MODELING BALTIC MARKET INDICES: A COMPARISON OF MODELS

I. Belovas
Vilnius University Institute of Data Science and Digital Technologies
Vilnius Gediminas Techncal University
Vilnius, LITHUANIA
e-mail: igoris.belovas@mii.vu.lt

Abstract

In this paper we perform a statistical analysis of the returns of Baltic market indices. We construct symmetric α-stable, non-standardized Student's t and normal-inverse Gaussian models, using maximum likelihood method for the estimation of the parameters of the models. The adequacy of the modeling is evaluated with the Kolmogorov tests for composite hypothesis. The results of the study indicate that the normal-inverse Gaussian model provides the best overall fit for the data.
Keywords: data science, market index, α-stable model

1 Introduction

Modeling of the returns of stock indices of developed and emerging markets always has been significant and controversial topic, since principal models in financial theory (mean-variance portfolio, capital asset pricing, prices of derivative securities) critically rely on underlying stock returns distribution form. A summary of the literature, covering the history of mainstream models, one may find in [1, 5, 6, 9] and references therein.

In our previous research [2] we have investigated the goodness-of-fit of ten major models (normal, mixture of normals, Student's t, logistic, exponential power, mixed diffusion-jump model, normal-inverse Gaussian, scaled symmetrized gamma, α-stable and symmetric α-stable) to five Standard & Poor's stock market indices. These indices covered the period of ten years (from 2006-04-28 to 2016-05-31). Only α-stable, Student's-t and normal-inverse Gaussian distributions properly described daily returns of all five Standard & Poor's indices.

In this paper we compare three above mentioned distributions, representing empirical returns of three Baltic market indices. These indices (OMX Tallinn, OMX Riga and OMX Vilnius) are the components of the Nasdaq Baltic index family. They include the shares listed on the Main and Secondary lists of the Baltic exchanges and reflect the current status and changes in each market or the Baltic Market as a whole.

2 Candidate models

Non-standardized Student's t-distribution (NSS) is a three-parameter generalization of classical Student's t-distribution ("arguably the simplest and the most well known model for stock returns" [1]) with the density function

$$f_{NSS}(x;\nu,\mu,\sigma)=\frac{1}{\sigma}f_S\left(\frac{x-\mu}{\sigma};\nu\right)=\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\sigma\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{1}{\nu}\left(\frac{x-\mu}{\sigma}\right)^2\right)^{-\frac{\nu+1}{2}},$$

where $f_S(x;\nu)$ is Student's t-distribution (with $\nu>0$ degrees of freedom) probability density function, μ is a location parameter, and $\sigma>0$ is a scale parameter [1].

Symmetric Normal-Inverse Gaussian distribution (SNIG) has the density function [5]

$$f_{SNIG}(x;\alpha,\mu,\sigma)=\frac{\alpha\sigma e^{\alpha\sigma}}{\pi\sqrt{\sigma^2+(x-\mu)^2}}K_1\left(\alpha\sqrt{\sigma^2+(x-\mu)^2}\right),$$

where $K_1(x)$ denotes the modified Bessel function of the third kind, $\alpha>0$ is a tail heaviness (shape) parameter, μ is a location parameter, and $\sigma>0$ is a scale parameter. In order to standardize the NIG distrubution, we modify the parametrization of the distribution by setting $\bar{\alpha}=\alpha\sigma$. The modified (now scale-invariant) representaion (MSNIG) has the density function (cf. [6])

$$f_{MSNIG}(x;\bar{\alpha},\mu,\sigma)=\frac{1}{\sigma}f_{SMSNIG}\left(\frac{x-\mu}{\sigma};\bar{\alpha}\right),$$

where $f_{SMSNIG}(x;\bar{\alpha})$ stands for the standard MSNIG density,

$$f_{SMSNIG}(x;\bar{\alpha})=\frac{\bar{\alpha}e^{\bar{\alpha}}}{\pi\sqrt{1+x^2}}K_1\left(\bar{\alpha}\sqrt{1+x^2}\right)=\frac{\bar{\alpha}e^{\bar{\alpha}}}{2\pi}\int_0^\infty t^{-2}e^{-\frac{\bar{\alpha}}{2}\left(t+\frac{1+x^2}{t}\right)}dt.$$

Thus, the cumulative distribution function of SMSNIG distribution can be written in terms of the cumulative distribution function of the standard normal distribution $\Phi(x)$,

$$F_{SMSNIG}(x;\bar{\alpha})=\sqrt{\frac{\bar{\alpha}}{2\pi}}e^{\bar{\alpha}}\int_0^\infty t^{-3/2}\exp\left(-\frac{\bar{\alpha}}{2}\left(t+\frac{1}{t}\right)\right)\Phi\left(x\sqrt{\frac{\bar{\alpha}}{t}}\right)dt.$$

Symmetric α-stable distribution (SαS) has the density function [4]

$$f_{S\alpha S}(x;\alpha,\mu,\sigma)=\frac{1}{\sigma}f_{SS\alpha S}\left(\frac{x-\mu}{\sigma};\alpha\right)=\frac{1}{\pi\sigma}\int_0^\infty e^{-t^\alpha}\cos\left(\frac{x-\mu}{\sigma}t\right)dt,$$

where $\alpha\in(0,2]$ is a stability parameter, μ is a location parameter, $\sigma>0$ is a scale parameter and $f_{SS\alpha S}(x;\alpha)$ stands for the standard SαS density. The cumulative distribution function for the standard SαS distribution is

$$F_{SS\alpha S}(x;\alpha)=\frac{1}{2}+\frac{1}{\pi}\int_0^\infty t^{-1}e^{-t^\alpha}\sin(xt)dt.$$

We estimate parameters of these three models by the maximum likelihood method, maximizing the log-likelihood function (ML),

$$\hat{l}(\Theta)=\frac{1}{n}\sum_{k=1}^n\log f(x_k;\Theta), \tag{1}$$

where Θ is a vector of parameters. It is the most accurate (albeit time consuming) method, if applied without parallel computing (cf. [3, 4]).

3 Empirical data

The empirical data sets under consideration are daily logarithmic returns of Baltic stock market indices beloging to the Nasdaq Baltic index family. The countries and their representative indices are: Estonia (OMX Tallinn), Latvia (OMX Riga) and Lithuania (OMX Vilnius) [8]. These indices include the shares listed on the Main and Secondary lists of the Baltic exchanges and reflect the current status and changes in each market or the Baltic Market as a whole. The data covers the period of ten years (from 2009-07-01 to 2019-02-19) with the lengths of the series $n = 4878$.

4 Results

The estimated parameters of the models, obtained by maximizing the log-likelihood function (1), can be found in Table 1. To quantify the goodness-of-fit of the model-

Table 1: Estimated parameters and goodness-of-fit statistics ("*" means rejected).

2009-2019		Estimated parameters, $\hat{\Theta}$				Goodness-of-fit st.	
Index	Model	Shape	Location	Scale	$\hat{l}(\Theta)$	$D_n(\hat{\Theta})$	$D_P(n,\hat{\Theta})$
OMX Tallinn	SαS	1.3951	0.0004	0.0040	3.4052	0.0152	0.0153
	NSS	2.1349	0.0004	0.0048	3.4139	0.0141	0.0155
	MSNIG	0.2520	0.0004	0.0049	3.4184	0.0092	0.0182
OMX Riga	SαS	1.4922	0.0003	0.0055	3.1593	0.0179*	0.0152
	NSS	2.1035	0.0003	0.0066	3.1686	0.0131	0.0149
	MSNIG	0.3497	0.0003	0.0070	3.1713	0.0108	0.0208
OMX Vilnius	SαS	1.3983	0.0004	0.0031	3.6432	0.0167*	0.0156
	NSS	2.0108	0.0004	0.0037	3.6483	0.0161*	0.0153
	MSNIG	0.1928	0.0003	0.0036	3.6480	0.0162	0.0191

ing, Kolmogorov tests are used. However, for a composite hypothesis testing (a model belongs to a distribution family, and parameters of a model are estimated from the sample we use to quantify the goodness-of-fit) the classical Kolmogorov test is not applicable, since the limiting distribution of the Kolmogorov test statistics no longer distribution-free. It is influenced by the law, corresponding to the null hypothesis, the type and the number of parameters of the law, the method of parameter estimation. Because of that, critical values $D_P(n, \hat{\Theta})$ of Kolmogorov tests for composite hypotheses are calculated (with the significance level $P = 0.05$) by Monte-Carlo methodology, proposed by Lemeshko et al. [7]. Results of the comparison of the models are summarized in Table 1.

5 Conclusions and discussion

As we can see from Table 1, the symmetric α-stable model performed the worst. Two time it was rejected. We can see that the non-standardized Student's t and the modified symmetric normal-inverse Gaussian distribution had better goodness-of-fit values.

The normal-inverse Gaussian outperforms alternative heavy-tailed models (note that it corroborates with the recent findings, see [9]), while the non-standardized Student's t model provides the second best overall fit for the data (cf. [5]).

It should be noted that the four-parameter normal-inverse Gaussian family could be reduced to the three-parameter symmetric normal-inverse Gaussian or modified symmetric normal-inverse Gaussian model without much loss (cf. [9]). The problem of stability of the distribution of returns over different time periods requires special attention.

References

[1] Afuecheta E., Chan S., Nadarajah S. (2018) Flexible models for stock returns based on Student's t distribution. *The Manchester School*, pp. 1–25.

[2] Belovas I., Sakalauskas L. (2016). Modeling financial data distributions: A comparison of models. In: *Computer Data Analysis and Modeling: Theoretical and Applied Stochastics: Proceedings of the XI International Conference, Minsk, September 6-10, 2016*, pp. 194–197.

[3] Belovas I., Starikovičius V. (2007). Parallelization of α-stable modelling algorithms. *Mathematical Modelling and Analysis*. Vol. **12**(4), pp. 409–418.

[4] Belovas I., Starikovičius V. (2015). Parallel computing for mixed-stable modelling of large data sets. *Information Technology and Control*. Vol. **44**(2), pp. 148-154.

[5] Corlu C.G., Corlu A. (2015) Modelling exchange rate returns: which flexible distribution to use? *Quantitative Finance*. Vol. **15**(11), pp. 1851–1864.

[6] Eriksson A., Ghysels E., Wang F. (2009) The normal inverse Gaussian distribution and the pricing of derivatives. *The Journal of Derivatives*. Vol. **16**(3), pp. 23–37.

[7] Lemeshko B., Lemeshko S., Rogozhnikov A. (2011) Real-Time Studying of Statistic Distributions of Non-Parametric Goodness-of-Fit Tests when Testing Complex Hypotheses. *The International Workshop "Applied Methods of Statistical Analysis. Simulations and Statistical Inference" – AMSA-2011*. Vol. **1**, pp. 19–27.

[8] Nasdaq Baltic. Mode of Access: https://www.nasdaqbaltic.com/market/. Date of Access: 18.05.2019.

[9] Suárez-García P., Gómez-Ullate D. (2013) Scaling, stability and distribution of the high frequency returns of the IBEX 35 index. *Physica A*. Vol. **392**(6), pp. 1409-–1417.