Trendy v podnikání, roč. 6, číslo 3, s. 21-27, 2016.
Business Trends, vol. 6, no. 3, p. 21-27, 2016.

INVESTING IN HIGH FREQUENCY DATA

Jiří Málek[1], Tran Van Quang[2]

[1] doc. Jiří Málek, Ph.D., Faculty of Finance and Accounting, University of Economics Prague, malek@vse.cz.
[2] Ing. Tran Van Quang, CSc, Ph.D., Faculty of Finance and Accounting, University of Economics Prague, tran@vse.cz

Abstract: The article analyzes the 10 min high frequency data of certain financial instruments (gold, exchange rate USD / Euro, stock quotes Boeing and Microsoft). As approximation of the empirical distribution of log-returns is used alpha stable distribution. This distribution allows to measure "power" tails and then compare them with the Gaussian distribution. The parameters of stable distributions are estimated using methods MLE (Maximum Likelihood Estimation), which proved to be the most accurate. At the same time four basic moments (mean, standard deviation, skewness and kurtosis) are also estimated .Using of the analysis of all these parameters are given investment characteristics considered instruments.

Keywords: High frequency data, skewness, kurtosis, α-stable distribution

JEL Classification: G10, G120

INTRODUCTION

Normal distribution due to its simplicity and easy processing is often used for modeling the returns on financial instruments. However, it appears that the tails of the empirical distributions tend to be thicker than the normal ones, which in turn has implications in the management of financial risks: the probability of a loss is much higher than in a normal distribution.

The heavy tailed character of the distribution of financial asset price has been repeatedly observed in various markets. In response to the empirical evidence Mandelbrot (1963) and Fama (1965) proposed the stable distribution as an alternative model to Gaussian law. Although there are other heavy-tailed alternatives to the Gaussian law, like Student's t, hyperbolic distribution, there is at least one good reason for modeling financial variables using stable distributions. Namely, they are supported by the generalized Central Limit Theorem, which states that stable laws are the only possible limit distributions for sum of properly normalized of independent, identically distributed random variables. Since stable distributions can model the fat tails and asymmetry, give a good fit to empirical data and even extreme events like market crash. Even though they are not universal, they are a useful tool of an analyst working in finance.

Examples of using stable distribution in economics and finance can be found for example in Mandelbrot (1963), Fama (1965), Embrechts at al. (1997), Cheng and Rachev (1995), McCoulloch (1996), Bouchaud and Potters (2000), Carr et al., (2002) Guillaume et al. (1997) Mantegna and Stanley (1995), Rachev, (2003); Weron, (2004), Nolan (2012)

In this paper we analyze the log-returns of some different instruments of international financial market (exchange rates USD/EUR, gold, and Boeing and Microsoft stocks) in 10 minutes high frequency data. Using Maximum Likelihood Estimation method (MLE) the estimates of the tail indexes and subsequent stable distribution parameters are derived.

1. DEFINITION OF STABLE DISTRIBUTION

Let $X, X_1, X_2, X_3, \ldots, X_n \ldots$ are independent and identically distributed (i.i.d) random variables. A random variable X is said to have the α-stable distributions if there is for any $n \geq 2$ a positive number cn and a real number dn such that

$$X_1 + X_2 + \cdots + X_n \overset{d}{=} c_n X + d_n \qquad (1)$$

This expression means, that any sum of i.i.d. random variables have the same distribution except for the „mean"and „variance".

Unfortunately there is no general form of the probability density function (pdf), we know only the general form of the characteristic function:

$$\Phi(t) = \exp\left\{-\sigma^\alpha |t|^\alpha \left(1 - i\beta \operatorname{sgn}(t) \tan\frac{\pi\alpha}{2}\right) + i\mu t\right\} \quad for \;\; \alpha \neq 1$$

$$\Phi(t) = \exp\left\{-\sigma^\alpha |t| \left(1 - i\beta \frac{2}{\pi} \operatorname{sgn}(t) \log|t|\right) + i\mu t\right\} \quad for \;\; \alpha = 1 \qquad (2)$$

where

α …tail power (tail index), as α decreases tail thickness increases,

β…skewness parameter, determines asymmetry, a positive β indicates that right tail is father than left one and vice versa, β = 0 corresponding to a symmetric distribution,

μ…location parameter, corresponding to mean value for α > 0,

σ…scale parameter, generalized standard deviation, for α = 2 corresponding to a standard deviation of normal distribution.

There is another (equivalent) parametrization of the characteristic function that differs only in location parameter

$$\Phi(t) = \exp\left\{-\sigma^\alpha |t|^\alpha \left(1 - i\beta \operatorname{sgn}(t) \tan\frac{\pi\alpha}{2}\right)\left(|\sigma t|^{1-\alpha} - 1\right) + i\mu_1 t\right\} \quad for \;\; \alpha \neq 1$$

$$\Phi(t) = \exp\left\{-\sigma^\alpha |t| \left(1 - i\beta \frac{2}{\pi} \operatorname{sgn}(t) \log|\sigma t|\right) + i\mu_1 t\right\} \quad for \;\; \alpha = 1 \qquad (3)$$

Location parameters are related by

$$\mu = \mu_1 - \beta\sigma \tan\frac{\pi\alpha}{2} \quad for \;\; \alpha \neq 1 \qquad (4)$$

$$\mu = \mu_1 - \frac{2\beta}{\pi}\sigma \log\sigma \quad for \;\; \alpha = 1$$

Parametrization (3) is preferred from the computational view in the case the parameter α is very close to one. But this is not our situation so we will use parametrization (2).

Although there is no general explicit formula for pdf for some special cases exists:

1. Normal distribution with the pdf

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \qquad (5)$$

is a stable distribution $S(2, 0, \sigma/2, \mu)$

2. Cauchy distribution with the pdf

$$f(x) = \frac{1}{\pi} \frac{\sigma}{\sigma^2 + (x-\mu)^2} \qquad (6)$$

is a stable distribution $S(1, 0, \sigma, \mu)$

3. Lévy distribution with pdf

$$f(x) = \sqrt{\frac{\sigma}{2\pi}} \frac{1}{(x-\mu)^{3/2}} \exp\left(-\frac{\sigma}{2(x-\mu)}\right) \qquad (7)$$

is a stable distribution $S\left(\frac{1}{2}, 1, \sigma, \mu\right)$

1.1 Properties of stable distributions

The following properties holds for stable distributions:

1. Let X_1, X_2 are the independent stable random variables, with $X_i = S(\alpha, \beta_i, \sigma_i, \mu_i)$ $i = 1,2$ then $X_1 + X_2 \approx S(\alpha, \beta, \sigma, \mu)$ with

$$\sigma = \sigma_1 + \sigma_2$$
$$\beta = \frac{\beta_1 \sigma_1^\alpha + \beta_2 \sigma_2^\alpha}{\sigma_1^\alpha + \sigma_2^\alpha} \qquad (8)$$
$$\mu = \mu_1 + \mu_2$$

2. If $X \approx S(\alpha, \beta, \sigma, \mu)$ and $a \in R$ then

$$X + a \approx S(\alpha, \beta, \sigma, \mu + a) \qquad (9)$$

3. If $X \approx S(\alpha, \beta, \sigma, \mu)$ and $a \in R$ then

$$aX \approx S(\alpha, \operatorname{sgn}(a)\beta, |a|\sigma, a\mu) \qquad (10)$$

2. MAXIMUM LIKELIHOOD ESTIMATION

According Borak, Hardle, Weron (2005), after substitution $\zeta = -\beta \tan \frac{\pi\alpha}{2}$ the density of standard α- stable random variable (μ=0, σ =1) for α≠1 can be expressed as: for $x > \zeta$:

$$f(x;\alpha,\beta) = \frac{\alpha(x-\zeta)^{\frac{1}{\alpha-1}}}{\pi|\alpha-1|} \int_{-\xi}^{\frac{\pi}{2}} V(\theta;\alpha,\beta) \exp\left(-(x-\zeta)^{\alpha/\alpha-1} V(\theta;\alpha,\beta)\right) d\theta, \qquad (11)$$

for $x = \zeta$:

$$f(x;\alpha,\beta) = \frac{\Gamma\left(1+\frac{1}{\alpha}\right)\cos\xi}{\pi(1+\zeta^2)^{\frac{1}{2\alpha}}} \qquad (12)$$

and for $x < \zeta$:

$$f(x;\alpha,\beta) = f(-x;\alpha,-\beta) \qquad (13)$$

Where

$$V(\theta;\alpha,\beta) = (\cos\alpha\xi)^{\frac{1}{\alpha-1}} \left(\frac{\cos\theta}{\sin\alpha(\xi+\theta)}\right)^{\alpha/\alpha-1} \frac{\cos[\alpha\xi(\alpha-1)\theta]}{\cos\theta} \qquad (14)$$

$$\xi = \frac{1}{\alpha}\arctan(-\zeta) \qquad (15)$$

In MLE we have to find from observation data x_i a maximum of the likelihood function

$$\sum_{i=1}^{n} \log f(z_i;\alpha,\beta,\delta,\mu) \qquad (16)$$

with respect to parameters $\alpha, \beta, \delta, \mu$, where $z_i = \frac{x_i - \mu}{\delta}$.

3. DATA AND RESULTS

The 10 min data used include the time period from December 11, 2015 to March 11, 2016 for Boeing, US $ / Euro and gold and from 7 December 2015 to March 4, 2016 for Microsoft.

In the following tables we show descriptive statistics original series and log-returns of this ones. The table 3 contains estimated parameters of stable distributions as the best approximation of the empirical ones. The figures are shown all the empirical distribution histograms compared to the corresponding Gaussian and stable distribution.

Table 1: Descriptive statistics of original series

	GOLD	USDEUR	MSFT	BOEING
Mean	1142.680	1.096815	52.90656	126.8995
Median	1112.460	1.093000	52.44000	122.9250
Maximum	1282.510	1.137400	56.81500	149.7400
Minimum	1049.910	1.071300	48.36900	102.6000
Std. Dev.	71.35896	0.013163	2.029411	11.19913
Skewness	0.468060	0.893298	0.060464	0.512325
Kurtosis	1.607672	3.087842	2.004403	2.048421
Obs.	8010	9377	2349	2359

Source: own

Table 2: Descriptive statistics of logarithmic returns series

	GOLD	USDEUR	MSFT	BOEING
Mean	1.96E-05	1.57E-06	-2.77E-05	-6.06E-05
Median	0.000000	0.000000	0.000000	0.000000
Maximum	0.008019	0.007490	0.049467	0.020149
Minimum	-0.008663	-0.008578	-0.034708	-0.080866
Std. Dev.	0.000949	0.000553	0.003209	0.003635
Skewness	-0.015408	0.003704	0.404325	-6.277240
Kurtosis	9.998906	28.08650	42.62444	125.7190
Obs.	8009	9376	2348	2358

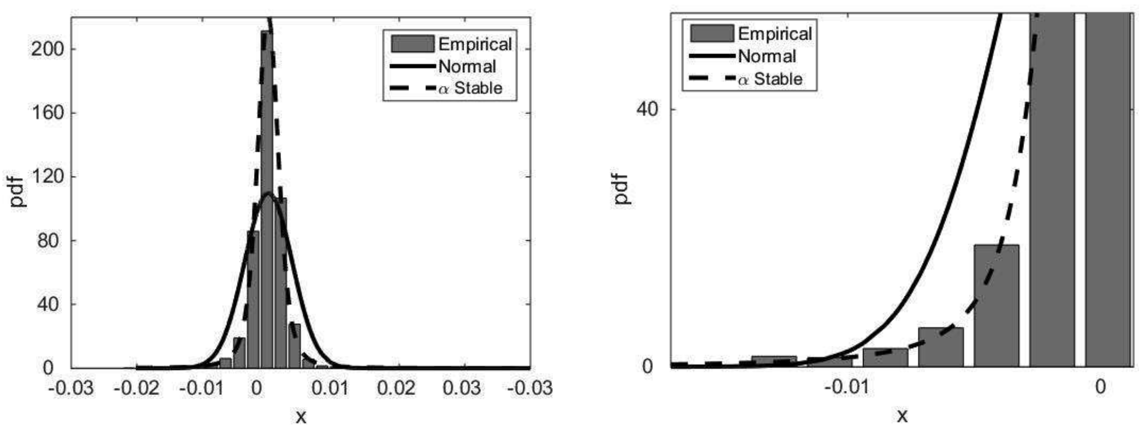
Source: own

Table 3: Estimation results: parameter values

Series	alpha	beta	sigma	mu
GOLD	**1.5381071**	-0.0155465	4.5955e-04	-2.2531e-05
USDEUR	1.6350977	0.0269907	2.7962e-04	2.8436e-06
MICROSOFT	1.6247460	0.0090409	1.5120e-03	2.0713e-05
BOEING	1.5414582	-0.0041070	1.2944e-03	-8.9006e-06

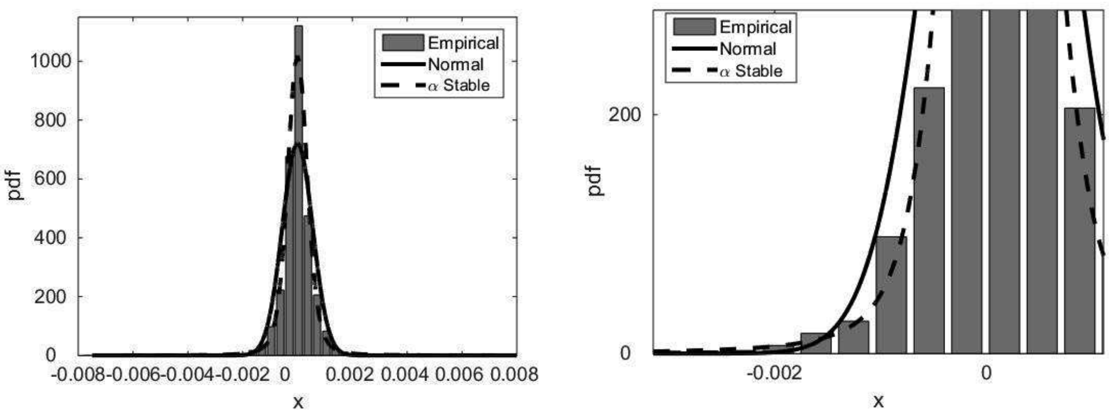
Source: own

Fig. 1: Boeing log-returns histograms



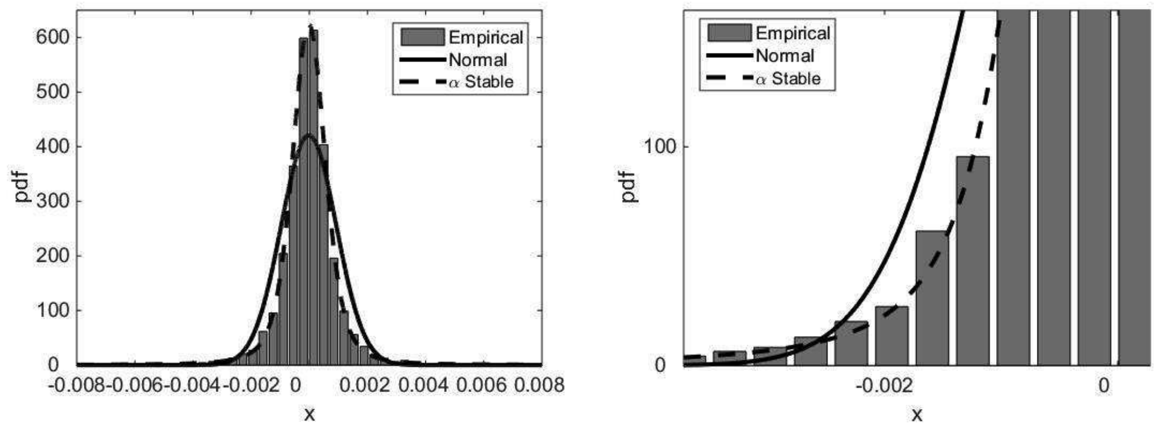
Source: own

Fig. 2: Ex. Rate USD/EUR log returns histograms



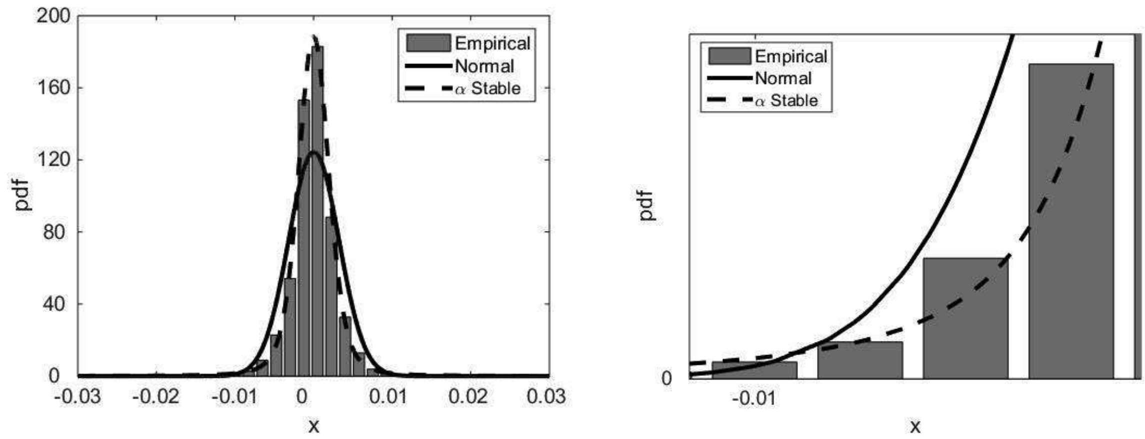
Source: own

Fig. 3: Gold log-returns histograms



Source: own

Fig. 4: Microsoft log-returns histograms



Source: own

4. INTERPRETATION AND DISCUSSION

It turns out that the probability distribution of log-returns all considered assets differs significantly from the normal distribution. The smallest tail index has gold, closely followed by Boeing. All α are in the interval [1.5381071 1.6350977]

Boeing stocks show the biggest negative change which also have the second largest positive change. This stocks are proving (in the considered period) very volatile. This corresponds to the highest standard deviation and a large negative skewness and the highest kurtosis, which is 125,7190, almost three times larger than the second largest Microsoft stocks (42. 6244). Additionally, estimated mean value is negative (-6.06E-05). All correspond with the relevant stable distribution parameters

Gold has thickest tails even if the tail index of other titles is very close. Skewness (and the corresponding parameter β) are negative. Estimated return is positive, but the coefficient μ is negative. Estimated standard deviation and the coefficient δ are relatively low. However, as an investment instrument gold seems relatively risky, since low alpha coefficient indicates the possibility of significant losses.

Microsoft has the highest coefficient α, positive skewness, but a higher standard deviation and coefficient sigma. Although the expected return is positive but μ coefficient is negative. In our opinion, it is better to follow the estimated rate of return. In this respect investment appears to be a moderate risky

Exchange rate USD / EUR seems the best from the considered investment. Its coefficient α is very close to Microsoft, skewness slightly positive (second highest), β coefficient is high. Standard deviation is the smallest and also the corresponding coefficient δ. Estimated return and coefficient μ are positive.

CONCLUSION

We have analyzed the distribution of ten minute intraday price data of four different financial assets: gold, exchange rate EURUSD and two stocks Boeing and Microsoft from the stable distribution perspective. The four parameters of this distribution have been estimated from data using the maximum likelihood method. As the tail parameter α of the distribution of returns of these instruments are substantially different from 2, it is clear that the returns obtained from intraday data at frequency ten minute also exhibit the fat tail phenomenon. From the investment perspective, the results of our analysis can be also used to make some relevant inference as they bear important more accurate information on behavior of each individual assets concerning their expected profit, volatility and the possible occurrence of extreme events.

ACKNOWLEDGMENTS

Jiří Málek has been supported by the Czech Science Foundation Grant 15-00036S "Credit Risk Modeling for Financial and Commodity Assets Portfolios" and by the VŠE institutional grant IP100040. Tran Van Quang has been supported by the VŠE institutional grant IP100040.

REFERENCES

Borak, S. and Härdle, W. and Weron,R.(2005): *Stable Distributions*, SFB649 Discussion Papers SFB 649DP2005-008, Sonderforschungsbereich 649, Humboldt University, Berlin, Germany

Bouchaud, J. P., Potters, M. (2000): *Theory of Financial Risk,* Cambridge University Press, Cambridge.

Carr, P., Geman, H., Madan, D. B., and Yor, M. (2002): The fine structure of asset returns: an empirical investigation, *Journal of Business* 75: 305-332.

Embrechts, P., Kluppelberg, C., and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance, Springer*.

Fama, E. F. (1965). The behavior of stock market prices*, Journal of Business* 38: 34-105.

Guillaume, D. M., Dacorogna, M. M., Dave, R. R., Muller, U. A., Olsen, R. B., and Pictet, O. V. (1997): From the birds eye to the microscope: A survey of new stylized facts of the intra-daily foreign exchange markets, *Finance & Stochastics* 1: 95-129.

Mandelbrot, B. B. (1963). The variation of certain speculative prices, *Journal of Business* 36: 394-419.

McCulloch, J. H. (1996). Financial applications of stable distributions, in G. S. Maddala, C. R. Rao (eds.), *Handbook of Statistics,* Vol. 14, Elsevier, 393-425.

Mantegna, R. N., Stanley, H. E. (1995). Scaling behavior in the dynamics of an economic index, *Nature* 376: 46-49.

Nolan, J. P. (2012). Maximum likelihood estimation and diagnostics for stable distributions, in O. E. Barndorff-Nielsen, T. Mikosch, S. Resnick (eds.), *Levy Processes*, Brikhauser, Boston.

Rachev, S., ed. (2003). *Handbook of Heavy-tailed Distributions in Finance*, North Holland.

Weron, R. (2004). Computationally intensive Value at Risk calculations, in J. E. Gentle, W. Hardle, Y. Mori (eds.) *Handbook of Computational Statistics,* Springer, Berlin, 911-950.