tr>
<th>Statistic</th>
<th>BSE Closing Index</th>
<th>S&amp;P High Index</th>
<th>S&amp;P Low Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residuals</td>
<td>0.004958879</td>
<td>0.2233965</td>
<td>1.515389</td>
</tr>
<tr>
<td>Squared Residuals</td>
<td>1.658886</td>
<td>0.1583074</td>
<td>0.9549111</td>
</tr>
</tbody>
</table>
Figure: Histogram of residuals with superimposed standard normal density.
Remark:

- The Weibull - GEV model is more suitable for volatility related to extremes, such as daily maxima, minima, etc.
- Another distribution on the positive support useful for modeling volatility is Birnbaum-Sauners.
- We briefly discuss the model and illustrate with an example.
Birnbaum-Saunders SV Model

Let $y_t$ be the return at time $t$. Define the SV model

$$y_t = \sqrt{h_t} \varepsilon_t,$$

$$h_t = \beta \left[ \frac{1}{2} \alpha X_t + \sqrt{\left( \frac{1}{2} \alpha X_t \right)^2 + 1} \right]^2,$$

$$X_t = \rho X_{t-1} + \eta_t \quad ; \quad |\rho| < 1, \quad t = 1, 2, \ldots,$$

with $\{X_t\}$ be a stationary Gaussian AR(1) sequence with standard normal marginal distribution. \(\{\varepsilon_t\}\) is a sequence of independent and identically distributed standard normal random variables. We assume that the sequence $\{\varepsilon_t\}$ is independent of $h_t$ and $\eta_t$ for every $t$. 

\[\text{(19)}\]


This is a stochastic volatility model for the return series \( \{r_t\} \) whose volatilities are generated by a stationary Markov sequence of BS random variables with marginal probability density function

\[
f (h_t; \alpha, \beta) = \frac{1}{2\alpha\beta\sqrt{2\pi}} \left[ \left( \frac{\beta}{h_t} \right)^{1/2} + \left( \frac{\beta}{h_t} \right)^{3/2} \right] \exp \left( -\frac{1}{2\alpha^2} \left[ \frac{h_t}{\beta} + \frac{\beta}{h_t} - 2 \right] \right),
\]

where \( h_t > 0, \alpha, \beta > 0 \).
Figure: The plot of kurtosis of return and the ACF of squared return
Estimation: Method of moments

Let \((y_1, y_2, \ldots, y_T)\) be a realization of length \(T\) from the SV model (19), \(\Theta = (\alpha, \beta, \rho)\) be the parameter vector to be estimated. We use the moments

\[
E(y_t^2) = \beta \left(1 + \frac{\alpha^2}{2}\right), \quad E(y_t^4) = 3\beta^2 \left(1 + 2\alpha^2 + \frac{3}{2}\alpha^4\right),
\]

\[
E(y_t^2 y_{t-1}^2) = \beta^2 \left(1 + \alpha^2 + \frac{\alpha^4}{4}(1 + 2\rho^2) + \alpha^2 I_1\right)
\]

to estimate the parameters.
The resulting moment equations for $\alpha$, $\beta$ and $\rho$ are expressed as

$$\frac{\bar{Y}_2^2}{\bar{Y}_4} = \frac{(1 + \frac{\hat{\alpha}^2}{2})^2}{3 \left(1 + 2\hat{\alpha}^2 + \frac{3}{2}\hat{\alpha}^4\right)}; \quad \hat{\beta} = \frac{\bar{Y}_2}{\left(1 + \frac{\hat{\alpha}^2}{2}\right)}$$

and

$$\bar{Y}_{22} = \hat{\beta} \left(1 + \hat{\alpha}^2 + \frac{\hat{\alpha}^4}{4} (1 + 2\hat{\rho}^2) + \hat{\alpha}^2 \hat{I}_1\right),$$

where

$$\bar{Y}_2 = \frac{1}{T} \sum_{t=1}^{T} y_t^2, \quad \bar{Y}_{22} = \frac{1}{T} \sum_{t=1}^{T} y_t^2 y_{t-1}, \quad \bar{Y}_4 = \frac{1}{T} \sum_{t=1}^{T} y_t^4.$$

To be solved numerically.
The method of moments estimates are easy to compute, but not very efficient.

We tried a likelihood based computation method known as Efficient Important Sampling (EIS) proposed by Richard and Zang (2007).

The procedure is based on a suitable decomposition of the likelihood function and then following an algorithm based on MCMC method.

The diagnosis method is same as before using state-space representation.

We applied this method for data on USD/INR exchange rate and S&P500 opening index.
Data Analysis

We apply the BS-SV model to analyse the daily returns for
(1) the rate of exchange on the Rupee/Dollar from July 25, 1998 to
May 22, 2015 obtained from Data base on Indian Economy, Reserve
Bank of India and
(2) the opening index of Standard and Poors 500 (S&P 500) from
January 02, 2008 to May 22, 2015 obtained from Yahoo Finance.
The plots of time series and the corresponding centered return series
are given in the next Figure.
**Figure:** The plot of return and the ACF of squared return
<table>
<thead>
<tr>
<th>Statistic</th>
<th>Dollar Exchange rate</th>
<th>Ex-</th>
<th>S&amp;P500 Opening Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
<td>4051</td>
<td>1861</td>
<td></td>
</tr>
<tr>
<td>Minimum</td>
<td>-3.0164</td>
<td>-9.1349</td>
<td></td>
</tr>
<tr>
<td>Maximum</td>
<td>4.0100</td>
<td>10.1193</td>
<td></td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.4225</td>
<td>1.3496</td>
<td></td>
</tr>
<tr>
<td>Kurtosis</td>
<td>11.3384</td>
<td>12.7281</td>
<td></td>
</tr>
<tr>
<td>$Q(20)$</td>
<td>1.7088</td>
<td>1.4721</td>
<td></td>
</tr>
<tr>
<td>$Q^2(20)$</td>
<td>87.1655</td>
<td>68.1816</td>
<td></td>
</tr>
</tbody>
</table>

**Table:** Descriptive statistics of the return series
Figure: ACF of the returns and the squared returns
Computed the residuals using State space representation.

Figure: ACF of the residuals
### Table: Estimates of parameters and Ljung-Box statistic for residuals

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Dollar rate</th>
<th>Exchange rate</th>
<th>S&amp;P 500 Opening Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>2.9411</td>
<td>2.5030</td>
<td></td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.1721</td>
<td>0.4404</td>
<td></td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.9101</td>
<td>0.7942</td>
<td></td>
</tr>
<tr>
<td>$Q^*(20)$</td>
<td>0.6820</td>
<td>0.3428</td>
<td></td>
</tr>
<tr>
<td>$Q^{2*}(20)$</td>
<td>5.0400</td>
<td>1.1539</td>
<td></td>
</tr>
</tbody>
</table>
Figure: Histogram of residuals with superimposed standard normal density
Some References

Some References


Thank You