Backtesting of market risk estimation assuming various copula functions

Aleš Kresta[1]

Abstract. Market risk estimation is a challenging and no less important task of all financial institutions, which requires the modeling of portfolio returns. When modeling the portfolio returns, we are interested in modeling both the distributions of individual assets returns and dependency of these marginal distributions. The useful tool for dependency modeling are copula functions. The task of this article is to compare the utilization of various copula functions, specifically Gaussian, Student and some other types of copula functions, for portfolio returns modeling and subsequent VaR estimation. As marginal distributions normal inverse Gaussian model (NIG) and also normal distribution are assumed in the paper. These two marginal distributions both joined by chosen copula functions are backtested on time series of historical returns of portfolios dependent on both stock market indices and foreign exchange rates.

Keywords: backtesting, market risk, model validation, subordinated Lévy model, copula function.

JEL Classification: G15, G21, G22
AMS Classification: 60G51

1 Introduction

Financial risk management is an important part of all financial institutions such as banks and insurance companies. In order to manage the risk well we have to be able to measure the risk soundly. Since the Gaussian model is not suitable for returns modeling, some alternative models for market risk modeling have been tested recently. Rank [13] analyzed various marginal distributions coupled together by copula functions for risk estimation, Alexander and Sheedy [1] assumed Gaussian/Student/GARCH/Empirical models for a simple positions. Also Lévy models are suitable for marginal risk modeling as showed Tichý [15], who assumed a variance gamma model (VG model) and a normal inverse Gaussian model (NIG model) coupled together by elliptical copula functions. As the author showed the VG model and the NIG model provided almost the same results. Hence Kresta and Tichý [9; 10] assumed only the NIG model as it is computationally less costly to evaluate its inverse distribution function which is needed in copula modeling framework.

In this paper we extend the previous researches also on Archimedean copula functions. The goal of the paper is to backtest various elliptical and Archimedean copula functions with marginal distributions in form of the Gaussian normal distribution and the NIG model and to compare the quantity of observed exceptions obtained utilizing this models.

We proceed as follows. First, the normal inverse Gaussian model is characterized. Then the copula functions are defined with the focus on the most important elliptical and Archimedean copula functions. Afterwards the backtesting procedure is introduced and statistical test concerning the results of the backtesting procedure is described. In empirical part of the paper we present the results obtained by utilization of the Gaussian normal distribution and the NIG model coupled together by different copula functions for Value at Risk (VaR) estimation of chosen portfolios.

2 Methodology

In this paper we utilize the copula functions for market risk estimation in terms of VaR. Hence, we assume several distinct risk factors, i.e. a marginal (independent) distributions, for which the NIG model (subsection 2.1) is utilized. These marginal distributions are grouped together by various copula functions (subsection 2.2) and market risk models are formed. These models are backtested on historical data of chosen portfolios of FX rates and equity indices. The backtesting procedure is described in subsection 2.3.

---

[1] VŠB – Technical University of Ostrava/Faculty of Economics, Department of Finance, Sokolská tř. 33, Ostrava, ales.kresta@seznam.cz.

2.1 Normal inverse Gaussian model

The normal inverse Gaussian process belongs to family of Lévy models. The most recent and complete monographs on the theory and application of Lévy models are e.g. [2; 5].

Generally, a Lévy process is a stochastic process, which is zero at origin, its path in time is right-continuous with left limits and its main property is that it is of independent and stationary increments. Another common feature is a so called stochastic continuity. Moreover, the related probability distribution of increments must be infinitely divisible. Concerning the probability distribution the crucial theorem is the Lévy-Khintchine formula:

$$\Phi(u) = i\gamma u - \frac{1}{2}\sigma^2 u^2 + \int_{-\infty}^{\infty} \left(\exp(iux) - 1 - iux I_{|x|<1}\right) \nu(dx). \tag{1}$$

For a given infinitely divisible distribution, we can define the triplet of Lévy characteristics,

$$\left\{\gamma, \sigma^2, \nu(dx)\right\}.$$

The former two define the drift of the process (deterministic part) and its diffusion. The latter is a Lévy measure. If it can be formulated as $\nu(dx) = u(x)dx$, it is a Lévy density.

Let X be a Brownian motion. If we substitute a standard time t in Brownian motion X,

$$X(t;\mu;\sigma) = \mu t + \sigma Z(t), \tag{2}$$

by a suitable function $l(t)$ as follows,

$$X(l(t);\theta;\vartheta) = \theta l(t) + \vartheta Z(l(t)) = \theta l(t) + \vartheta \varepsilon \sqrt{l(t)}, \tag{3}$$

we get a subordinated Lévy model. Due to the simplicity (tempered stable subordinators with known density functions in the closed form), the most suitable candidates for function $l(t)$ seem to be either (i) the VG model – the overall process is driven by a gamma process from the gamma distribution or (ii) the NIG model – the subordinator is given by an inverse Gaussian process based on the inverse Gaussian distribution.

In financial literature the NIG model was introduced by Barndorff-Nielsen [3]. Assuming parameters $\alpha > 0$, $-\alpha < \beta < \alpha$ and $\delta > 0$, the NIG model is defined by its characteristic function,

$$\phi_{NIG}(u, \alpha, \beta, \delta) = \exp\left(\delta\sqrt{\alpha^2 - \beta^2} - \delta\sqrt{\alpha^2 - (\beta + iu)^2}\right), \tag{4}$$

or it can be viewed as a subordinated Lévy model by assuming inverse Gaussian process $I(t;\nu)$ as $l(t)$ in (3),

$$X(I(t;\nu);\theta,\vartheta) = \theta I(t;\nu) + \vartheta \varepsilon \sqrt{I(t;\nu)}. \tag{5}$$

The parameters θ, ϑ and ν can be calculated from parameters α, β and δ (and vice versa) as follows,

$$\theta = \frac{\delta\beta}{\sqrt{\alpha^2 - \beta^2}}, \quad \vartheta = \frac{\sqrt{\delta\sqrt{\alpha^2 - \beta^2}}}{\sqrt{\alpha^2 - \beta^2}}, \quad \nu = \frac{1}{\delta\sqrt{\alpha^2 - \beta^2}}. \tag{6}$$

2.2 Copula functions

A useful tool for dependency modeling are the copula functions, i.e. the projection of the dependency among particular distribution functions into $[0,1]$,

$$C : [0,1]^n \rightarrow [0,1] \text{ on } R^n, n \in \{2,3,...\}. \tag{7}$$

Basic reference for the theory of copula functions can be found in [12], while [8; 13] target mainly on the application issues in finance. Alternatively, Lévy processes can be coupled on the basis of Lévy measures by Lévy copula functions. However, this approach is not necessary in our case.

Actually, any copula function can be regarded as a multidimensional distribution function with marginals in the form of standardized uniform distribution.

For simplicity, assume two potentially dependent random variables with marginal distribution functions F_X, F_Y and joint distribution function $F_{X,Y}$. Then, following the Sklar's theorem [14]:

$$F_{X,Y}(x,y) = C(F_X(x), F_Y(y)).\tag{8}$$

If both F_X and F_Y are continuous, a copula function C is unique. Sklar's theorem implies also an inverse relation,

$$C(u,v) = F_{X,Y}\left(F_X^{-1}(u), F_Y^{-1}(v)\right).\tag{9}$$

The formulation above should be understood such that the joint distribution function gives us two distinct information: (i) marginal distributions of random variables, (ii) dependency function of the distributions. Hence, while the former is given by $F_X(x)$ and $F_Y(y)$, copula function specifies the dependency. Only when we put both information together, we get sufficient knowledge about the pair of random variables X, Y.

Elliptical and Archimedean copula functions

With some simplicity we can distinguish the elliptical copula functions and Archimedean copula functions. The elliptical copula functions are based on some elliptical joint distribution, such as Gaussian copula function,

$$C_R^{Ga}(u,v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-R^2}} e^{\left[\frac{2Rst - s^2 - t^2}{2 - 2R^2}\right]} ds\, dt\tag{10}$$

where R is the correlation coefficient, or Student copula function based on the Student t distribution,

$$C_{R,\upsilon}^{St}(u,v) = \int_{-\infty}^{t_\upsilon^{-1}(u)} \int_{-\infty}^{t_\upsilon^{-1}(v)} \frac{1}{2\pi\sqrt{1-R^2}} \left[1 + \frac{s^2 + t^2 - 2Rst}{\upsilon(1-R^2)}\right]^{-\frac{\upsilon+2}{2}} ds\, dt\,,\tag{11}$$

where again R is the correlation coefficient and υ stands for degrees of freedom of the Student t distribution.

On the other hand, Archimedean copula functions are defined on the basis of function ϕ called generator. Generator is continuous, decreasing and convex function such that $\phi(1) = 0$ and for a strict generator also stands that $\phi(0) = +\infty$. Archimedean copula functions can be defined then as follows,

$$C_{\phi,\phi^{[-1]}}^{Arch}(u,v) = \phi^{[-1]}(\phi(u), \phi(v)),\tag{12}$$

where $\phi^{[-1]}$ is the pseudo-inverse function such that $\phi^{[-1]}(\phi(v)) = v$ for every $v \in \langle 0;1 \rangle$. The most known Archimedean copula functions are: Gumbel copula function [7],

$$C_a^{Gl}(u,v) = \exp\left\{-\left[(-\ln u)^a + (-\ln v)^a\right]^{\frac{1}{a}}\right\},\tag{13}$$

Clayton copula function [4],

$$C_a^{Cl}(u,v) = \max\left[\left(u^{-a} + v^{-a} - 1\right)^{-\frac{1}{a}}, 0\right],\tag{14}$$

and Frank copula function [6],

$$C_a^{Fr}(u,v) = -\frac{1}{a}\ln\left[1 + \frac{(e^{-au}-1)(e^{-av}-1)}{e^{-a}-1}\right].\tag{15}$$

Parameters estimation

There exist three main approaches to parameter estimation for copula function based dependency modeling: exact maximum likelihood method (EMLM), inference function for margins (IFM), and canonical maximum likelihood (CML). While for the former all parameters are estimated within one step, which might be very time consuming (mainly for high dimensional problems or complicated marginal distributions), the latter two methods

are based on the estimation of the parameters for the marginal distribution and parameters for the copula function separately. While assuming IFM, marginal distributions are estimated in the first step and the copula function in the second one, for CML instead of parametric margins empirical distributions are used.

2.3 Backtesting

Within the backtesting procedure, the ability of a given model to estimate the future losses is tested. Loosely speaking, backtesting is based on the estimation of the risk (mostly measured as VaR) at time t for time $t + \Delta t$, where Δt is usually (in line with the standards for bank supervision defined within Basel II) set to one business day, and compared with the true loss observed at time $t + \Delta t$. This procedure is applied for moving time window over the whole utilized data set.

Within the backtesting procedure on a given time series two situations can arise – the loss is higher or lower than its estimation (from the stochastic point of view, the equality shouldn't arise). While the former case is denoted by 1 as an exception, the latter one is denoted by zero. In this way, we get the sequence of logical values corresponding to the fact whether the exception occurred or not. On this sequence, it can be tested whether the number of ones (exceptions) corresponds with the assumption, i.e. $\alpha \cdot n$ (where n is the length of the sequence), whether the estimation is valid either unconditionally or conditionally, whether bunching is present, etc.

The quantity of exceptions can be tested by Kupiec's test [11] which is derived from a relative amount of exceptions. Thus it tests whether the number of exceptions is from the statistical point of view different from the assumption. A given likelihood ratio on the basis of χ^2 probability distribution with one degree of freedom is formulated as follows:

$$LR = -2\ln\left[\frac{\pi_{ex}^{n_1}\left(1-\pi_{ex}\right)^{n_0}}{\pi_{obs}^{n_1}\left(1-\pi_{obs}\right)^{n_0}}\right], \tag{16}$$

where π_{ex} is expected probability of exception occurring, π_{obs} is observed probability of exception occurring, $\pi_{obs} = \dfrac{n_1}{n_0+n_1}$, n_0 is the number of zeros and n_1 is the number of ones (the quantity of exceptions).

3 Results

The data set we consider in this study comprises of daily closing prices of four well established equity indices – Down Jones Industrial Average (DJI) from the US market, FTSE 100 (FTSE) from London (UK), Nikkei 225 (N225) from Tokyo (Japan) and Swiss Market Index (SMI) from Switzerland – over preceding 20 years (January 1, 1991 to August 31, 2011). However, the indices are denominated in four distinct currencies, in particular the US dollar (USD), British pound (GBP), Japan yen (JPY) and Swiss franc (CHF). This fact extends our data set to 8 distinct time series. For all currencies we assume the foreign exchange rate to euro (EUR). Since the trading days at particular markets are not always harmonized, we have to interpolate the missing data. In this way we get eight time series of 5,376 log-returns.

Basic descriptive statistics of daily log-returns are apparent from Table 1. In particular for each index and FX rate the minimal and maximum return, its mean (expected value), median, standard deviation and two higher moments, the skewness and kurtosis, are recorded.

Characteristic	DJI	FTSE	N225	SSMI	USD	GBP	JPY	CHF
minimum	-8.20%	-9.26%	-12.11%	-8.38%	-4.06%	-3.89%	-3.90%	-4.58%
maximum	10.51%	9.38%	10.09%	10.79%	4.82%	2.83%	5.93%	3.26%
mean	0.03%	0.02%	-0.02%	0.03%	0.00%	0.00%	0.01%	0.01%
median	0.05%	0.04%	0.00%	0.07%	-0.01%	0.00%	-0.03%	0.00%
standard deviation	1.10%	1.13%	1.47%	1.17%	0.65%	0.48%	0.76%	0.36%
skewness	-0.105	-0.111	-0.323	-0.128	0.142	-0.421	0.415	-0.098
kurtosis	11.740	9.644	8.283	9.323	5.808	7.863	6.937	17.593

Table 1 Basic descriptive statistics of daily log-returns of utilized dataset

From these assets we construct two types of simple portfolios: (i) we assume the equity index and the corresponding FX rate, thus the investment into the index from euro investor point of view, and (ii) the portfolios of USD/EUR currency pair and one from the other three currency pairs. The difference in these two types of portfolios is in the sign of the correlation coefficient. While the indices and corresponding currencies are correlated negatively, the correlations between the FX rates are positive.

For modeling purpose we assume two types of marginal distributions: (i) normal distribution for its simplicity and (ii) normal inverse Gaussian distribution for its ability to model the higher moments. According to backtesting procedure for each of 3,376 days (first 2,000 days was left for initial parameters estimation) we estimate the parameters of the models (both marginal distributions and copula functions), simulate 50,000 random returns for each asset and then estimate the portfolio VaR for chosen day. By comparison of estimated VaR with true losses we get the sequence of logical values indicating, whether the exception occurred.

The sums of exceptions for models with normal distribution and Gaussian/Student/Clayton/Gumbel/Frank copula function are summarized in Table 2. The numbers which can be statistically accepted on 10% probability level are denoted in bold. In italics we denoted the numbers of exceptions which are closest to the assumption, we call this cases as the winning. We can see, that the normal distribution is truly not the good model for marginal distributions. For $\alpha = 15\%$ the number of observed exceptions is low – the model overestimate the risk. For $\alpha = 1\%$ and $\alpha = 0.5\%$ the number of exceptions is too high to be statistically accepted. These probability levels are assumed in financial sector for risk estimation, thus normal distribution even coupled together with any copula can not be utilized for risk estimation in financial sector. On the other hand for probability level $\alpha = 5\%$ the normal distribution works good, while almost for all chosen portfolios the number of exceptions can be statistically accepted on 10% probability level. The best results are obtained when Gaussian or Frank copula function is utilized. As normal distribution and Gaussian copula function is nothing more than joint normal distribution, we can conclude, that for estimation of VaR (only) at 5% probability level the joint normal distribution is sufficient.

Portfolio	$\alpha = 0.005$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.15$
Assumption	16,88	33,76	168,8	506,4
DJI & USD	47/*41*/42/44/46	63/60/*55*/56/60	**173**/**174**/**162**/**160**/***172***	471/***477***/452/450/471
FTSE & GBP	57/56/*50*/52/54	80/78/*74*/75/79	**190**/191/***181***/**186**/**190**	456/*464*/449/448/453
N225 & JPY	38/35/*28*/31/34	53/50/***40***/*43*/52	**162**/**164**/127/131/***165***	437/443/375/379/*447*
SSMI & CHF	56/55/46/*45*/50	82/81/*71*/*71*/79	**186**/**184**/158/***161***/**185**	450/449/403/403/*455*
USD & GBP	39/36/*29*/50/53	57/55/*45*/67/69	***169***/168/158/179/174	462/469/470/470/442
USD & JPY	29/**23**/***19***/36/40	45/**43**/***35***/53/63	***168***/*168*/161/174/167	439/451/*457*/447/419
USD & CHF	46/*36*/38/41/48	69/*64*/65/69/74	**187**/194/***168***/**181**/**190**	456/*475*/439/435/460
Sum of significant cases	0/1/1/0/0	0/1/2/1/0	7/5/6/6/7	0/2/0/0/0
Sum of winning cases	0/2/4/1/0	0/1/6/1/0	2/1/2/1/2	0/3/2/1/2

Table 2 The numbers of exceptions for normal distribution and Gaussian/Student/Clayton/Gumbel/Frank copula functions. The closest numbers to the assumption is denoted in italics and on 10% significance level statistically significant numbers of exceptions are denoted in bold.

In Table 3 the quantity of exceptions for NIG model and various copula functions are summarized. Again we denoted the statistically acceptable numbers of exceptions in bold and the winning numbers of exceptions in italics. Compared to the normal distribution there is generally more cases for which the quantity of exceptions can be statistically accepted. For important probability levels 1% and 0.5% the NIG model coupled together with the Clayton copula can be accepted for almost all portfolios (except portfolio USD & CHF). Also utilizing Student copula function the results are very good (the model is insufficient for FTSE & GBP portfolio and again USD & CHF). Other copulas show worse results, but the numbers of exceptions are still half the numbers when normal distribution is utilized. Also for $\alpha = 5\%$ the results of Clayton copula function are very good – only for N225 & JPY the number of observer exceptions is, surprisingly, too low to be statistically accepted. On the other hand, for $\alpha = 15\%$ the results are poor – only for 4 out of 7 portfolios the results can be accepted for Gaussian/Gumbel/Frank copulas and results of Student/Clayton copulas are even worse.

From the results presented in Table 3 we can conclude that the most appropriate copula for VaR estimation is the Clayton copula function. Except the VaR at probability level $\alpha = 15\%$ the Clayton copula is the best from all chosen copulas in terms of the statistically accepted numbers of exceptions and the winning cases (i.e. the closest numbers of observed exceptions to the assumed numbers). Also Student copula function shows good results.

Portfolio	$\alpha = 0.005$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.15$
Assumption	16,88	33,76	168,8	506,4
DJI & USD	***24/23/21/22/22***	***35/36/36/37/36***	**190**/196/*172*/**179**/193	560/567/544/*542*/561
FTSE & GBP	25/*23*/**20**/22/25	45/44/***43***/46/44	193/194/***184***/**190**/191	559/566/*545*/548/561
N225 & JPY	**20/18/17/20/20**	***33/33/28/31/33***	***162***/**160**/126/136/**158**	**517**/**525**/444/445/**521**
SSMI & CHF	*17/17/14/16/17*	**41/42/*33*/35/36**	206/197/**172**/*169*/194	575/584/*535*/**536**/580
USD & GBP	**18**/*17*/**11/21**/26	38/*34*/**26**/46/47	**173**/174/*165*/**183**/175	*515*/**535**/548/**521**/495
USD & JPY	**24/22/*18*/30/31**	42/38/***33***/47/50	187/193/*182*/198/191	**527**/543/553/**535**/*499*
USD & CHF	27/*22*/24/22/29	52/47/***44***/50/54	200/207/***187***/192/207	**520**/544/*509*/499/**518**
Sum of significant cases	5/7/7/6/3	5/5/6/3/3	4/2/6/4/2	4/2/2/4/4
Sum of winning cases	1/3/4/1/1	2/2/4/0/1	1/0/5/1/0	2/0/3/1/1

Table 3 The numbers of exceptions for NIG model and Gaussian/Student/Clayton/Gumbel/Frank copula functions. The closest numbers to the assumption is denoted in italics and on 10% significance level statistically significant numbers of exceptions are denoted in bold.

4 Conclusion

Unexpectedly high decreases in the prices of financial assets are very challenging task for any risk model. In this paper we compared the risk models composed of NIG model (or Gaussian normal distribution) and various copula functions on the basis of the quantity of exceptions observed. From the provided results it is apparent that the most accurate risk estimations in terms of VaR at probability levels 0.5% and 1% are provided by NIG model and Clayton copula function. Also Student copula function provides good results. We also concluded, that for the estimation of VaR at 5% probability level the joint Gaussian normal distribution is sufficient.

Acknowledgements

References

[1] Alexander, C., and Sheedy, E: Developing a stress testing framework based on market risk models. *Journal of Banking and Finance* **32** (2008), 2220-2236.

[2] Applebaum, D.: *Lévy Processes and Stochastic Calculus.* Cambridge University Press, Cambridge, 2004.

[3] Barndorff-Nielsen, O. E.: *Normal inverse Gaussian distributions and the modeling of stock returns.* Aarhus University, Aarhus, 1995. Doctoral dissertation.

[4] Clayton, D. G.: A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* **65** (1978), 141-151.

[5] Cont, R., and Tankov, P.: *Financial Modelling with Jump Processes.* Chapman & Hall/CRC, London, 2003.

[6] Frank, M. J.: On the simultaneous associativity of F(x, y) and x+y-F(x, y). *Aequationes Mathematicae* **19** (1979), 194-226.

[7] Gumbel, E. J.: Bivariate exponential distributions. *Journal of the American Statistical Association* **55** (1960), 698-707.

[8] Cherubini, G., Luciano, E., and Vecchiato, W.: *Copula Methods in Finance.* Wiley, Chichester, 2004.

[9] Kresta, A., and Tichý, T.: International Equity Risk Modeling by NIG Model. In: *Mathematical methods in economics* (Dlouhý, M., and Skočdopolová, V., eds.). Professional Publishing, Prague, 2011, 401-406.

[10] Kresta, A., and Tichý, T.: International Equity Portfolio Risk Modeling: The case of NIG model and ordinary copula functions. *Finance a úvěr – Czech Journal of Economics and Finance* **61** (2012), 141–151.

[11] Kupiec, P.: Techniques for verifying the accuracy of risk measurement models. *Journal of Derivative* **3** (1995), 73-84.

[12] Nelsen, R. B.: *An Introduction to Copulas.* 2nd ed. Springer, New York, 2006.

[13] Rank, J.: *Copulas: From Theory to Application in Finance.* Risk Books, London, 2006.

[14] Sklar, A.: Fonctions de repartition à n dimensions et leurs marges. *Publications de l'Institut de statistique de l'Université de Paris* **8** (1959), 229-231.

[15] Tichý, T.: Posouzení odhadu měnového rizika portfolia pomocí Lévyho modelů. *Politická ekonomie* **58** (2010), 504-521.