Journal of Statistical and Econometric Methods, vol.2, no.2, 2013, 39-50
ISSN: 1792-6602 (print), 1792-6939 (online)
Scienpress Ltd, 2013

A Copula-GARCH Model of Conditional Dependencies: Estimating Tehran Market Stock Exchange Value-at-Risk

Sedigheh Shams[1] **and Fatemeh K. Haghighi**[2]

Abstract

Modeling the dependency between stock market returns is a difficult task when returns follow a complicated dynamics. It is not easy to specify the multivariate distribution relating two or more return series. In this paper, a methodology based on fitting ARIMA, GARCH and ARMA-GARCH models and copula functions is applied. In such methodology, the dependency parameter can easily be rendered conditional and time varying. This method is used to the daily returns of five major stock markets (Telecom (TE), Sina darou (SI), Motojen (MO), Mellat bank (ME), and Esfahan oil refinery (ES)). Then Value-at-Risk of Tehran Stock Exchange portfolio including mentioned assets, is estimated.

Mathematics Subject Classification: 62H05, 62M15
Keywords: ARMA-GARCH model, Copula, Dependency structure, GARCH model, Portfolio, Value-at-Risk

[1] Department of Mathematics, Alzahra University.
[2] Department of Mathematics, Alzahra University.

Article Info: *Received* : January 16, 2013. *Revised* : March 22, 2013
Published online : June 1, 2013

1 Introduction

Value-at-Risk (VaR) has been one of the most popular tools for estimating market risk. VaR can provide information about the loss of a portfolio with a given confidence level. Modeling dependence through estimating time-varying conditional correlations between variables, is the key of portfolio construction and VaR (Alexander J. McNeil, Rudiger Frey and Paul Embrechts. (2005)). Traditionally, normal distribution was frequently assumed and correlation was used to describe the dependence between random variables. But in many cases normal distribution could not perform well. So researchers have suggested copula functions as an effective tool for modeling the correlation between variables. Regarding to data structure, time-varying data, usually GARCH model (Franco Ch. and Zakoian J.(2010) and Cherubin U.,Luciano E. and Vecchiato W. (2004)) is used to describe the dependency and together with copula, copula-GARCH model is used to model of conditional dependencies (Eric Jondeau and Micheal Rockinger (2006)). In section 2, we outline the methodological approach. Section 3 presents a step-by-step programme for estimating VaR of a portfolio of Tehran stock market price index.

2 Methodology

In this section some necessary concepts for estimating Tehran Market Stock Exchange Value-at-Risk, are briefly presented.

2.1 Copula

Consider a random vector $(X_1, X_2, \ldots, X_d)$. Suppose its margins,$F_1, F_2, \ldots, F_d$, are continuous functions. By applying the probability integral transform to each component, the random vector

$$(U_1, U_2, \ldots, U_d) = (F_1(X_1), F_2(X_2), \ldots, F_d(X_d))$$

has uniform margins.

The copula of $(X_1, X_2, \ldots, X_d)$ is defined as the joint cumulative distribution function of $(U_1, U_2, \ldots, U_d)$:

$$C(u_1, u_2, \ldots, u_d) = P(U_1 \leq u_1, U_2 \leq u_2, \ldots, U_d \leq u_d)$$

The copula C contains all information on the dependence structure between the components of $(X_1, X_2, \ldots, X_d)$ whereas the marginal cumulative distribution functions F_i contain all information on the marginal distributions. The above formula for the copula function can be rewritten to correspond to this as:

$$C(u_1, u_2, \ldots, u_d) = P(X_1 \leq F_1^{-1}(u_1), X_2 \leq F_2^{-1}(u_2), \ldots, X_d \leq F_d^{-1}(u_d))$$

In probabilistic terms, $C : [0,1]^d \to [0,1]$ is a d-dimensional copula if C is a joint cumulative distribution function of a d-dimensional random vector on the unit cube $[0,1]^d$ with uniform marginals.

In analytic terms (Nelsen(1999)), $C : [0,1]^d \to [0,1]$ is a d-dimensional copula if

(i) $C(u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_d) = 0$,

(ii)$C(1, \ldots, 1, u, 1, \ldots, 1) = u$,

(iii) C is d-increasing.

Sklar's theorem provides the theoretical foundation for the application of copulas. Sklar's theorem states that a multivariate cumulative distribution function $H(x_1, \ldots, x_d)$ of a random vector $(X_1, \ldots, X_d)$ with marginals F_i can be written as

$$H(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)) \tag{1}$$

where C is a copula.

The converse is also true: given a copula $C : [0,1]^d \to [0,1]$ and margins F_i then $C(F_1(x_1), \ldots, F_d(x_d))$ defines a d-dimensional cumulative distribution function.

For example the Gaussian copula is a distribution over the unit cube $[0,1]^d$. It is constructed from a multivariate normal distribution over $\mathbf{R}^d$ by using the probability integral transform. For a given correlation matrix $\Sigma \in \mathbf{R}^{d \times d}$, the Gaussian copula with parameter matrix Σ can be written as

$$C_\Sigma(u) = \Phi_\Sigma(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d)) \tag{2}$$

where Φ^{-1} is the inverse cumulative distribution function of a standard normal and Φ_Σ is the joint cumulative distribution function of a multivariate normal distribution with mean vector zero and covariance matrix equal to the correlation matrix Σ .

Another famous copula family is Archimedean copula family. In practice, Archimedean copulas are popular because they allow modeling dependence in arbitrarily high dimensions with only one parameter, governing the strength of dependence.

A copula C is called Archimedean if it admits the representation

$$C(u) = \Psi_\Sigma(\Psi^{-1}(u_1) + \ldots + \Psi^{-1}(u_d)) \tag{3}$$

where Ψ is the so called generator.

The above formula yields a copula if and only if Ψ is d-monotone on $[0,\infty)$, that is, if the kth derivatives of Ψ satisfy $(-1)^k\Psi^{(k)}(x) \geq 0$ for all $x \geq 0$ and $k = 0, 1, \ldots, d-2$ and $(-1)^{d-2}\Psi^{d-2}(x)$ is non-increasing and convex.

The generators in Table 1 are the most popular ones. All of them are completely monotone.

Table 1: The most popular Archimedean copulas

name	$\psi(t)$	$\psi^{-1}(t)$	parameter
Ali-Mikhail-Haq	$\frac{1-\theta}{e^t-\theta}$	$\log\left(\frac{1-\theta+\theta t}{t}\right)$	$\theta \in [0,1)$
Clayton	$(1+\theta t)^{-1/\theta}$	$\frac{1}{\theta}(t^{-\theta}-1)$	$\theta \in (0,\infty)$
Frank	$-\frac{log(1-(1-e^{-\theta}))e^{-t}}{\theta}$	$-log(\frac{e^{-\theta t}-1}{e^{-\theta}-1})$	$\theta \in (0,\infty)$
Gumbel	$exp(-t^{1/\theta})$	$(-log(t))^\theta$	$\theta \in [1,\infty)$
Joe	$1-(1-e^{-t}))^{1/\theta}$	$-log(1-(1-t)^\theta)$	$\theta \in [1,\infty)$

2.2 Value-at-Risk

Given a confidence level $\alpha \in (0,1)$, the VaR of the portfolio at the confidence level α is given by the smallest number such that the probability that the loss $\mathbf{L}$ exceeds $\mathbf{l}$ is at most $(1-\alpha)$ (Alexander J. McNeil, Rudiger Frey and Paul Embrechts. (2005)). Mathematically, if $\mathbf{L}$ is the loss of a portfolio, then $VaR_\alpha(\mathbf{L})$ is the level α-quantile, i.e.

$$VaR_\alpha(\mathbf{L}) = \inf\{\mathbf{l} \in \mathbf{R}; P(\mathbf{L} > \mathbf{l}) \leq 1-\alpha\} = \inf\{\mathbf{l} \in \mathbf{R}; F_L(\mathbf{l}) \geq \alpha\} \tag{4}$$

Risk managers typically assume that some fraction of the bad events will have undefined losses, either because markets are closed or illiquid, or because the entity bearing the loss breaks apart or loses the ability to compute accounts. Therefore, they do not accept results based on the assumption of a well-defined probability distribution. On the other hand, many academics prefer to assume a well-defined distribution, usually one with fat tails. This point has probably caused more contention among VaR theorists than any other.

In this paper the VaR of a portfolio return is considered. This portfolio includes some assets, $X_i, (i = 1, ..., d)$ and the portfolio return, Rp_t, is equal to

$$Rp_t = \sum_{i=1}^{d} \gamma_i X_{i,t} \tag{5}$$

where $X_{i,t}$ is the value of ith asset at time t, and the coefficient $\gamma_i (i = 1, ..., d)$ should be such that $\sum_{i=1}^{d} \gamma_i = 1$.

2.3 Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) models

AutoRegressive Conditional Heteroskedasticity (ARCH) models are used to characterize and model observed time series. They are used whenever there is reason to believe that, at any point in a series, the terms will have a characteristic size, or variance. In particular ARCH models assume the variance of the current error term or innovation to be a function of the actual sizes of the previous time periods error terms: often the variance is related to the squares of the previous innovations.

Suppose one wishes to model a time series using an ARCH process. Let ϵ_t denote the error terms (return residuals, with respect to a mean process) i.e. the series terms. These ϵ_t are split into a stochastic piece and a time-dependent standard deviation σ_t characterizing the typical size of the terms so that $\epsilon_t = \sigma_t z_t$. The random variable z_t is a strong White noise process. The series σ_t^2 is modeled by

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + ... + \alpha_q \epsilon_{t-q}^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 \tag{6}$$

Where $\alpha_0 > 0$ and $\alpha_i \geq 0, i > 0$ (Franco Ch. and Zakoian J.(2010)).

An ARCH (q) model can be estimated using ordinary least squares. A methodology to test for the lag length of ARCH errors using the Lagrange multiplier test was proposed by Engle (1991). This procedure is as follows:

1. Estimate the best fitting autoregressive model AR(q)

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + ... + \alpha_q y_{t-q} + \epsilon_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i y_{t-i} + \epsilon_t$$

2. Obtain the squares of the error $\hat{\epsilon}^2$ and regress them on a constant and q lagged values

$$\hat{\epsilon}_t^2 = \hat{\alpha}_0 + \sum_{i=1}^{q} \hat{\alpha}_i \hat{\epsilon}_{t-i}^2$$

where q is the length of ARCH lags.

3.The null hypothesis is that, in the absence of ARCH components, we have $\alpha_i = 0$ for all $i = 1, ..., q$. The alternative hypothesis is that, in the presence of ARCH components, at least one of the estimated α_i coefficients must be significant.

If an AutoRegressive Moving Average model (ARMA model) is assumed for the error variance, the model is a Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) model.

In that case, the GARCH(p,q) model (where p is the order of the GARCH terms σ^2 and q is the order of the ARCH terms ϵ^2 is given by

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + ... + \alpha_q \epsilon_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + ... + \beta_p \sigma_{t-p}^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \tag{7}$$

Generally, when testing for heteroskedasticity in econometric models, the best test is the White test. However, when dealing with time series data, this means to test for ARCH errors in (6) and GARCH errors in (7).

The lag length p of a GARCH (p,q) process is established in three steps:

1. Estimate the best fitting AR(q) model of equation 3.

2. Compute and plot the sample autocorrelations of ϵ^2 by

$$\rho(i) = \frac{\Sigma_{t=i+1}^{T} (\hat{\epsilon}_t^2 - \hat{\sigma}_t^2)(\hat{\epsilon}_{t-1}^2 - \hat{\sigma}_{t-1}^2)}{\Sigma_{t=1}^{T} (\hat{\epsilon}_t^2 - \hat{\sigma}_t^2)^2}$$

3. For large samples, standard deviation of $\rho(i)$ is $1/\sqrt{T}$. Individual values that are larger than this indicate GARCH errors. To estimate the total number of lags, use the Ljung-Box test until the value of these are less than, say, 10% significant. The Ljung-Box Q-statistic follows χ^2distribution with n degrees of freedom if the squared residuals $\varepsilon_t^2$are uncorrelated. It is recommended to consider up to T/4 values of n. The null hypothesis states that there are no ARCH or GARCH errors. Rejecting the null thus means that there are existing such errors in the conditional variance.

Assuming the density f of the strong white noise z_t is known, and given the observations $\epsilon_1, \ldots, \epsilon_T$ the conditional likelihood function is defined as:

$$L_{T,f}(\theta|\{\epsilon_t\}) = \Pi_{t=1}^{T} \frac{1}{\hat{\sigma_t^2}} f(\frac{\epsilon_t}{\hat{\sigma_t^2}}) \tag{8}$$

Where θ is the vector of the parameters that is estimated for conditional mean and variance and $\hat{\sigma_t^2}$ is recursively defined by (7). A maximum likelihood estimator (MLE) is obtained by maximizing the likelihood function defined by (8).

3 Data analysis

The applications of copulas in quantitative finance are numerous, both in the real-world probability of risk/portfolio management and in the risk-neutral probability of derivatives pricing. Here, we estimate the joint density of the five main assets in the Tehran stock exchange, based on the data from 22 June 2010 to 23 March 2011, using the methodology described above, and we use this estimated joint density to compute the VaR. Each univariate process is modeled with GARCH(p,q) model selected using the AIC criterion and an appropriate marginal distribution is chosen for the residuals. After estimating the parameters, an appropriate copula , is fitted for all series residuals.

The data set is the value of the following assets during 22 June 2010 to 23 March 2011:

Telecom (TE), Sina daroo (SI), Motojen (MO), Mellat bank (ME) and Esfahan oil refinery (ES).

Table 2 shows the summary statistics on daily returns about the mentioned assets.

Table 2: Summary statistics on daily returns

	TE	SI	MO	ME	ES
Mean	0.0031	0.0026	0.0046	0.0013	-0.0008
	(0.0015)	(0.0042)	(0.0074)	(0.0021)	(0.0043)
Std	0.0194	0.0538	0.0944	0.0273	0.0546
Skewness	-0.1845	-0.1368	-0.5235	-1.1429	0.5720
Kurtosis	0.0975	3.0769	8.7104	9.6452	6.3256
Correlation matrix					
TE	1	0.1418	0.0951	-0.0381	0.0228
SI	0.1418	1	0.0260	0.2286	0.1689
MO	0.0951	0.0260	1	-0.08148	-0.1859
ME	-0.0381	0.2286	-0.08148	1	0.2909
ES	0.0228	0.1689	-0.1859	0.2909	1

3.1 Fitting ARMA-GARCH

Each of time series has been considered separately. Table 3 shows the values of test statistics for ARIMA, GARCH and ARMA-GARCH models residuals and Table 4 shows the best model characterized for each of them.

Table 3: Values of test statistics for ARMA/GARCH models residuals

	TE	SI	MO	ME	ES
LM Arch Test(TR^{2}2)	2.4696	[a]	0.4608	2.8766	5.8340
p-Values	(0.9983)	[a]	(0.9999)	(0.9963)	(0.9242)
QW(10) for R	4.8990	[a]	8.2820	4.6730	10.5474
p-Values	(0.8978)	[a]	(0.6003)	(0.9119)	(0.3938)
QW(10) for R^2	1.7618	[a]	0.3510	2.9921	6.7104
p-Values	(0.9979)	[a]	(0.9999)	(0.9816)	(0.7525)

a.ARMA(0,4)

Table 4: The best model for each of the assets

asset	model
TE	arma(0,1)+garch(1,1)
SI	arma(0,4)
MO	garch(1,1)
ME	garch(1,1)
ES	arma(0,1)+garch(1,0)

3.2 Selecting the Marginal Distributions

After specifying model for each asset (TE, SI, MO, ME and ES) it is necessary to fit an appropriate marginal distribution to the residuals. Some known distributions (normal, t-Student, skewed t-Student, skewed generalized error, generalized error, Cauchy, skewed normal and ...) has been fitted to them, and by Anderson-Darling statistic the model adequacy has been considered. The best marginal distribution has been characterized by AIC criteria showing in Table 5.

Table 5: The marginal distribution of asset residuals

Asset	Marginal distribution (estimated parameter)	Anderson-Darling test statistic
TE	T-Student (mean=0.053, sd=1.07,d.f.=3)	1.437077
SI	Normal (mean= 6.181236e-06,sd= 0.045)	1.945678
MO	T-Student (mean= 0.044, sd= 0.86, d.f.=3)	1.037849
ME	Skewed generalized error (mean= -8.173e-14, sd= 1.22e-01, d.f.=1, $\gamma = 1$)	2.511184
ES	Skewed T-Student (mean=0.167, sd= 1.13, d.f=3, $\gamma = 1.55$)	1.656944

3.3 Modeling dependency (Copula selecting)

The copula families used in modeling the data are: Frank, Joe, Clayton, AMH, Gumble, T-student and Normal, for details on these classes see Nelsen (1999). The choice of copula models is partly based on previous analysis, tractability. In this step, to all assets residuals, with selected marginal distributions, seven mentioned copula models were fitted (Fermanian,J.D. (2005)), the suitable ones are determined in Table 6 and by Cramer-Von Mises test criteria, there were two suitable copulas, T-student and Normal. Between these two copulas, by AIC criteria (AIC(T-student)= -1.881099 that is less than AIC(Normal)= 5.877649), it was found that Normal copula is more suitable for five assets residuals as their joint distribution.

Table 6: The Cramer-Von Mises test statistic values

copula	Cramer-Von Mises test statistic	sig.
Frank	0.3456018	$sig. < 0.01$
Joe	0.3530452	$sig. < 0.01$
Clayton	0.3570421	$sig. < 0.01$
AMH(Ali-Mikhail-Hag)	0.367845	$sig. < 0.01$
Gumble	0.3465347	$sig. < 0.01$
T- Student	0.04091816	$sig. > 0.2$
Normal	**0.03493014**	$sig. > 0.2$

3.3.1 Estimating the Copula Parameters

After selecting the Normal copula as the joint distributions for five assets residuals marginals, it is necessary to estimate its parameters. The maximum likelihood method is used. Maximization can be performed by Newton method on the first derivative or by an interval search. The parameter estimates for Normal copula models are reported in Table 7.

Table 7: Parameters estimates for Normal copula

ρ	TE	SI	MO	ME	ES
TE	1	0.1917	0.1008	-0.0027	-0.0342
SI		1	0.1384	0.0591	0.1013
MO			1	-0.0736	0.0277
ME				1	-0.0016
ES					1
Mean	0.7693	0.01768	0	0	-0.0001
Variance	0.0349	1	0.0461	0.0066	0.01791

3.4 Estimating Value at Risk

Now the question is how to compute the Value-at- Risk of a portfolio using copulas. In order to minimize the portfolio variance, lagrange multipliers method is used and estimation of coefficients are obtained in the following table:

Table 8: The results of Coefficient of assets

Asset	Coefficient of assets
TE	4.484788e-10
SI	9.245571e-01
MO	2.101806e-04
ME	7.523272e-02
ES	5.681134e-09

By assuming Normal copula fitted to the data and substituting the table 8 coefficients, VaR is estimated for two levels $\alpha = 0.01$ and $\alpha = 0.05$ as follows.

$$P(R_{portfolio} < VaR_{portfolio} = 0.1021410904) = 0.99$$

and

$$P(R_{portfolio} < VaR_{portfolio} = 0.0592324264) = 0.95$$

where

$$\begin{aligned} R_{portfolio} &= (4.484788e-10)TE + (9.245571e-01)SI + (2.101806e-04)MO + \\ &\quad + (7.523272e-02)ME + (5.681134e-09)ES. \end{aligned}$$

Acknowledgements. This research was partially supported by Alzahra University.

References

[1] Alexander J. McNeil, Rudiger Frey and Paul Embrechts, *Quantitative Risk Management: Concepts, Techniques, and Tools*, Princeton Series in Finance, 2005.

[2] U. Cherubini and E. Luciano, *Copula Method in finance*, Wiley, 2004.

[3] R.F. Engle and V.K. Ng, Measuring and testing the impact of news on volatility, *Journal of Finance*, **48**(5), (1991), 17491778.

[4] J.D. Fermanian, Goodness of Fit Tests for Copulas, *Journal of American Statistical Assiciation*, **95**, (2005), 119-152.

[5] Ch. Franco and J. Zakoian, *GARCH models, structure, statistical inference and financial application*, Wiley, United Kongdom, 2010.

[6] Roger B. Nelsen, *An Introduction to Copulas*, New York, Springer, 1999.