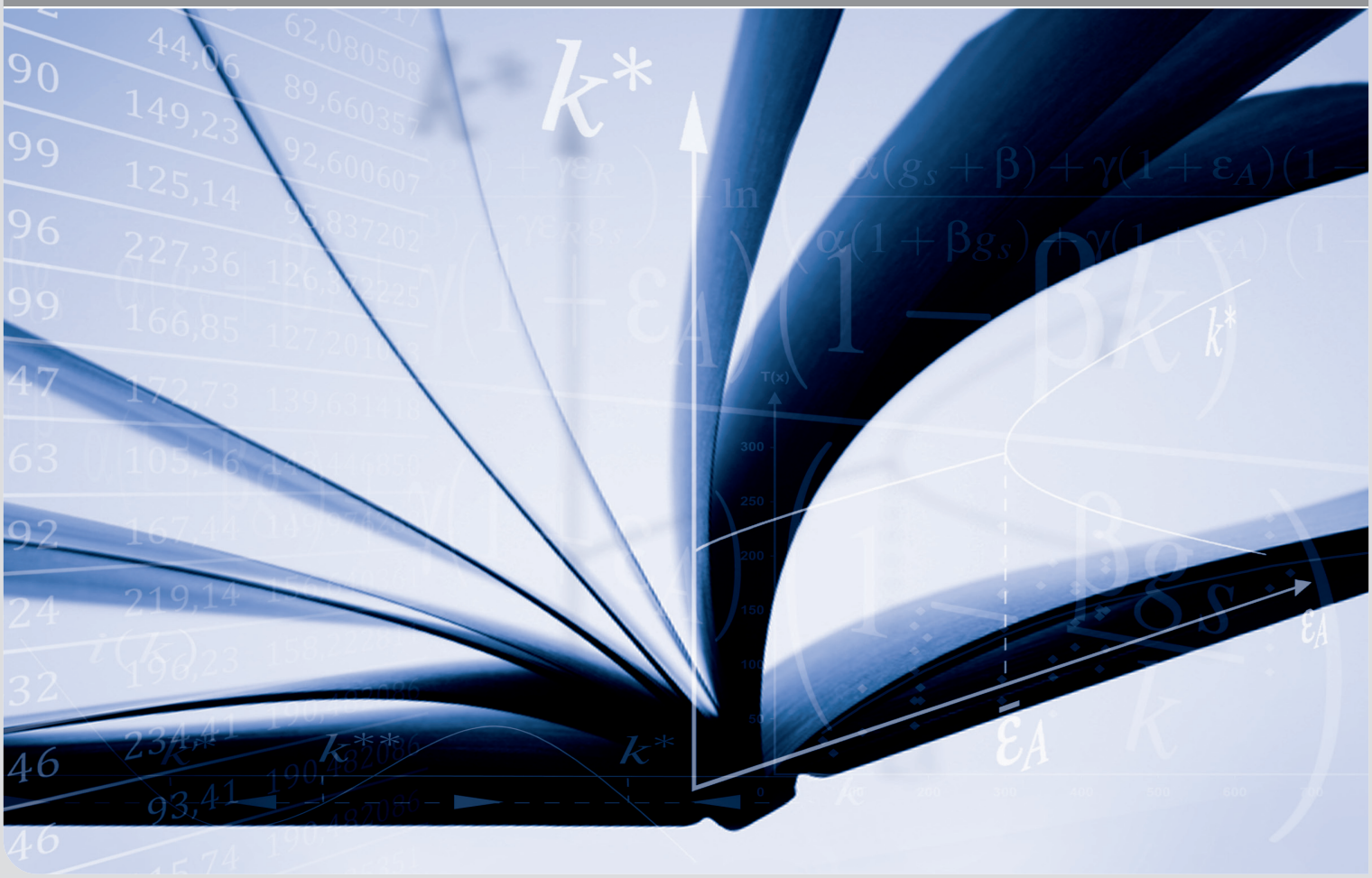
Karlsruher Institut für Technologie

On portfolio risk estimation

by Mher Safarian

No. 52 | DECEMBER 2013

WORKING PAPER SERIES IN ECONOMICS



KIT – University of the State of Baden-Wuerttemberg and
National Laboratory of the Helmholtz Association

econpapers.wiwi.kit.edu

Impressum

Karlsruher Institut für Technologie (KIT)
Fakultät für Wirtschaftswissenschaften
Institut für Volkswirtschaftslehre (ECON)

Schlossbezirk 12
76131 Karlsruhe

KIT – Universität des Landes Baden-Württemberg und
nationales Forschungszentrum in der Helmholtz-Gemeinschaft

Working Paper Series in Economics
No. 52, December 2013

ISSN 2190-9806

econpapers.wiwi.kit.edu

On Portfolio Risk Estimation

Mher Safarian*

Karlsruhe Institute of Technology (KIT)

Institute of Economics (ECON)

Chair of Econometrics and Statistics

December 3, 2013

*E-mail: mher.safarian@kit.edu, Schlossbezirk 12, 76131 Karlsruhe

Contents

1 Statement of the problem

The present work considers the problem of investment portfolio risk estimation, including dynamic adjustment for each new transaction. Any Bank portfolio has a complex structure. It consists of stocks, bonds and a set of derivative securities. A portion of bonds and loans is riskless. For some of these assets, the methods offered cannot be applied without additional consideration of the term structure of interest rates and credit risks features. The risk estimation of this part of the portfolio containing some peculiar financial tools represents a separate issue, solving which exceeds the limits of the present research.

We use as an estimation of a portfolio risk the amount of probable losses that can be sustained in case of a complete asset sale, related to current market value of these assets.

The investment portfolio includes a number of shares, sale of which can significantly affect the market for a brief period of time, *making the calculated estimation of risk insolvent*. Thus it is necessary to estimate the quantity of shares that can be sold without having a material influence on the prices dynamics. Knowing this size, it is easy to calculate a time interval during which this portfolio can be sold. Definition of the stability of the *concrete market* is directly concerned with its specificity. This represents a separate practical problem, which is not considered in the submitted paper.

Consequently, for a portfolio risk calculation, it is necessary to estimate dynamics of price behaviour for the time period during which controllable realization of portfolio assets is possible. Such an approach is described in many papers where estimation of risk is based on studying prices dynamics of stocks included in a portfolio (VaR - "$RiskMetrics$", $RiskManagement^{+}$).

However, forecasting such processes represents a complicated problem itself. For example, on NASDAQ the prices of the most liquid stocks have large volatility. Deviation from average value of a stock price can run up to several percent even on ordinary days. To circumvent this problem, a new approach, which is not considered earlier, is offered in the given research.

2 Analysis of the empirical data

2.1 Volume weighted average price (VWAP)

Volume weighted average price (VWAP) is an analysis tool professional traders systematically use to assist themselves in decision making. To define this value let us consider a sequence of transactions with the prices $p_1, \ldots, p_n$ and appropriate volumes $V_1, \ldots, V_n$, then VWAP is calculated as follows:

$$(VWAP)_n = \frac{\sum\limits_{i=1}^{n} p_i V_i}{\sum\limits_{i=1}^{n} V_i}.$$

For example, if there are two transactions for the KRFT symbol with the prices 56\$, 55\$ and volumes 200 stocks and 300 stocks accordingly, then the VWAP price for KRFT is equal to

$$55.4\$ = \frac{56 * 200 + 55 * 300}{200 + 300}.$$

The VWAP is often a benchmark for traders. If you bought a stock today at a price lower than the current cumulative VWAP, you bought the stock at a good price (you did better than the average price of the asset). On the other hand, if you sell a stock higher than the current cumulative VWAP you again did better than the market is.

The stability to sharp jumps is a distinctive feature of such representation of the price. This property is perfectly illustrated in figure 1. It is obvious, that VWAP diagram is much smoother than the diagram of the price. One of the most important shortcomings of VWAP representation for the stock price is the impossibility to determine the actual stock price at any moment of time without additional information. However, VWAP allows to essentially reduce a degree of estimated risk due to it's stability and small volatility in comparison with behaviour of an actual stock price.

Despite the obvious advantages of the average price, there has not been developed a technical system of selling stocks at VWAP. The system developed by us allows to sell stocks for some period of time intraday in such a manner that with some error the average price of transactions coincides with VWAP for the same period of time. This program is used at work on NASDAQ and has shown the following results. For example, during the work since December 12 2012 to February 8 2013 the statistics for a difference between real VWAP value and the value forecasted by this system was as follows:

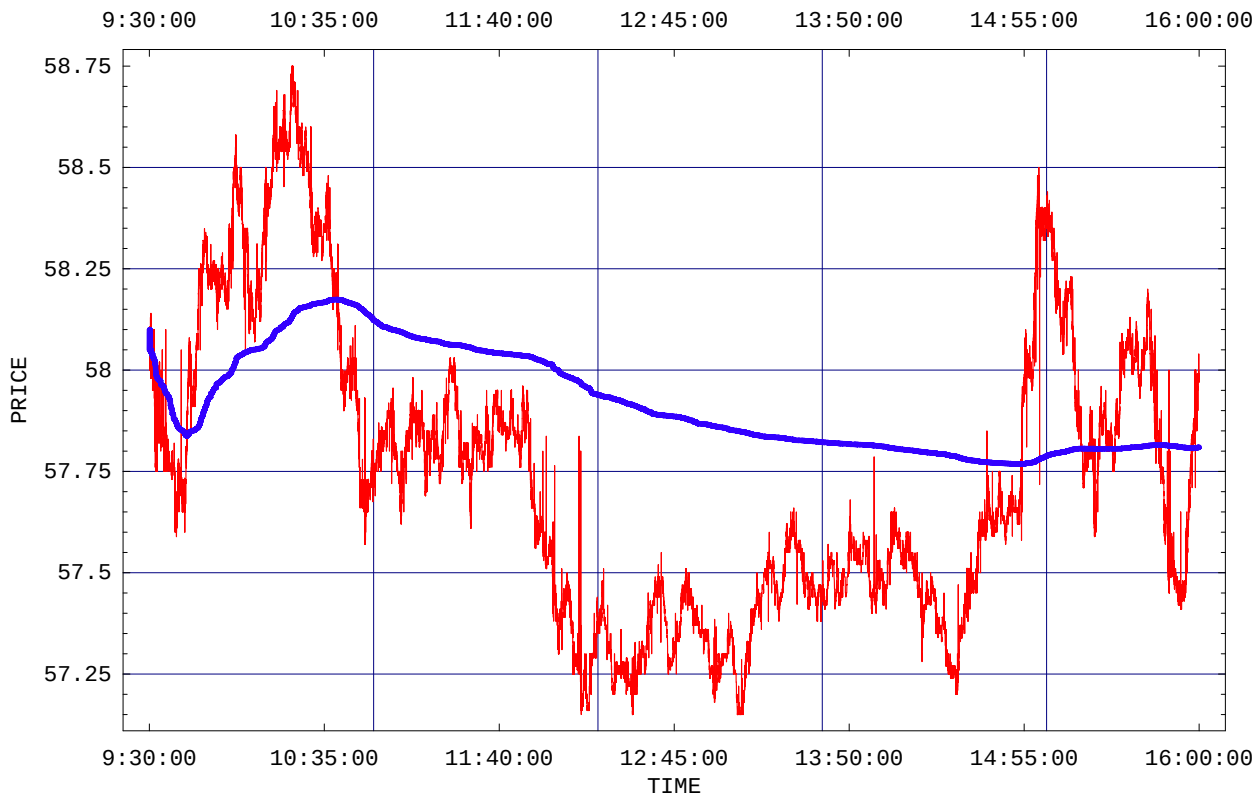


Figure 1: *Diagram of the transaction prices and VWAP for the Kraft Foods Groups stock on NASDAQ (July, 22, 2013). The dynamics of the stock price is represented with red color, and VWAP behavior is represented by dark blue line*

mean: 0,

standard deviation: 5,8 cents.

There was used a sample of 20 stocks from Nasdaq100 for this research.

From the analytical point of view what plays the main role is the assumption about normality (gaussianity) of brownian motion distribution. Thus, naturally, there is a question about the degree of compatibility of the real statistical data with (1) hypothesis about the normality and (2) hypothesis about the independence of successive increments for $H_t = \ln S_t$.

Traditionally the statistical analysis for the given case is based on the analysis of $y_t = H_t - H_{t-1} = \ln \frac{S_t}{S_{t-1}}$ or $x_t = \frac{S_t}{S_{t-1}} - 1$ where $S_1, S_2, \ldots$ are the successive prices of some asset.

Under the hypothesis of fitness of $S = (S_t)$ to the geometrical brownian motion, the sequence $y = (y_t)$, $t = 1, 2, \ldots$, consists of identically independent distributed random variables, $y_t \sim \mathcal{N}(\mu, \sigma)$.

In the financial literature devoted to analysis of properties for $y = y(y_t)$, $t = 1, 2, \ldots$, sequence, it was repeatedly specified, that the assumption about normality is rather disputable,

due to some properties of the empirical density function which has a higher excess and heavier tails than density function of normal distribution. Furthermore the shape of density function of the empirical data can be asymmetric, i.e. it has different left and right tails.

In other words, the empirical data can not be approximated with normal distribution, and there are specific types of processes to be considered for aforementioned reasons. Hyperbolic distributions, ARCH-type processes and Levy's processes are considered in the represented paper.

2.2 ARCH-type models

In the course of numerous researches it was found out that the conditional variance of the financial time series is not constant through the time. This property turned out to bevery important for the construction of risk estimation procedures, and consequently caused huge interest among both researchers and experts. As a result plenty of volatility estimating models were created on the basis of ARCH model introduced by Engle (1982).

ARCH

Let y_t be a stochastic variable which is defined in some information set $\mathcal{F}_{t-1}$. Formally, $\mathcal{F}_{t-1}$ is the sigma-algebra, generated by the set of all variables which appeared by the $t-1$ moment of time. Let us define y_t as the logarithm of the successive increments of some financial variable (the stock price, bond yield, volume weighed average price etc.) ratio

$$y_t = \ln \frac{p_t}{p_{t-1}}.$$

Actually the problem is reduced to modelling the conditional density function for y_t, which can be denoted as

$$f(y|\mathcal{F}_{t-1}) \equiv \frac{d}{dy} P(y_t \leqslant y|\mathcal{F}_{t-1}),$$

where $P(y_t \leqslant y|\mathcal{F}_{t-1})$ is the conditional probability. It is necessary to note that functions of the conditional mean and conditional variance are parametrized with a finite dimensional vector $\theta \in \Theta \subseteq \mathbb{R}^n$, θ_0 is a vector of intrinsic values. The most important thing in ARCH-type models is the conditional variance, while the conditional mean just of secondary importance. It can be modelled in a rather simple way:

$$\mu_t = \mu_0 + \mu_1 \zeta(\sigma_{t-1}),$$

where $\zeta(x) = x^2$. As a rule, a conditional mean can be represented as a constant or just a zero, otherwise there occurs superfluous complexities at modelling which do not result in essential improvement of the result.

Let us define that process $\{\varepsilon_t(\theta)\}$ corresponds to ARCH model if the conditional mean of it is equal to zero,

$$\mathrm{E}_{t-1}(\varepsilon_t(\theta)) = \mathrm{E}[\varepsilon_t(\theta)|\mathcal{F}_{t-1}] = 0, \quad t = 1, 2.,$$

and a conditional variance

$$\sigma_t^2(\theta_0) = \mathrm{Var}_{t-1}(\varepsilon_t(\theta_0)) = E_{t-1}(\varepsilon_t^2(\theta)), \quad t = 1, 2$$

non-trivially depends on the previous values $\{\varepsilon_{t-1}(\theta),\ \varepsilon_{t-2}(\theta), \ldots\}$.

Let us denote the conditional mean as $\mu_t(\theta_0) = E_{t-1}(y_t), \quad t = 1, 2, \ldots$ The process $\{\varepsilon_t(\theta)\}$ will be defined as follows

$$\varepsilon_t(\theta_0) = y_t - \mu_t(\theta_0) \quad t = 1, 2, \ldots$$

The conditional variance of $\{\varepsilon_t\}$ equals to the conditional variance of $\{y_t\}$, and its conditional expectation is equal to zero. It is necessary to resort to such substitution, because the conditional mean of real time series used in economy is equal to zero only in very rare cases.

Let us define

$$z_t(\theta_0) = \frac{\varepsilon_t(\theta_0)}{\sqrt{\sigma_t^2(\theta_0)}},$$

it is easy to see that its conditional mean of $z_t(\theta_0))$ is equal to zero, and a conditional variance — to unity. This normalization allows us to compare our process with some standard distributions, for example, standard normal distribution.

According to linear ARCH(q) model the conditional variance should linearly depend on last q available squared values.

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2$$

For the model to be well defined and the conditional variance to be positive it is necessary for the parameters to satisfy the following conditions

$$\alpha_0 > 0, \quad \alpha_1, \ldots, \alpha_q \geqslant 0$$

If define $\nu_t = \varepsilon_t^2 - \sigma_t^2$, the model can be rewritten in the following way:

$$\varepsilon = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \nu_t^2.$$

Since $E_{t-1} = 0$ the model directly corresponds to AR(q) model for squared increments ε_t^2. Such processes have a property of covariant stability in the only case when the sum of positive parameters of autoregression is less than the unity. In this case the variance of the process is equal to

$$\mathrm{Var}(\varepsilon_t) = \sigma^2 = \frac{\alpha_0}{1 - (\alpha_1 + ... + \alpha_q)}$$

Even in spite of the fact that ε_t are serially uncorrelated, i.e. the coefficient of autocorrelation is close to zero, there is no obvious independence of process through the time. According to some features of successive price increment logarithms distribution ratio there is a small evidence to expect that the big (small) changes can be followed by big (small) changes of an unpredictable sign. For the same reasons it is possible to say that if the distribution of the normalized process z_t is a stable, then unconditional distribution of ε_t has heavier tails than the distribution for z_t.

GARCH

Plenty of parameters and a large lag are often required for the experimental implementation of ARCH(q). To circumvent this problem Bollerslev offered GARCH(p, q) (Generalized ARCH) model,

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2 = \alpha_0 + \alpha(L)\varepsilon_{t-1}^2 + \beta(L)\sigma_{t-1}^2,$$

where L denotes the lag or backshift operator, $L^i y_t = y_{t-i}$.

For the conditional variance in the GARCH(p, q) model to be well defined all the coefficients in the corresponding infinite order linear ARCH model must be positive. Provided that $\alpha(L)$ and $\beta(L)$ have no common roots and the roots of polynomial $\beta(x) = 1$ lie outside the unity circle, this positivity constraint is satisfied if and only if all the coefficients in the infinite power series expansion for $\alpha(x)/(1 - \beta(x))$ are non-negative. For the simplest GARCH(1,1) model almost sure the positivity of σ_t^2 requires that $\alpha_0 > 0$, $\alpha_1 \geqslant 0$ and $\beta_1 \geqslant 0$.

GARCH(p, q) can be rewritten as follows:

$$\varepsilon_t^2 = \alpha_0 + (\alpha(L) + \beta(L))\varepsilon_{t-1}^2 - \beta(L)\nu_{t-1} + \nu_t,$$

that defines ARMA($\max(p, q)$) model for ε_t^2. It follows from the aforementioned reasons that the model is covariantly stable, if and only if all the roots of $\alpha(x)+\beta(x) = 1$ lie outside the unity circle. In many applications with high frequency financial data the estimation for $\alpha(1) + \beta(1)$ turn out to be very close to unity.

Other ARCH type models

The event described above provides an empirical motivation for the so-called Integrated GARCH(p,q) or IGARCH(p,q) model introduced by Engle and Bollerslev (1986) which can be written with the following way:

$$\sigma_t^2 = \alpha_0 + \varepsilon_{t-1}^2 + \sum_{i=2}^{p} \alpha_i(\varepsilon_{t-i}^2 - \varepsilon_{t-1}^2) + \sum_{j=1}^{q} \beta_j(\sigma_{t-j}^2 - \varepsilon_{t-1}^2)$$

In the IGARCH class of models the autoregressive polynomial has a single root which means that the impact on the conditional variance is persistent in the sense that it remains important for future forecasts of all horizons.

Empirical researches have shown, that GARCH successfully captures thick tails of the successive data increments distribution and volatility clustering, apart from this it is easy to modify it in such a way that it could work with another features of the market such as non-trading period and forecastable events. However it is not well suited to capture the "leverage effect", since the conditional variance is a function depending only on the magnitudes of the lagged residuals and not on their signs.

Nelson (1991) has offered EGARCH (Exponential GARCH) model, which depends both on magnitudes and sings of lagged residuals:

$$\ln(\sigma_t^2) = \alpha_0 + \sum_{i=1}^{p}[\alpha_i\varepsilon_{t-i} + \gamma_i(|\varepsilon_{t-i}| - E|\varepsilon_{t-i}|)] + \sum_{j=1}^{q} \beta_j \ln \sigma_{t-j}^2$$

Thus the $\ln \sigma_t^2$ corresponds to ARMA(p,q) process with already considered stability conditions. As well as in the GARCH case it is possible to make α_0 function of time. By this manner the model can be adapted to effects of non-trading period and forecastable events. Apart from the above mentioned one,there are plenty of another ways for volatility evaluation. For example, to satisfy the property of volatility clustering GARCH assumes that conditional variance equals to distributed lag of squared residuals. An equally natural assumption is that the conditional variance is the distributed lag of the absolute residuals:

$$\sigma_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i|\varepsilon_{t-i}| + \sum_{j=1}^{q} \beta_i\sigma_{t-j.}$$

this assumption was brought forward by Taylor and Schwert. Higgins and Bera have combined GARCH and the model just considered, and put them to the NGARCH (Non-linear GARCH) class of models,

$$\sigma_t^\delta = \alpha_0 + \sum_{i=1}^{p} \alpha_i|\varepsilon_{t-i}|^\delta + \sum_{j=1}^{q} \beta_j\sigma_{t-j}^\delta$$

However this model does not capture "leverage effect", therefore A-PARCH modification was proposed

$$\sigma_t^\delta = \alpha_0 + \sum_{i=1}^{p} \alpha_i [|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i}]^\delta + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^\delta$$

When formulating the equation with $\gamma = 2$ one obtains a special case of QARCH model, in which volatility is modelled as a quadratic form in the lagged residuals. The simplest version of this model was named AARCH (Asymmetric ARCH),and in the first order case it becomes:

$$\sigma_t^2 = \alpha_0 + \alpha \varepsilon_{t-1}^2 + \delta \varepsilon_{t-1} + \beta \sigma_{t-1}^2,$$

where the negative value of δ means that positive returns increase volatility less, than the negative ones.

Another way to take into account the effect of asymmetry is:

$$\sigma_y^\gamma = \alpha_0 + \sum_{i=1}^{p} [\alpha_i^+ I_{\{\varepsilon_{t-i} > 0\}} |\varepsilon_{t-i}|^\gamma + \alpha_i^- I_{\{\varepsilon_{t-i} \leqslant 0\}} |\varepsilon_{t-i}|^\gamma] + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2,$$

where $I_{\{\cdot\}}$ denotes the indicator function. A special case of this model is TARCH (Threshold ARCH) which corresponds to previous equation with $\gamma = 1$:

$$\sigma_y = \alpha_0 + \sum_{i=1}^{p} \alpha_i [(1 - \gamma_i)\varepsilon_{t-i}^+ - (1 + \gamma_i)\varepsilon_{t-i}^-] + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2,$$

another special case with $\gamma = 2$ is GJR model.

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} [\alpha_i + \gamma_i I_{\{\varepsilon_{t-i}^2 > 0\}}] \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$$

This model allows a quadratic response of volatility to news using different coefficients for good and bad news, but maintains the assertion that the minimum volatility will result when there is no news.

There exists plenty of ARCH type models modifications, some of which are represented below:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} [\alpha_i \varepsilon_{t-i}^2 + \gamma_i \varepsilon_{t-i}] + \sum_{i=1}^{q} \beta_j \sigma_{t-j}^2 \qquad \text{A-GARCH}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i (\varepsilon_{t-i} + \gamma_i \sigma_{t-i})^2 + \sum_{i=1}^{q} \beta_j \sigma_{t-j}^2 \qquad \text{NA-GARCH}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i(\varepsilon_{t-i} + \gamma_i)^2 + \sum_{i=1}^{q} \beta_j \sigma_{t-j}^2 \qquad \text{V-GARCH}$$

$$\ln \sigma_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i |\varepsilon_{t-i}| + \sum_{i=1}^{q} \beta_j \ln \sigma_{t-j}^2 \qquad \text{log-GARCH}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i} + \sum_{i=1}^{p} \alpha_i i \varepsilon_{t-i}^2 + \sum_{i<j}^{p} \alpha_i j \varepsilon_{t-i} \varepsilon_{t-j} + \sum_{i=1}^{q} \beta_j \sigma_{t-j}^2 \qquad \text{GQ-GARCH}$$

2.3 GARCH(1,1)

The simplest model of the class of GARCH models is GARCH(1,1) which is represented with the formula

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

This model is especially appealling because of its simplicity, as it contains only three parameters. There exist many ways to estimate the parameters of the model, such as the maximum likelihood method,the method of quasi-maximum likelihood etc.

The researches carried out during a long period of time have shown that GARCH(1,1) does not so well describe the big time series, but it can be used on small lags. Moreover, the researchers found out that this model estimates volatility with high accuracy, and its precision is just in some cases lower than the one of the competing models. Just due to its simplicity for practical implementation and high accuracy GARCH(1,1) is widely adopted among the experts.

When using GARCH(1,1) for a long period of time it is necessary to update parameters of the model. In this relation, it is necessary to resolve the problem of how to calculate the moment of time at which the parameters should be revalued. It has to be noted that the distribution of GARCH processes have Pareto-like tails, i.e.

$$P(X > x) \sim c_0 x^{-k} \text{ under } x \to \infty \text{ for any } c_0,\ k > 0$$

Experience shows that for long lags $\phi_1 = \alpha_1 + \beta_1$ is usually close to 1, for example Bollerslev and Mikkelsen applied GARCH (1,1) to S&P500, and using daily data from 1958 to 1990, have received value $\phi_1 = 0,995$, when used on the data with higher frequency, ϕ take on a lower value.

2.4 Maximum Likelihood method

The parameters of GARCH(1,1) model is vector $\theta_0 = (\alpha_0,\, \alpha_1,\, \beta_1)$. One of the most often used methods in estimating θ_0 for ARCH models is based on maximization of a likelihood function constructed under the auxiliary assumptions of an i.i.d. distribution for the standardized innovations.

Let $\{y_T,\, y_{T-1}, \ldots, y_1\}$ denote a set of empirical data which satisfy the conditions of ARCH model. The log likelihood function for t^{th} observation is given by

$$l_t(y_t, \theta) = \ln(f(z_t(\theta))) - \frac{1}{2}\ln\sigma_t^2(\theta),$$

where $f(z_t(\theta))$ is a density function for $z_t(\theta)$ distribution. The second term on the right hand side is a Jacobian that arises in the transformation from the standardized innovations $z_t(\theta)$ to the observable data $y_t(\theta)$. By a standard way, the log likelihood function for the sample should be denoted as a sum of the conditional log likelihood functions as follows

$$L_T(y_T, y_{T-1}, \ldots, y_1) = \sum_{t=1}^{T} l_t(y_t)$$

The maximal likelihood estimator for the true parameters θ_0 is found by maximization of this function.

The actual implementation of the maximum likelihood procedure requires an explicit definition for the conditional density $f(z_t(\theta))$. The distribution most commonly employed in literature is a standard normal one which looks as

$$f(z_t(\theta)) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z_t^2}{2}}$$

consequently

$$l_t(y_t) = \ln\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{z_t^2}{2}}\right) - \frac{1}{2}\ln\sigma_t^2 = \ln\left(\frac{1}{\sqrt{2\pi\sigma_t^2}} e^{-\frac{(y_t-\mu_t)^2}{2\sigma_t^2}}\right) =$$

$$= -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln\sigma_t^2 - \frac{(y_t-\mu_t)^2}{2\sigma_t^2}$$

Following GARCH(1,1) model conditional variance equals

$$\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2,$$

due to constancy of the parameters it can be rewritten as follows

$$\sigma_t^2 = \alpha_0(1-\beta_1) + \alpha_1\varepsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2,$$

where each σ_t^2 has the same representation, and $\sigma_1^2 = \alpha_0$. Taking into consideration the sum of geometric series let us write down the next formula for σ_t^2

$$\sigma_t^2 = \alpha_0 + \alpha_1 \sum_{i=0}^{t-2} \beta_1^i \varepsilon_{t-i-1}^2$$

Consequently expression for the likelihood function is

$$L_T(\{y_t\}) = \sum_{t=1}^{T} l_t(y_t) = -\frac{1}{2} T \ln 2\pi - \frac{1}{2} \ln \alpha_0 - \frac{(y_1 - \mu_1)^2}{2\alpha_0} - \frac{1}{2} \sum_{t=2}^{T} \ln \sigma_t^2 - \sum_{t=2}^{T} \frac{(y_t - \mu_t)^2}{2\sigma_t^2}$$

2.5 Hyperbolic distributions

As it was mentioned above, empirical density can not always be adequately described with the help of a normal distribution. Let us consider two assumptions related to this fact: (1) there exists statistically stable nongaussian distribution of empirical data; (2) there is no any statistically stable distributions, in this case it is implied that actual data can be approximated with some mixtures of distributions (may be Gaussian), with different parameters. As a result of such representation of the empirical data summary distribution is something more complex then distributions of summands.

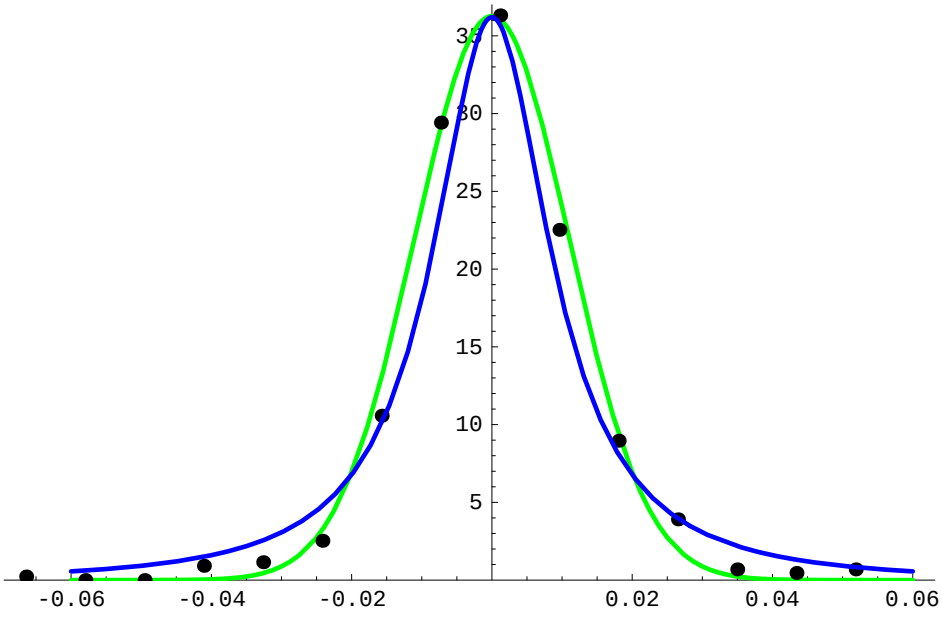


Figure 2: *The comparative diagram of the empirical data, normal and gaussian distributions. Empirical data are represented with points, normal distribution is a light green curve and a gaussian distribution is represented with dark blue color.*

The class of *hyperbolic* distributions was considered for the approximation of empirical data.

Implementation of these distributions in financial mathematics was offered by Eberlein E. and Keller U. (see [6], [7] and [8] for example).

In the simplest case (actually *hyperbolic* distribution) density function is represented as follows:

$$h(x;\alpha,\beta,\delta,\mu) = C(\alpha,\beta,\delta)\, e^{-\alpha\sqrt{\delta^2+(x-\mu)^2}+\beta(x-\mu)}. \tag{1}$$

$C(\alpha,\beta,\delta)$ is the normalizing constant which has the following form

$$C(\alpha,\beta,\delta) = \frac{\sqrt{\alpha^2-\beta^2}}{2\alpha\delta K_1(\delta\sqrt{\alpha^2-\beta^2})} \tag{2}$$

where $K_1(x)$ is a modified Bessel function of the third order with index 1.

The origin of the term "hyperbolic" is related to the following circumstance. Logarithm $\ln\phi(x)$ of the normal density

$$\phi(X) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-m)^2}{2\sigma^2}}$$

is a parabola. In the same way logarithm $\ln h(x)$ of the density function of the "hyperbolic" distribution is a hyperbola. Indeed, the equation

$$y = -\alpha\sqrt{\delta^2+(x-\mu)^2}+\beta(x-\mu) \tag{3}$$

with asymptotes

$$-\frac{\alpha}{2}|x-\mu|+\beta(x-\mu) = 0 \tag{4}$$

after the substitution $z = y-\beta(x-\mu)$ and squaring it will take the form of a hyperbola equation.

It is supposed that parameters $(\alpha,\beta,\delta,\mu)$ in (1) satisfy the following conditions

$$\alpha > 0, \qquad 0 \leqslant |\beta| < \alpha, \qquad \mu \in \mathbb{R}, \qquad \delta \geqslant 0.$$

The condition $|\beta| < \alpha$ cause a decrease of the density function (1) under $x \to \pm\infty$, which is a true property of probabilistic density. Density function decrease has the same speed as an exponential function $\exp\{-(\alpha\pm|\beta|)x\}$. The normal density has a decrease speed $\exp\{-x^2/2\sigma^2\}$, i.e. it is much faster than the exponential function. This property of the gaussian density function give an evidence to believe that it can approximate empirical data with their "heavy" tails better the the normal one.

Each parameter in formula (1) has its own meaning, μ is just the shift parameter (x is replaced with $x-\mu$), δ is the scale parameter. It should be mentioned that the hyperbolic distribution is invariant under shift and scale changes. Parameters α and β are responsible

for the shape of function. β is related to skewness of the distribution function with respect to $x = \mu$ point, if $\beta > 0$, the right tail decreases slower than the left one (vice versa,if $\beta < 0$).

It is offered to apply distribution (1) for the description of increments process $\Delta \ln S_t = \ln S_{t+\Delta} - \ln S_t$. The set of examples of empirical data well fitted with hyperbolic distribution is represented in papers [6]-[8].

Let us consider another representative of the class of generalized hyperbolic distributions, it is so-called "Gaussian—Inverse Gaussian" distribution (see [9]).

The density of GIG-distribution is given

$$g(x) = C_1(\alpha, \beta, \delta, \mu) \left(q\left(\frac{x-\mu}{\delta}\right)\right)^{-1} K_1\left(\alpha\delta q\left(\frac{x-\mu}{\delta}\right)\right) e^{\beta(x-\mu)}, \tag{5}$$

where

$$C_1(\alpha, \beta, \delta) = \frac{\alpha}{\pi}\, e^{\delta\sqrt{\alpha^2+\beta^2}},$$
$$q(x) = \sqrt{1+x^2}.$$

It should be noted, that as

$$K_1(x) \sim \sqrt{\frac{\pi}{2}}\, x^{-1/2}\, e^{-x}, \qquad x \to \infty,$$

then under $|x| \to \infty$

$$g(x) \sim \left(\frac{\alpha}{2\pi\delta}\right)^{1/2} \cdot \frac{1}{1+\left(\frac{x-\mu}{\delta}\right)^2} \exp\left\{-\alpha\sqrt{\delta^2+(x-\mu)^2} + \beta(x-\mu)\right\} \tag{6}$$

and, concequently,

$$\ln \frac{h(x)}{g(x)} \ln\left(1+\left(\frac{x-\mu}{\delta}\right)^2\right), \qquad |x| \to \infty. \tag{7}$$

The checking of how the hyperbolic distribution fits the empirical data was undertaken in papers [8] and [10]. In paper [9]it was marked that GIG-distribution is well fitted to the sample received from hyperbolic distribution, and hence it can also approximate empirical data y_t well.

It is possible to use the technique offered in paper [11] to estimate the hyperbolic distribution parameters.

Hyperbolic distribution is significantly simpler than "Gaussian—Inverse Gaussian" distribution (to ascertain this just compare densities $h(x)$ (1) and $g(x)$ (5)). However, there is one basic circumstance giving preference to the second distribution.

Let the random variable Y inhere of GIG-distribution with $g(X) = g(x; \alpha, \beta, \delta, \mu)$ density. Accordingly, generating function of the moments

$$\mathsf{E}\, e^{\lambda Y} = \exp\left\{\delta\left[\sqrt{\alpha^2-\beta^2} - \sqrt{\alpha^2-(\beta+\lambda)^2}\right] + \mu\lambda\right\}. \tag{8}$$

It is obvious from the above mentioned that if $(Y_i)_{i=1,\dots,m}$ independent variables related to GIG-distribution with the same α and β, but generally speaking different μ_i and δ_i, then the sum $Y_+ = Y_1 + \cdots + Y_m$ is again GIG-distributed value with the same α and β and

$$\mu_+ = \mu_1 + \cdots + \mu_m, \qquad \delta_+ = \delta_1 + \cdots + \delta_{m.}$$

In other words, GIG-distribution is closed (in the specified sense) concerning convolution.

If X has a hyperbolic distribution, then having put $\beta = \mu = 0$ for simplicity we find, that

$$\mathsf{E}\, e^{\lambda X} = \frac{\alpha}{K_1(\alpha\delta)} \frac{K_1(\delta\sqrt{\alpha^2 - \lambda^2})}{\sqrt{\alpha^2 - \lambda^2}} \tag{9}$$

it is obvious that the property of closure which GIG-distributions have, does not exist for a hyperbolic distribution.

It should be noted in the conclusion that both GIG and hyperbolic distributions are infinitely divisible.

Thus, appealling properties of hyperbolic and GIG-distributions and their good compliance with empirical data gives a proof to use them at more detailed description of the processes $H = (H_t)_{t\geqslant 0}$, incluided into the definition of stock prices $S_t = S_0\, e^{H_t}$.

In the conclusion we will note that the implementation of the hyperbolic distributions class with the purpose of VAR estimation which will be considered later, represents interesting and almost unexplored issue. And as hyperbolic distributions are better fitted to the empirical data than normal distributions it opens wide prospects for solving VAR estimation problem.

2.6 Levy processes

The random process $X = (X_t)_{t\geq 0}$, given on probabilistic space (Ω, F, P) and attaining values in d-dimensional Euclidean space $\mathbb{R}^d$, is called a d-dimensional *Levy process*, if the following conditions are met:

1. $X_0 = 0$ (P-a.s.);

2. for any $n > 0$ and the set $0 \leq t_0 < t_1 < \ldots < t_n$, the values $X_{t_0}, X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent (property of "independence of increments");

3. for each $s \geq 0$ and $t \geq 0$ it is true that $X_{t+s} - X_s \stackrel{d}{=} X_t - X_0$ (property of "stability", "uniformity" of increments);

4. for each $t \geq 0$ and $\varepsilon \geq 0$ $\lim\limits_{s\to t} P(|X_s - X_t| > \varepsilon) = 0$ (property of "stochastic continuity");

5. trajectories $(X_t(\omega))_{t\geq 0}$ on $\omega \in \Omega$ -a.s. belong to $\mathbb{D}^d$ space, consisting of vector functions $f = (f_t)_{t\geq 0}, f_t(f_t^1, f_t^2, \ldots, f_t^d$, each component of which $f^i = (f_t^i)_{t\geq 0}, i = 1, \ldots, d$ is continuous on the right and has limits on the left under $t > 0$.

Levy processes $X = (X_t)_{t\geq 0}$ are the processes with homogeneous increments, and consequently their distributions are completely determined by one-dimensional distribution $P_t(dx) = P(X_t \in dx)$. From the very definition of Levy processes it follows that distribution $P_t(dx)$ is infinitely divisible for each t.

Let $\phi_t(\theta) = Ee^{i(\theta,X_t)} = \int\limits_{\mathbb{R}^d} e^{i(\theta,x)} P_t(dx)$ be a characteristic function. Then, using the formula of Levy-Khinchina

$$\phi_t(\theta) = exp\{i(\theta, B_t) - \frac{*1}{*2}(\theta, C_t\theta) + \int\limits_{\mathbb{R}^d} (e^{i(\theta,x)} - 1 - i(\theta, x)I(|x| \leq 1))\nu_t(dx)\},$$

where $B_t \in \mathbb{R}^d, C_t$ is a symmetric non-negatively definite matrix of the $d \times d$ order, and $\nu_t(dx)$ is a Levy measure (for each t) with the property of $\int\limits_{\mathbb{R}^d} (|x|^2 \wedge 1)\nu(dx) < \infty$.

Because of the uniformity and independence of increments

$$\phi_{t+s}(\theta) = \phi_t(\theta)\phi_s(\theta),$$

consequently

$$\phi_t(\theta) = exp\{t\psi(\theta)\}.$$

Since triplets (B_t, C_t, ν) are unambiguously defined through the characteristic function, they can be represented as

$$B_t = t \cdot B, C_t = t \cdot C, \nu_t(dx) = t \cdot \nu(dx),$$

where $B = B_1, C = C_1, \nu = \nu_1$.

Hence

$$\psi(\theta) = i(\theta, B) - \frac{*1}{*2}(\theta, C\theta) + \int\limits_{\mathbb{R}^d} (e^{i(\theta,x)} - 1 - i(\theta, x)I(|x| \leq 1))\nu(dx.)$$

The sense of the triplet components, figuratively speaking, is as follows:

1. The first component is a "trend" component responsible for the average movement of the process.

2. The second component determines the dispersion of a gaussian continuous process component.

3. The third component is responsible for the behaviour of spasmodic process component. It shows frequency and value of a jump.

A classical example of a continuous Levy process is a standard brownian motion (with $B_t = 0, C_t = t, \nu_t \equiv 0$).

A classical example of the process having a finite Levy measure case is Poisson's process with parameter $\lambda > 0$. In this case $B_t = \lambda \cdot t, C_t = 0, \nu(dx) = \lambda_{\{1\}}(dx)$. A characteristic function has the following form

$$\phi_t(\theta) = exp\{\lambda t(e^{i\theta)} - 1)\}.$$

It is possible to get a wide class of purely spasmodic Levy processes basing on Poisson's process. Let $N = (N_t)_{t\geq 0}$ be Poisson's process with parameter $\lambda > 0$, $\xi = (\xi_j)_{j\geq 1}$ is a sequence of independent identically distributed random variables (independent from N), the distribution of which is

$$P(\xi_i \in A) = \frac{\nu(A)}{\lambda}, A \in \mathbb{B}(\mathbb{R})$$

where $\lambda = \nu(\mathbb{R}) < \infty, \nu(\{0\}) = 0$. Let's form the process $X_t = \sum\limits_{j=1}^{N_t} \xi_t, t > 0$. Direct calculation shows that

$$\phi_t(\theta) = e^{t\int(e^{i\theta x}-1)\nu(dx)}.$$

This process is called compound Poisson's process and it is also a Levy process.

The simplest example of a Levy process with $\nu(\mathbb{R}) = \infty$ measure can be derived as follows. Let $\lambda = (\lambda_k)_{k\geq 1}$ be a sequence of positive numbers, $\beta = (\beta_k)_{k\geq 1}$ be a sequence of such numbers that $\sum\limits_{k=1}^{\infty} \lambda_k \beta_k^2 < \infty$. Let $\nu(dx) = \sum\limits_{k=1}^{\infty} \lambda_k I_{\{\beta_k\}}(dx)$, then denote $N^{(k)} = (N_t^{(k)})_{t\geq 0}, k \geq 1$ as a sequence of independent Poisson's processes with parameters $\lambda_k, k \geq 1$, accordingly.

Assuming $X_t^{(n)} = \sum\limits_{k=1}^{n} \beta_k(N_t^{(k)} - \lambda_k \cdot t)$ it is easy to see that for each $n \geq 1$ process $X^{(n)} = (X_t^{(n)})_{t\geq 0}$ is a Levy process with Levy measure $\nu^{(n)}(dx) = \sum\limits_{k=1}^{n} \lambda_k I_{\{\beta_k\}}(dx)$ and

$$\phi_t^{(n)} = Ee^{i\theta X_i^{(n)}} = e^{t\int(e^{i\theta x}-1-i\theta x)\nu^n(dx)}$$

Limit process is also a Levy process.

In connection with "explicit" representations of some (spasmodic) Levy processes, a way for their modelling can be obtained. This method is based on modelling random variables ξ_j, β_j only and on modelling exponentially distributed values which determine the intervals between the jumps of Poisson's process.

The second way of modelling is described in O.E. Brandorff-Nielsen "Probability densities and Levy densities" article. Non-decreasing positive random variables are considered in this paper. In this case it turns out that the density function of the process $X = (X_t)_{t\geq 0}$ distribution has the following form

$$f(x,t) = \sum_{n=1}^{\infty} \frac{t^n}{n!} u_n(x),$$

where $u_n = \lim\limits_{s\downarrow t} u_{n\varepsilon}(x), u_{n\varepsilon}(x) = \sum\limits_{k=1}^{n} (-1)^{n-k} C_k^n c(\varepsilon)^{n-k} u_\varepsilon^{*k}(x)$, $u_\varepsilon^{*n}(x)$ is an n-dimensional convolution and $u_\varepsilon(X)1_{(s,\infty)}u(x)$.

From such representation of distribution density function it directly follows that $\lim\limits_{t\downarrow 0} t^{-1} f(x,t) = u(x)$. This limit transformation allows to model distribution function with the Levy measure for small lags. The speed of convergence of this transformation can not be evaluated theoretically, and it should be calculated for each separate type of processes.

2.7 Analysis of the empirical data distribution

All the researches cover the period of time since October, 2012 till February, 2013. Daily data could not be used for the statistical analysis due to the insufficient size of the sample. For this reason in the given paper were used hourly intervals intraday for a set of stocks included in Nasdaq100 index.

For the following research natural logarithms of the successive price changes were used

$$S_1 = \ln \frac{V_2}{V_1}, S_2 = \ln \frac{V_3}{V_2}, \ldots, S_{n-1} = \ln \frac{V_n}{V_{n-1}},$$

where V_i is a VWAP of i-th hour.

2.7.1 Graphical methods of normality testing

Frequency distributions

One of the simplest ways to analyse the distribution of price logarithms changes is to construct frequency distributions for each separate stock. For each stock there can be constructed quantitative ratios of number of hits of price logarithms changes during some interval and the total number of the changes considered. The size of the interval can be determined, for example, by the numbers representing a product of standard deviation by some coefficients. The coefficients, for example, can be 0.5; 1; 2; 3; 4; 5. Then, the obtained results can be compared with results obtained under the assumption that the distribution of price logarithms changes is normal (see Appendix 1, fig.6).

Normal probability paper

Other graphical tool for estimating the deviation from normality is plotting on the normal probability paper. If u is a random gaussian variable with mean μ and variance σ^2, then normalized variable

$$z = \frac{u - \mu}{\sigma}$$

will refer to the normal distribution with the mean equal to zero and the variance equal to unity, and the diagram of $z(u)$ dependence will be a straight line.

The relation between z and u can be used for estimating The deviation from normality for u distribution. Let us consider the sorted sample of values $u_i,\ i = \overline{1, N}$, where each u_i is an estimation for f fractile of u distribution. The value of the fractile is specified as:

$$f = \frac{3i - 1}{3N + 1}$$

The exact z value for fractile f of normal distribution can be easily calculated with the help of the computer. If u distribution is normal, then dependence $u(z)$ has to be a straight line. Some deviations from linearity are possible due to the errors of the sample. If the observable deviation is very big it is possible to conclude that u distribution is not normal.

It should be noted that plotting these graphs gives too rough an estimation for normality and besides they are too subjective (see appendix 1, figure 5). For this reason there appears a necessity of using more strict statistical criterions, two of which are described below.

2.7.2 Kolmogorov-Smirnov's criterion of consent

As it has already been mentioned, the size of the sample used is not very large. For this reason Kolmogorov-Smirnov's criterion of consent is implemented for checking the empirical data normality.

The parameters of a theoretical distribution which the empirical data are compared to have to be known. In the given case a normal distribution is used.

Let us consider the realization of this criterion:

1. Formulation of null and alternative hypotheses

$$H_0 : \hat{F}(x) = F_{mod}(x, \vec{\theta})$$

$$H_1 : \hat{F}(x) \neq F_{mod}(x, \vec{\theta}),$$

where $\hat{F}(x)$ is an empirical distribution, $F_{mod}(x, \vec{\theta})$ is the given theoretical distribution with the set of parameters $\vec{\theta}$.

2. Assignment of a confidence level α.

3. Formulation of critical statistics.

   In the criterion considered are implemented the following types of statistics to evaluate the measure of deviation between empirical and modelling distributions:

$$D_n = \max|\hat{F}(x) - F_{mod}(x,\vec{\theta})|$$
$$D_n^+ = \max(\hat{F}(x) - F_{mod}(x,\vec{\theta}))$$
$$D_n^- = \max(F_{mod}(x,\vec{\theta}) - \hat{F}(x))$$

   Statistics $\sqrt{n}D_n$ and $\sqrt{n}D_n^-$ are called Kolmogorov and Smirnov's statistics accordingly, where

$$D_n = \max|D_n^+, D_n^-|.$$

   Exact distributions for all the tree types of statistics are known, though D_n statistics is usually enough for being used in practical implementation. For these reasons let us put the criterion distribution equal to

$$\psi_{cr} = \sqrt{n}D_n = \sqrt{n}\max|\hat{F}(x) - F_{mod}(x,\vec{\theta})|$$

   A.N.Kolmogorov has shown, that if function $F_{mod}(x,\vec{\theta})$ is continuous, then distribution ψ_{cr} has function

$$K(\psi_{cr}) = \sum_{i=-\infty}^{\infty}(-1)^i \cdot e^{-2i^2\cdot\psi_{cr}^2}$$

   as its limit.

   This function was called Kolmogorov's function and it does not depend on $F_{mod}(x,\vec{\theta})$ function shape.

   It is possible to speak about a consistency of Kolmogorov-Smirnov's criterion implementation in the case, when the modelling distribution has the parameters of shift and scale only.

4. From the definition of the distribution function it follows that under any $\psi_{cr} > 0$ and large enough n the probability of the event that $\sqrt{n}D_n$ takes ψ_{cr} value, is represented as

$$P\sqrt{n}D_n \geqslant \psi_{cr} = 1 - K(\psi_{cr}) = 1 - \sum_{i=-\infty}^{\infty}(-1)^i \cdot e^{-2i^2\cdot\psi_{cr}^2} = \alpha.$$

The value of $\psi_{cr.e.}$ can be calculated with the help of the Kolmogorov-Smirnov's function values table. The lower critical border is not used in this criterion.

5. ψ_{ev} can be determined from the expression of the criterion by substituting n and D_n values for the empirical data used. If the following condition is satisfied

$$\psi_{ev} < \psi_{cr.e},$$

the hypothesis about the consent of empirical and modelling distributions is accepted.

Kolmogorov-Smirnov's criterion of consent can be used not only with small samples but with big ones too. For this purpose it is necessary to represent the sample in aggregated form and thevalues of $\hat{F}(x)$ and $F_{mod}(x, \vec{\theta})$ should be calculated on the borders of selected intervals of aggregation.

2.7.3 Pirson's chi-square criterion of consent

The chi-square criterion of consent allows to carry out checking the hypothesis about the consent when the parameters of the model are unknown.

Unknown parameters of the model can be replaced by their estimations obtained from a sample. The method of moments or the maximum likelihood method can be used to estimate the model's parameters.

The chi-square criterion of consent is applicable if the sample size is $n \geqslant 200$ and it demands grouping of thesample. The number of grouping intervals should satisfy $L \geqslant 8$ condition, and the amount of hits in each interval μ_j should not be less than 7-10. Otherwise the neighboring intervals should be concatenated with an appropriate correction of L.

Let us consider the realization of the criterion:

1. Formulation of null and alternative hypotheses

$$H_0 : \hat{F}(x) = F_{mod}(x, \vec{\theta})$$

$$H_1 : \hat{F}(x) \neq F_{mod}(x, \vec{\theta}),$$

where $\hat{F}(x)$ is an empirical distribution, $F_{mod}(x, \vec{\theta})$ is theoretical distribution with the set of parameters $\vec{\theta}$

2. Assignment of confidence level α.

3. Formulation of critical statistics.

$$\psi_{cr} = \sum_{j=1}^{L} \frac{(\mu_j - n \cdot p_j)^2}{n \cdot p_j},$$

where $\mu_j,\ j = \overline{1, L}$ is the amount of hits in each interval of grouping j, p_j is the theoretical probability of hitting interval j.

$$p_j = F_{mod}(x_{j+1}; \hat{\vec{\theta}}) - F_{mod}(x_j; \hat{\vec{\theta}})$$

Here $x_{j+1}$and x_j are the upper and lower borders of the current interval of grouping accordingly.

The limit distribution of statistics ψ_{cr} under $n \to \infty$ has the following form

$$\lim_{n\to\infty} \sum_{j=1}^{L} \frac{(\mu_j - n \cdot p_j)^2}{n \cdot p_j} = \chi^2(\psi_{cr}; L - S - 1),$$

where S is the amount of modelling distribution parameters, the consent with which should be checked. $\chi^2(\psi_{cr}; L-S-1)$ is a function of chi-square distribution with $(L-S-1)$ degrees of freedom.

4. Definition of upper and lower critical points with the help of the percentage points table of χ^2 distribution:

$$\psi_{cr.u} = \chi^2_{\alpha/2\cdot 100\%}(L - S - 1)$$

$$\psi_{cr.l} = \chi^2_{(1-\alpha/2)\cdot 100\%}(L - S - 1)$$

5. Evaluation of critical statistics

$$\psi_{ev} = \sum_{j=1}^{L} \frac{(\mu_j - n \cdot p_j)^2}{n \cdot p_j},$$

If the following condition is satisfied

$$\chi^2_{\alpha/2\cdot 100\%}(L - S - 1) < \psi_{ev} < \chi^2_{(1-\alpha/2)\cdot 100\%}(L - S - 1),$$

than the hypothesis H_0 is true with an error of the first sort α. In the opposite case the hypothesis H_0 should be rejected.

Rejecting H_0 hypotheses at $\psi_{ev} < \chi^2_{\alpha/2\cdot 100\%}(L - S - 1)$ contradicts common sense at first sight. However, it is necessary to note that statistics ψ_{ev} is also a random variable with its own variance, this means that very big and very small values of ψ_{ev} statistics are equally improbable.

Too small values of ψ_{ev} can result from a number of reasons. One of them is an unsuccessful choice of $F_{mod}(x_{j+1};\hat{\vec{\theta}})$ (for example, when the model parameters number has been artificially overestimated), another one can be an incorrect realization of the experiment with a deformation of the sample (grouping procedure), for example, the desire to "adjust" empirical data to the result.

The results of implementing both methods for checking the normality of the considered empirical data from Nasdaq100 you can see in table 1.

Stock	α	χ^2
AMAT	0.69	11.88
BRCD	0.98	9.86
CIEN	0.90	42.38
CSCO	0.98	18.20
DELL	0.89	9.84
INTC	0.95	24.64
JNPR	0.99	10.21
MSFT	0.67	18.05
ORCL	0.95	11.20
SUNW	0.91	35.92

Table 1: *The results of normality test of VWAP distribution. Column α consists of the values of the confidence level obtained with the help of Kolmogorov-Smirnov's test. Column χ^2 consists of the chi-square test results.*

2.8 Checking stochastic independence of successive price logarithms increments

2.8.1 Median based criterion

Let us consider sample $x_i,\ i = \overline{1,n}$ from any total population. It is necessary to check up the randomnicity and independence of the sample elements. The median based criterion will be

used for this purpose. This criterion is a rank test.

The procedure of data checking with this criterion consists of the following steps.

1. Formulation of null and alternative hypotheses.

   H_0: sample elements $x_i,\ i = \overline{1,n}$ are stochastically independent,

   H_1: sample units $x_i,\ i = \overline{1,n}$ are not stochastically independent.

2. Assignment of confidence level α.

3. Formulation of critical statistics.

   To determine the form of critical statistics, the following sequence of operations has to be executed.

   (a) Variational series have to be shaped

   $$x_{(1)} \leqslant x_{(2)} \leqslant \ldots \leqslant x_{(i)} \leqslant \ldots \leqslant x_{(n)}.$$

   (b) Evaluation of a median

   $$\hat{x}_{med} = \begin{cases} x_{((n+1)/2)}, & \text{if } n \text{ is even} \\ \dfrac{1}{2}[x_{\frac{n}{2}} + x_{n/2+1}], & \text{if } n \text{ is odd} \end{cases}$$

   (c) Each element of the initial sample $x_{(i)}$ should be replaced with "+", if $x_{(i)} > \hat{x}_{med}$, and with "—", if $x_{(i)} < \hat{x}_{med}$, step to the next element if $x_{(i)} = \hat{x}_{med}$.

   The received sequence of "+" and "—" can be characterized by the quantity of series $\nu(n)$ and by the length of the longest series $\tau(n)$. The series consists of successive "+" or "—" only. The series can consists of only one "+" or "-" sign.

   A pair of critical statistics is considered in this criterion simultaneously (bidimentional critical statistics)

   $$\psi_{cr} = \psi\{\nu(n), \tau(n)\}.$$

   The limit distribution of ψ_{cr} statistics is bidimentional with particular $\nu(n)$ and $\tau(n)$ limit distributions.

4. The definition of upper and lower critical points of distribution is carried out in the following way

   $$\nu_{cr}(n) = \frac{1}{2}(n + 1 - \sqrt{n-1} \cdot u_{1-\alpha/2}),$$

$$\tau_{cr}(n) = 3.3 \cdot \lg(n+1),$$

where $u_{1-\alpha/2}$ is a $1 - \dfrac{\alpha}{2}$ fractile of normal distribution.

5. Evaluation of critical statistics.

   $\nu_{ev}(n)$ denotes the amount of series in the initial sample, and $\tau_{ev}(n)$ is the length of the longest series. If the following conditions are satisfied simultaneously.

$$\begin{cases} \nu_{ev}(n) > \nu_{cr}(n), \\ \tau_{ev}(n) < \tau_{cr}(n), \end{cases}$$

   then H_0 hypothesis can be accepted with an error of the first kind α. Otherwise the sample elements are not stochastically independent.

This criterion has a feature of catching only monotonous change of the average (estimation of a population mean). For this reason the results obtained from the criterion considered are not reliable and more valid tests have to be carried out. One of them is the criterion of "ascending" and "descending" series.

2.8.2 Criterion of "ascending" and "descending" series

By analogy with the described above criterion of series based on the median of the sample, in the rank test of "ascending" and "descending" series the sequence of series "+" and "—" is also formed. For this purpose each i-th element of the initial sample $x_i,\ i = \overline{1,n}$ from the total population will be replaced with "+" if $x_{i+1} > x_i$, and with "—" if $x_{i+1} < x_i$, step to the next element if $x_{(i)} = \hat{x}_{med}$.

Let us consider the realisation of the criterion.

1. Formulation of null and alternative hypotheses.

   H_0: elements of the sample $x_i,\ i = \overline{1,n}$ are stochastically independent,

   H_1: elements of the sample $x_i,\ i = \overline{1,n}$ are not stochastically independent.

2. Assignment of confidence level α.

3. Formulation of critical statistics.

$$\psi_{cr} = \psi\{\nu(n), \tau(n)\}.$$

   The limit distribution of ψ_{cr} statistics is bidimentional with particular $\nu(n)$ and $\tau(n)$ limit distributions.

Stock	$\nu_{ev}(n)$	$\tau_{ev}(n)$
AMAT	333	4
BRCD	325	6
CIEN	333	4
CSCO	324	6
DELL	319	5
INTC	331	3
JNPR	330	5
MSFT	310	3
ORCL	339	4
SUNW	320	3

Table 2: *Results obtained with the rank test on independence of "ascending" and "descending" series. In $\nu_{ev}(n)$ column there are calculated values of the amount of series, in $\tau_{ev}(n)$ column there is the calculated value of the longest series length. The values of critical statistics for the considered sample are: $\nu_{cr} = 324$, $\tau_{cr} = 7$.*

4. Definition of upper and lower limit points

$$\psi_{cr.u} = \nu_{cr}(n) = \frac{1}{3}(2n-1) - \sqrt{\frac{16n-29}{90}} \cdot u_{1-\frac{\alpha}{2}},$$

$$\psi_{cr.l} = \tau_{cr}(n) = \begin{cases} 5, \text{if } n \leqslant 26 \\ 6, \text{if } 26 < n \leqslant 153 \\ 7, \text{if } 153 < n \leqslant 1170 \end{cases}$$

where $u_{1-\frac{\alpha}{2}}$ is a fractile of normal distribution.

5. Evaluation of $\nu_{ev}(n)$ and $\tau_{ev}(n)$ statistics.

$\nu_{ev}(n)$ is the quantity of series.

$\tau_{ev}(n)$ is the length of the longest series.

If the following conditions are satisfied simultaneously

$$\begin{cases} \nu_{ev}(n) > \nu_{cr}(n), \\ \tau_{ev}(n) < \tau_{cr}(n), \end{cases}$$

H_0 hypothesis can be accepted with an error of the first kind α. Otherwise the elements of the sample can not be stochastically independent.

This criterion catches monotonous and periodic shift of the estimation of the mean. It is a more powerful criterion than the median based criterion of series.

2.8.3 Abbe's criterion of stochastic independence

If sample $x_i,\ i = \overline{1,n}$ belongs to the normal total population it is better to use the successive differences criterion of squares (Abbe's criterion) to determine whether the elements of this sample are stochastically independent or not. The decision to use this criterion is based on the results of normality tests of the empirical data distribution which were carried out in the given research. These results allow to speak about high degree of normality of the data considered.

Abbe's criterion allows to find out a regular shift of the mean during the sample investigation.

Stock	AMAT	BRCD	CIEN	CSCO	DELL	INTC	JNPR	MSFT	ORCL	SUNW
ψ_{ev}	1.336	1.626	3.603	1.397	1.891	1.865	6.926	7.291	2.156	10.81

Table 3: *Calculated values of critical statistics for Abbe's criterion. The value of the lower critical point for the sample used is* $\psi_{cr.l} = 1.004$.

The procedure of data checking with this criterion consists of the following steps.

1. Formulation of null and alternative hypotheses.

   H_0: sample elements $x_i,\ i = \overline{1,n}$ are stochastically independent,

   H_1: sample units $x_i,\ i = \overline{1,n}$ are not stochastically independent.

2. Assignment of a confidence level α.

3. Formulation of critical statistics.

$$\psi_{cr} = \gamma_{\alpha}^{(n)} = \frac{\hat{q}^2(n)}{\hat{\sigma}^2},$$

where

$$\hat{q}^2(n) = \frac{1}{2(n+1)} \sum_{i=1}^{n-1} (x_{i+1} - x_i)^2,$$

$\hat{\sigma}^2$ is the unbiased estimation for the variance of a sample.

The limit distribution of critical Statistics $\gamma_\alpha^{(n)}$ is tabulated for each $n \leqslant 60$. These data can be found in tables of critical points of Abbe's distribution for different values of α.

4. The definition of the lower critical point can be carried out in two ways.

   If $n > 60$,

$$\psi_{cr.l} = 1 + \frac{u_{1-\frac{\alpha}{2}}}{\sqrt{n + 0.5 \cdot (1 + u_{1-\frac{\alpha}{2}})^2}},$$

   where $u_{1-\frac{\alpha}{2}}$ is a fractile of the standard normal distribution.

   If $n \leqslant 60$ then $\psi_{cr.l}$ should be found in statistical tables.

5. Evaluation of critical statistics

$$\psi_{ev} = \frac{\hat{q}^2(n)}{\hat{\sigma}^2}.$$

   The hypothesis about stochastic independence of the sample units is accepted under $\psi_{ev} > \psi_{cr.l}$, otherwise the elements of sample cannot be accepted as random and independent.

3 Portfolio risk estimation

Portfolio risk can be estimated in two ways: a) considering each stock included into the portfolio separately; b) considering all of them in aggregate. That will correspond in some sense to a Value-at-Risk approach, but with a division between two different portfolios: a general portfolio and a current short-time portfolio.

In the first case, the risk of each stock should be calculated. Then, taking into account the correlation between stocks, the risk of the whole portfolio is estimated. This is a difficult problem largely because such correlations are not determined values.

In the simplest Gaussian case this method leads us to calculate the correlation matrix to the n-th order, where n is the number of stocks. This matrix depends on the time and state of the system considered .

Volatility of a portfolio takes the form

$$D^(\pi_t)[V_t] = [\pi_t^{S_1}, \ldots, \pi_t^{S_n}] \Sigma_t(S_1, \ldots, S_n)(\pi_t^{S_1}, \ldots, \pi_t^{S_n}),$$

where Σ is a correlation matrix and $\pi_t^{S_i}$ is the total number of stock S_i at the moment t.

This formula can be dynamically calculated for n not greater than 20 stocks for the general case, where we shall use some form of the ARCH-GARCH-EGARCH model.

It was decided to use the second approach for this reason. In the second case the portfolio is represented as a generalized stock the price of which is averaged according to the shares of the stocks, included into the portfolio.

This approach is quite good and can be improved for the dynamic portfolio case. One should take into account possible changes in the portfolio. These changes are considered as components of another portfolio, consisting of this generalized stock and some additional stocks (in short or long positions) which are the result of the current days activity. In this case we are able to make calculations in accordance with the first approach because the number of stocks is much more lower than in the whole portfolio.

3.1 Value-at-Risk methodology of risk estimation

Profit earning is inseparably linked with risk. One of the basic positions of modern financial theory is, that earning a higher profit is connected with a higher risk. It is intuitively clear to everyone what risk is, but obtaining quantitative estimations of risk involves serious difficulties.

One of the primary issues of modern financial institutions is estimation of market risks that arise due to fluctuations of stock prices, raw goods, exchange rates, interest rates etc. The simplest measure of investor dependence on market risk is the value of portfolio capital change, i.e. profits or losses arising due to asset prices movement. At present, the most common methodology for market risk estimation is *Value-at-Risk* (VAR). VAR is a summary measure of risk that compare risks of various portfolios (for example, of portfolios consisting of stocks and bonds) and of various financial instruments (for example, forwards and options).

For the last few years VAR has became one of the most popular means of risk management and risk control in companies of various types.

VAR is a *statistical approach*. The basic concept is the *distribution of probabilities* connecting all possible values of changes in market factors with their probabilities. VAR methodology has a number of doubtless advantages: it allows measurement of risk in terms of possible losses related to the probabilities of their occurrence; it allows to measurement of risks with a universal method in various markets; it allows aggregation of risks of separate positions into a single value for the whole portfolio, taking into account the size of positions, volatilitie in

the market and the duration positions are held. Thus, VAR is a really universal approach to market risk measurement.

What is VAR? VAR is a *statistical estimation of the maximum loss exposure of a financial organizations portfolio, measured under the given distribution of market factors for the given period of time for all cases, except for the given small number of situations.*

Development and introduction of VAR models occurs in a very intensive way. In investment companies and banks VAR methodology can be implemented in several business lines at the same time.

• *Internal market risks monitoring.* Institutional investors can calculate and monitor the values of VAR on several levels: aggregated portfolio, the asset class, the emitter, the contractor, the trader/portfolio manager etc. From the monitoring point of view the precision of VAR values estimation is of secondary importance. In this case the relative value of VAR is more important than the absolute one, i.e. manager's VAR or VAR of a portfolio compared to the VAR of the model portfolio, the index, to another manager or to the same manager at previous times.

• *External monitoring.* VAR allows creation of a representation of the market risk of a portfolio without disclosing the information about the structure of the portfolio (which can be very complex). Regular reports using VAR values can serve as one of the arguments that the risk accepted by managers was reasonable.

• *Hedge efficiency monitoring.* VAR values can be used for determining the degree of consistency of the *hedging strategy.* The manager can estimate the efficiency of the hedge comparing VAR values for portfolioes with and without a hedge. If, for example, the difference between these values is insignificant, there is a question of the expediency of the hedging or whether the hedging is applied correctly.

• *Analysis of possible trades.* VAR methodology allows giving more freedom and autonomy to managing personnel, because bureaucratic procedures connected with trading (especially with derivative instruments) can be reduced. This is achieved by monitoring transactions using VAR.

Thus, companies can use VAR values to create reports for managers, shareholders and external investors, because VAR allows aggregation of every possible market risk into it a single value having a monetary expression. Estimated risk can be calculated for different market segments, marking out the most risky positions. VAR estimations can be used to diversify capital holdings, to set limits and also to estimate the company's activity. In some

banks, an estimation of traders' operations (performance) and also compensation is calculated based on the rate of profitability per VAR unit.

3.1.1 Financial portfolio risk estimation

The modern investment portfolio can consist of tens of thousands of various financial instruments inform various world markets. How should the risk of such a portfolio be considered?

Portfolio risk could be considered by calculating the distribution function of the portfolio cost change during some period of time. This requires considering all possible scenarios of developing market events by evaluating portfolio costs of each scenario.

Even if it could be measured, there is a problem comparing the risks of two different portfolios. To compare their risks, each risk should be expressed by a single number. If a strictly determined estimation of the risk could be obtained, it would be a real breakthrough for monitoring of financial institutions activities. Unfortunately, it is impossible to obtain any precise estimation of the risk as a scalar value due to potentially infinite number of possible scenarios of an event developing. Therefore, when speaking about portfolio risk, one always means some estimation (generally probabilistic) of the value of the risk.

It might seem that the modern theory of financial portfolio management answers the question of what risk is and how to measure it. According to the theory, risk is a standard deviation from the portfolio cost. Representation of a risk as a standard deviation has some serious faults from a practical point of view:

Firstly, portfolio managers generally prefer to obtain risk information in terms of potential real monetary losses, instead of in the form of a standard deviation;

Secondly, the standard deviation takes into account favourable and adverse changes of the portfolio cost. If the distribution of change of the portfolio cost has a symmetric form, then a standard deviation gives the correct value of risk, but a modern portfolio includes options and similar financial instruments. The change of cost for such instruments is nonlinear with respect to market prices. This results in asymmetry of changes of the portfolio cost distribution function. Standard deviation gives an incorrect estimation of risk under this condition.

In conclusion, VAR methodology has been developed to represent information about the investment portfolio risk with a single number.

3.1.2 VAR definition

The exactly definition of VAR is formulated as follows.

Let the portfolio be fixed. *VAR of a portfolio under given confidence level α and the given period of positions holding t is defined as such value V, which provides a covering of possible losses x of the portfolio holder for time t with probability α, i.e.* $P(x, V) = \alpha$. From a probability theory point of view VAR is the fractile α for given distribution. In other words, having calculated VAR, it is possible to formulate the statement of the following type: "We are sure with α% probability that the losses will be within VAR for the next N days".

In general the problem of VAR calculation with the given parameters t (horizon of forecasting) and α (confidential probability) has the following steps:

1. The current cost of the investment portfolio is determined as V_0;
2. Portfolio cost in the time t we should define as $V_t = V_0\, e^r$;
3. Let us predict such $\tilde{r}$ that $\mathsf{P}(r < \tilde{r}) = \alpha$;
4. Let us obtain the worst variant of cost of portfolio cost $\widetilde{V}_t = V_0\, e^{\tilde{r}}$, then $VAR = V_0 - \widetilde{V}_t$.

Speaking the language of mathematics $VAR = VAR_{t,T}$ is defined as the upper border of a unilateral confidential interval

$$Probability(R_t(T) < -VAR) = 1 - \gamma,$$

where γ is the confident level, $R_t(T)$ is the rate of growth of the portfolio capital in an interval $[t, T]$ under "a continuous way of interests charges"

$$R_t(T) = \ln(V(t+T)/V(t)),$$

where $V(t+T)$ and $V(t)$ are values of the portfolio capital at the $t+T$ and t moments of time accordingly. In other words, $V(t+T) = V(t)\, e^{R_r(T)}$.

It should be noted that $R_{(T)}$ is a random variable and it is characterized by some *probabilistic distribution*. VAR value is calculated from the distribution of increments as follows:

$$1 - \gamma = F_R(-VAR) = \int_{-\infty}^{-VAR} f_R(x)dx,$$

where $F_R(X) = Probability(R \leqslant x)$ is the *distribution function* of portfolio growth rates, $f_R(x)$ is the *density function* for $R_t(T)$ distribution.

As itfollows from VAR definition, to calculate it, is necessary to know the structure of the portfolio, the interval of time for which VAR is calculated, and the distribution function of the portfolio cost change.

Besides, it is necessary to determine the number of the base elements influencing VAR value. First of all it is a probabilistic distribution of the market factors which directly influence the portfolio assets price changes. It is clear that some statistics of each of these assets' behaviour through the time is necessary to construct the distribution. It is enough to estimate only the volatility (i.e. a standard deviation) under the assumption that the logarithms of assets price changes belong to the normal gaussian distribution with the mean equal to zero. However in the real market, the assumption about the normality of Distribution does not work as a rule. After setting up market factors distribution it is necessary to choose a confidence level, i.e. a probability at which our losses should not exceed VAR. Then it is necessary to determine the Position's holding period during which the losses are estimated. Under some simplifying assumptions it is easy to see, that VAR of a portfolio is proportional to the square root from the position's holding period. That is why it is enough to calculate only one-day VAR. Then, for example, four-day VAR will be twice as much.

Besides if the portfolio contains complex derivative financial instrument (options, for example), it is necessary to choose their *pricing function* depending on the parameters of the market. At last, it is necessary to determine the *correlations* between various market factors.

Getting the information on the structure of a portfolio is not so trivial an issue, as it can seem. Large companies have very complex portfolios including thousands of various financial instruments which are in world markets, besides, as a rule these companies perform very active trading operations. For this reason they face the problem of how to quickly get the information on the current structure of a portfolio.

Another problem is the choice of the fixing time for the prices of assets included into the portfolio. Trading sessions in the world markets have different closure time, it creates a problem of choosing the price hich will be used to calculate the portfolio cost change. The closure time of that market where basic assets of the company are concentrated is usually accepted as the fixing time.

Having determined the structure of the portfilio and having chosen the desirable period of time, it is necessary to determine the distribution function for the portfolio cost change.

3.1.3 Methods of VAR estimation

The example explaining VAR concept is given in figure 3. The curve in the figure is the density function for the probabilities of profits and losses for the given portfolio and the positions' holding period. The light part of the figure corresponds to the chosen confidence level (97,5%)

in the sense that its area makes a 97,5% from the general area under the curve. VAR represents the value of possible losses meeting the given confidence level.

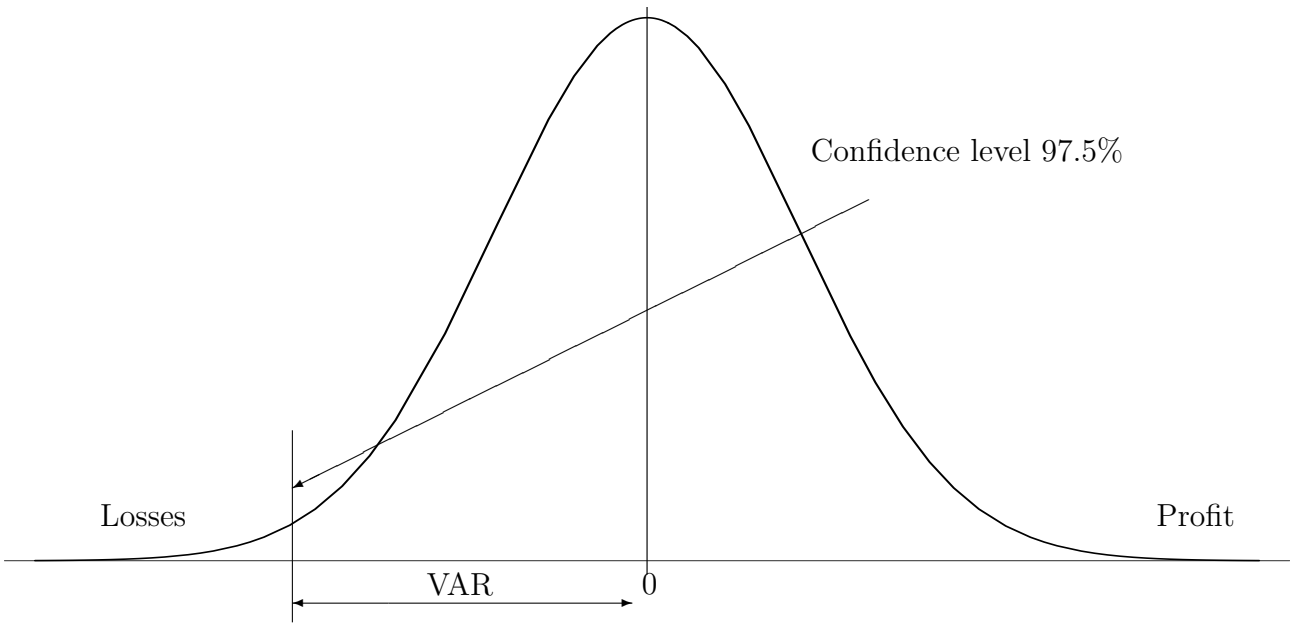


Figure 3: *Distribution of the profit and losses for a portfolio with indication of VAR.*

So after all the base elements have been marked, it is possible to address directly to Value-at-Risk calculation.

There are three basic methods of the distribution function parameters evaluation: historical, analytical and simulation ones.

Analytical method. The basic idea of the method consists in revealling the market factors influencing the cost of a portfolio, and approximating the portfolio cost on the basis of these factors. In other words the financial instruments which aggregate a portfolio, are divided into elementary assets, as far as it is possible, in such a way, that changes of each one depend only on one market factor. For example, the long-term coupon bond can be considered as a set of bonds with different maturity dates.

Further the assumption about the form of market Factors distribution is made. Usually it is supposed that the profitability of market factors submits to normal distribution. The mean and variance of the data as well as the correlations between market factors are calculated on the basis of historical data. Having estimated to some extend standard deviations of price logarithms changes for each asset included into the portfolio, VAR should be calculated as the production of standard deviations and the multiplier corresponding to the confidence level

(for the confidence level of 97,5% the multiplier is equal to 1.96, for example). It is necessary to know the correlations between the assets to calculate total VAR of the portfolio. The distribution of profitability of a portfolio will be normal if the approximation is linear. One can evaluate the distribution parameters for the whole portfolio, knowing the parameters of market factors distributions. All the necessary parameters of normal distribution are well-known for the majority of market factors and this is a serious advantage of the considered approach (see http://www.jpmorgan.com/- Risk Metrics).

Example 1.

Let the portfolio consist of a thousand futures per a USA dollar with maturity date on January, 15, 2013 and let the current futures price in some stock exchange be 6 s.u./dollars (then the cost of all portfolio is 6 million s.u.). Let there also be some statistics about the futures' prices F_i for the last N days. Let us consider variables $x_i = \ln \frac{F_i}{F_{i-1}}$ which are the logarithms of one-day changes of the futures' prices. Let x_i be a set of random variables having normal gaussian distribution with the mean equal to zero. Then it is possible to estimate its volatility (i.e. a standard deviation) according to the known formula $\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}$. Let us assume that σ takes on a value of 0,3%. Then VAR value for the given portfolio, corresponding to a 97,5% confidence level and one-day positions' holding period, will be equal to:

VAR = 1,96 $\cdot\sigma\cdot$6 000 000 s.u. = 35 280 000 s.u.

In the given example a normal distribution was chosen only for the illustrative purposes by force of calculation simplicity. In practice, as it is known, the increments of the asset prices have heavier "tails" in comparison with the normal distribution, i.e. in reality there are more "extreme" events on the market then it would be possible to expect at normal distribution. VAR by its nature, deals with the prediction of events from the "tails" of distribution. Therefore using Pareto-like distributions seems to be more consistent. For these distributions, the probability of large deviations is given by the following expression:

Probability($R \leqslant x$)=$F_R(x) \approx ax^{-\alpha}$ under $x \to \infty$.

Here $a > 0$ is a constant, α is a so-called "tail" index.

Apart from Pareto-like stable distributions, it is possible to use distributions belonging to the class of generalized hyperbolic distributions.

It should be noted that the estimation of VAR obtained with the help of analytical method most closely coincides with the estimation of risk offered by the modern portfolio theory.

The analytical method is simple in realization and allows to obtain VAR estimation very quickly (it is possible, even in real-time mode) almost on any computers. But the quality of the

estimation worsens if there is an increase in the share of instruments with non-linear payment function in the portfolio. Besides the necessity to make an assumption about the form of the distribution for base assets is a serious fault of this method.

Historical method. The historical method implies research of the portfolio cost change during the previous historical period.

For VAR calculation a database is created for some predetermined historical period of price values for the financial instruments which are included into the portfolio (or the appointed market factors, if a portfolio is approximated). After that it is necessary to calculate the instruments' price changes for the period of time which will be used for VAR calculation and to obtain the appropriate values of the portfolio cost change. Then it is necessary to arrange the obtained data to construct the histogram of the portfolio cost changes distribution and to find VAR value corresponding the to chosen value of probability.

The procedure of VAR calculation is as follows. A time period is chosen (100 trading days, for example) during which relative changes of the prices for all the assets included into today's portfolio, are traced. Then, for each of these changes a possible change of the portfolio cost is calculated, then 100 obtained numbers are sorted out by decrease. The number taken with the opposite sign corresponding to the chosen confidence level (for example, for 99% level it is necessary to take the value with number 99), will represent the VAR of a portfolio.

This method has two main advantages: firstly there is no any necessity to know the form of the distribution function of portfolio market factors, secondly it is a simple in realization. There are no any difficulties with the portfolios which include options and financial instruments of the same type.

One of the faults of the method in question is the necessity to do big work on gathering and processing historical data. Besides an estimation of possible changes of the portfolio cost is limited to a set of previous historical changes. The absence of the sufficient amount of historical data is a typical problem of the given method. To obtain more exact estimation of VAR, it is necessary to use as much data as it is possible, but using too old data may lead to estimating the risk on the basis which does not correspond to the current state of market.

Monte Carlo method. This method consists in modelling possible changes of a portfolio cost under some assumptions. The basic market factors influencing the cost of a portfolio are determined. Then the summary distribution of these factors is constructed somehow, using

historical data, for example, or the data based on any scenario of economy development. After that a big number of possible scenarios of the situation development are modelled, and the change of the portfolio price is calculated for each scenario of modelling. Further a histogram of the obtained data should be constructed and the value of VAR evaluated.

In contrast to historical modelling, in the Monte Carlo method The assets price changes are generated in a pseudo-random way according to predetermined parameters. The simulated distribution can be almost voluntary defined, and the number of scenarios can be rather big (up to tens of thousands). In other respects the method is similar to historical modelling.

The Monte Carlo method is notable for its high accuracy and suitability almost for any portfolios, but its implementation demands a certain mathematical base of the experts and sufficient computer resources.

This method has several advantages. It does not use any concrete model of parameters' definition ant it can be easy recustomized according to the economic forecast. The method does not model the final cost of a portfolio, but the whole scenario of a situation development. This circumstance allows to trace the change of the portfolio cost depending on how situation develops.

The fault of the ethod is its slow convergence that results in essential time and computing expenses.

Concrete models of VAR estimation are based on various combinations of considered methods.

Generally speaking, it is difficult to recommend one of the methods of VAR calculation. It is necessary to take into account the macroeconomic situation, and also the purposes and issues of the concrete organization when choosing any method.

3.1.4 The faults of VAR

A common fault of VAR is that all the models, irrespectively of the calculation methods used, use historical data, and under fast changing market conditions it will give a sufficient delay. For example, if volatility has a spasmodic behavior changes VAR can take these changes into account only in some lag, but up to this moment the estimation of VAR will be incorrect.

What is not taken into account when estimating VAR is such an important characteristic of markets as liquidity. At some moments it can result in serious difficulties in changing the portfolio structure to reduce the risk.

This or that model is used for VAR the estimation, and it means the presence of a modelling

risk in calculations. Therefore it is necessary to periodically check the adequacy of the model used.

VAR estimates the probability of occurrence of losses which exceed a certain level, it means that it estimates the "weight of the tail" of distribution, but it does not say anything about the possible value of these losses. Therefore in addition to VAR it is recommended to study the behaviour of a portfolio in stressful situations to estimate the "length of the tail" of the distribution.

The result of all these factors is that VAR works well in case of stable conditions in the markets and ceases to display the value of risk adequately if dramatic changes occur in the markets. It is necessary to remember that VAR is just one of tools in risk management, but it is not a universal method of estimation.

In conclusion it is desirable to note that VAR methodology is not a panacea from financial losses. It just helps companies to estimate risks and to decide whether these risks are accepted or just desirable to be accepted. VAR can and should be used in addition to other methods of risk analysis such as, for example, *Shortfall-at-Risk* (SAR) when one is interested in not just the boundary size of the capital, below which a loss could be expected with a certain share of probability, but also in the value of this loss.

3.2 VWAP of a portfolio

It would be logical to define the price and VWAP of a portfolio as follows:

$$p = \sum_{i=1}^{n} \alpha_i p_i,$$

where $\alpha_1 + \cdots + \alpha_n = 1$, α_i is the share of a stock in portfolio, and p_i is the price of this stock.

$$VWAP = \sum_{i=1}^{n} \alpha_i \, \mathrm{vwap}_i,$$

where vwap_i is VWAP of i-th stock.

As VWAP value of a portfolio is a linear composition of VWAP values of separate stocks, its distribution has the same type as those of its summands.

Statistical tests performed above showed that our data seem to belong to the normal total population, however, not all of the stocks considered satisfy the normal distribution with a sufficient confidence level, hence the statement about the normality of the whole portfolio has to be checked. For this purpose χ^2 and Kolmogorov-Smirnov's criterions of consent were used, the results of The testing are submitted in the table 5.

If the confidence level is sufficient, it is possible to accept the assumption of normality for successive price increments of the portfolio VWAP distribution and to use brownian motion for the volatility estimation. Otherwise it is necessary to use another methods of σ estimation, for example, those described in section 2.

Further we shall assume for all the examples, that the confidence level allows to accept a hypothesis about the normality of the distribution. Besides, it is supposed that volatility σ is a constant value, though many researches showed, that it is not absolutely so for the empirical data.

3.3 Portfolio risk estimation

VWAP for the whole day is summed up from hourly values arcwise. For this reason it is possible to make a rough assumption, that the normality of VWAP hourly value increments results in normality for VWAP daily increments. This fact is also proved with the χ^2 and Kolmogorov-Smirnov's criterions of consent.

Successive increments of VWAP logarithms are a wiener process:

$$VWAP_t = VWAP_0 \cdot e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t}$$

Using the rule of triple sigma[1], let us define the lower border of this value as follows:

$$VWAP_t = VWAP_0 \cdot e^{(\mu - \frac{9\sigma^2}{2})t - 3\sigma \cdot \sqrt{t}}$$

It is supposed, that the portfolio is sold by equal parts in n days, consequently its risk is (in percents) estimated with the formula:

$$R = (1 - \frac{\frac{1}{n} \cdot \sum_{i=1}^{n} VWAP_i}{VWAP_0}) \cdot 100, \quad \text{where} \quad VWAP_i = VWAP_0 \cdot e^{(\mu - \frac{9\sigma^2}{2})i - 3\sigma \cdot \sqrt{i}}.$$

The examples of risk estimation for various portfolios consisting of some stocks which belong to Nasdaq100 at n=5, are submitted in table 5, and also in diagrams in Appendix 2.

The procedure of risk estimation was the following. The distribution parameters were estimated in some predetermined historical period of data, and then with the help of the

[1]This rule is independent of probability distribution and in general case is equal to the assert that probability

$$\mathsf{P}\left\{|\mathsf{E}[\xi] - \xi| > 3\sqrt{\mathsf{D}[\xi]}\right\}$$

is lower than 1/9. It can be shown that for the Normal distribution this probability not greater than 1/100.

Portfolio	AMAT	BRCD	CIEN	CSCO	DELL	INTC	JNPR	MSFT	ORCL	SUNW
No.1	0.070	0.120	0.042	0.199	0.202	0.109	0.060	0.032	0.142	0.021
No.2	0.093	0.118	0.036	0.060	0.117	0.099	0.109	0.020	0.157	0.057
No.3	0.081	0.166	0.163	0.130	0.108	0.128	0.092	0.025	0.067	0.036
No.4	0.027	0.116	0.121	0.068	0.096	0.100	0.120	0.095	0.033	0.120
No.5	0.026	0.135	0.187	0.119	0.088	0.107	0.095	0.119	0.033	0.187

Table 4: *Shares of stocks in portfolios*

procedure described above lower and upper bounds of estimation were evaluated. After these procedures the actual trajectory of the price was considered for the period of time for which risk estimations were obtained.

In general case the independence of VWAP increments is sufficient for use of the similar procedure for lower and upper board calculation:

$$R = (1 - \frac{\frac{1}{n} \cdot \sum_{i=1}^{n} VWAP_i}{VWAP_0}) \cdot 100, \quad \text{where} \quad VWAP_i = VWAP_0 \cdot e^{(\widetilde{\mu})i - 3\widetilde{\sigma} \cdot \sqrt{i}},$$

and $\widetilde{\mu} = \mathsf{E}[\ln(\Delta VWAP_i)]$ is an expectation of the VWAP increment $\ln(\Delta VWAP_i)$ and $\widetilde{\sigma} = \sqrt{\mathsf{D}[\ln(\Delta VWAP_i)]}$ is a standard deviation of the VWAP increment $\ln(\Delta VWAP_i)$.

4 Conclusion

The present research shows that using VWAP as a risk estimation criterion strongly simplifies the procedure of estimation. The examples considered obviously show that the offered approach gives a good estimation of risk, even with quite rough assumptions of normality for successive VWAP increments and a triple sigma rule.

The assumption of normality is not crucialhere, because a triple sigma rule is independent of probabilitydistribution, so we can use it in other cases.

Also mentioned was a very important problem connected with market collapse from sale of an enormous quantity of shares. The correct estimation of market stability is crucial not only for risk estimation, but also for the technical realization of selling shares at VWAP. Accomplishing sale of shares at VWAP can be strongly dependent on the quantity of shares to be sold in a set

ID of a portfolio	Normality		Evaluation of risk				Evaluation of actual loss			
	α	χ^2	I	II	III	IV	I	II	III	IV
No.1	0.9945	2.04	-12.47	-12.15	-12.37	-12.43	10.52	-2.75	-5.66	2.89
No.2	0.9903	8.07	-13.23	-13.07	-13.40	-13.25	11.04	-7.14	-2.32	-0.76
No.3	0.8368	3.23	-14.08	-14.21	-14.51	-14.11	2.80	-7.01	-2.70	-0.74
No.4	0.9874	3.15	-12.6	-12.76	-13.08	-12.85	-0.73	-6.85	-1.02	-3.50
No.5	0.9926	5.28	-12.49	-12.96	-13.01	-12.84	-2.14	-4.79	1.25	-4.60

Table 5: *Portfolio risk estimation. α and χ^2 columns consist of the confidence level obtained from the Kolmogorov-Smirnov's criterion and the results of χ^2 criterion implementation accordingly. The corresponding estimations for risk at various moments of time are represented in the subsequent four columns. The actual values of loss, given for the comparison with the corresponding estimations of risk are represented in the last four columns.*

period of time.

In general, the assumed risk-estimation procedure looks as follows:

- Analysis of the distribution properties and VWAP process construction.
- Generalized stock calculation and risk estimation on a portfolio basis at the end of a trading day.
- Use of this generalized stock as an ordinary stock for short-time portfolio risk estimation on the next trading day.
- Combining the cumulative portfolio at the end of a trading day.

References

[1] *Abken P.A.* An Empirical Evaluation of Value at Risk by Scenario Simulation. March, 2000.

[2] *Eberlein E., Kallsen J., Kristen J.* Risk Management Based on Stochastic Volatility. University of Freiburg, 1995.

[3] *Dai Bo.* Value at Risk. Department of Mathematics National University of Singapore, 2001

[4] *Jorge Mina and Jerry Yi Xiao.* Return to RiskMetrics: The Evolution of a Standard RiskMetrics, 2001.

[5] *Shao J.* Mathematical statistics. Springer, 2008.

[6] *Eberlein E., Keller U., Prause K.* "New insights into smile,mispricing, and value at risk: the hyperbolic model". Journal of Business, 1998, v.71, No.3, pp. 371-405.

[7] *Eberlein E.* "Recent advances in more realistic market and credit risk management: the hyperbolic model". Preprint. Univ. of Freiburg, Inst. for Mathematic Stochastic. February 25, 2000.

[8] *Eberlein E., Keller U.* "Hyperbolic distributions in finance". Bernoulli 1, 1995, pp.281-299.

[9] *Barndorff-Nielsen O.E.* "Gaussian—Inverse Gaussian Processes and the Modeling of Stock Returns". Preprint: October, 1994, Aarhus University.

[10] *Barndorff-Nielsen O.E., Jiang W.* "An initial analysis of some German stock price series". Working Paper Series No. 15, 1998, Aarhus University.

[11] *Blæsild P., Sørensen M.K.* " A computer program for analyzing data by means of hyperbolic distribution". Research Report 248, Dept. Theot. Statistics, Aarhus University.

[12] *Barndorff-Nielsen O.E., Jiang W.* "An initial analysis of some German stock price series". Working Paper Series No. 15, 1998, Aarhus University.

[13] *Baillie R.T., Bollerslev T., Mikkelsen H.-O.* Fractionely Infrtegrat-ed Generalized Autoregressive Conditional Heteroskedasticity. Working paper No. 168. Evanston, IL: Kellogg Graduate School of Management, North-western University, 1993 // Journal of Econometrics. 1996. V.74. No.1.

[14] *Barndorff-Nielsen O.E.* Inverse Gaussian Distributions and Stochastic Volatility Modelling. Preprint. Aarhus: Aarhus University, Dept. of Mathematical Sciences, March 5, 1996.

[15] *Barnea A., Downes D.* A reexamination of the empirical distribution of stock price changes // Journal of American Statistical Association, 1973. V.68. P.348-350.

[16] *Bollerslev T.* Generalized autoregressive conditional heteroskedasticity // Journal of Econometrics. 1986. V.31. pp.307-327.

[17] *Cootner P.H.* (Ed.). The Random Character of Stock Market Prices. Cambridge, MA: MIT Press, 1964.

[18] *De Vries C.G.* On the relation between GARCH and stable processes // Journal of Econometrics. 1991. V.48. pp.313-324.

[19] *Ding Z., Granger C.W.J., Engle R.F.* A long memory property of stock market returns and a new model // Journal of Empirical Finance. 1993. V.1. pp.83-106.

[20] *Durrett R.* Brownian Motions and Martingales in Analysis. Belmont, CA: Wadsworth, 1984.

[21] *Einstein A.* Investigation on the Theory of the Brownian Movement. (Ed. R. Fürth.) New York: Dover, 1956.

[22] *Engle R.F.* Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation // Econometrica. 1982. V. 50. No.4. pp.987-1008.

[23] *Fama E.F.* The behavior of stock market prices // Journal of Business 1965. V.34. pp.420-429.

[24] *French K.R., Schwert G.W., Stambaugh R.* Expected stock returns and volatility // Journal of Financial Economics. 1987. V.19. pp.3-29.

[25] *Gnedenko B.V.* Course of the probability theory. Eddition 6-th. Science, 1988.

[26] *Hagerman R.* More evidence on the distribution of security returns // Journal of Finance. 1978. V.33. pp.1213-1221.

[27] *Hamilton J.D.* Time Series Analysis. Princeton, NJ: Princeton Univ. Press, 1994.

[28] *Hill B.M.* A simple general approach to inference about the tail of a distribution // The Annals of Statistics. 1975. V.5, No.3. pp.1163-1173.

[29] *Karatzas I., Shreve S.E.* Brownian Motion and Stochastic Calculus. Berlin: Springer-Verlag, 1988.

[30] *Küchler U., Neumann K., Sørensen M., Streller A.* Stock Returns and Hyperbolic Distributions. Discussion paper No.23. Berlin: Humboldt University, 1994.

[31] *Lin S.J.* Stochastic analysis of fractional Brownian motion // Stochastics and Stochastics Reports. 1995. V.55. pp.121-140.

[32] *Linton O.* Adaptive estimation in ARCH Models // Econometric Theory. 1993. V.9. pp.539-569.

[33] *Mandelbrot B.B., Taylor H.M.* On the distribution of stock price difference // Operations Research. 1967. V.15. No.6. pp.1057-1062.

[34] *Mandelbrot B.B., van Ness J.W.* Fractional Brownian motions, fractional noises and applications // SIAM Review. 1968. V.10. No.4. pp.422-437.

[35] *Nelson D.B.* Conditional heteroskedasticity in asset returns: a new approach // Econometrics. 1990. V.59. pp.347-370.

[36] *Osborne M.F.M.* Brownian motion in the stock market // Operations Research. 1959. V.7. pp.145-173.

[37] *Schwert G.W., Segnin P.J.* Heteroskedasticity in stock returns // Journal of Finance. 1990. V.45. pp.1129-1155.

[38] *Shiryaev A.N.* Probability. 2nd ed. New York: Springer-Verlag, 1995.

[39] *Stein E.M., Stein J.* Stock prices distributions with stochastic volatility: an analytic approach // Review of Financial Studies. 1991. V.4. No.4. pp.727-752.

[40] *Borovkov A.A.* Probability Theory. Springer, 2013.

[41] *Bollerslev T., Engle R. F., Nelson D.B.* "ARCH Models". Prepared for Handbook of Econometrics, Volume 4. July, 1993.

[42] *Hansen P. R., Lunde A.* "A comparison of volatility models: Does anything beat a GARCH(1,1)?". Centre for Analytical Finance, working paper series No.84. March, 2001.

[43] *Sang-Won Lee, Bruce E. Hansen* "Asymptotic theory for GARCH(1,1) Quasi-Maximum Likelihood Estimator". Indiana University December,1992.

[44] *Micosh T., Starica C.*, "Cange of structure in finansial time series, long range dependence and the GARCH model". Working paper series No.58, Cenre for Analytical Finance, April, 2000.

[46] *Shiryaev A.N.* Fundamentals of stochastical financial mathematics. Models. Facts. FAZIS, Moscow 2008.

[47] *Brandorff-Nielsen O.E.* Probability densities and Levy densities. CAF, July, 2000.

[48] *David Nualart, Wim Schoutens.* Chaotic and Predictable Representation for Levy Processes, Stochastic Processes and their Applications, 2001.

[49] *Brandorff-Nielsen O.E., Shepard N.* Modelling by Levy processes for financial econometrics. Birkhauser, Boston, 2002.

[50] *Klebanov, L., Rachev, S. T., Safarian M.* Local Pre-limit theorems and their applications to finance, Applied Mathematics Letters, 13, 2000, 73-78.

[51] *Brooks, C., Clare, A. D., Molle, J. and Persand, G.* A comparison of extreme value theory approaches for determining value at risk, Journal of Empirical Finance , Vol. 12, 2005, pp. 339-352.

Appendix 1

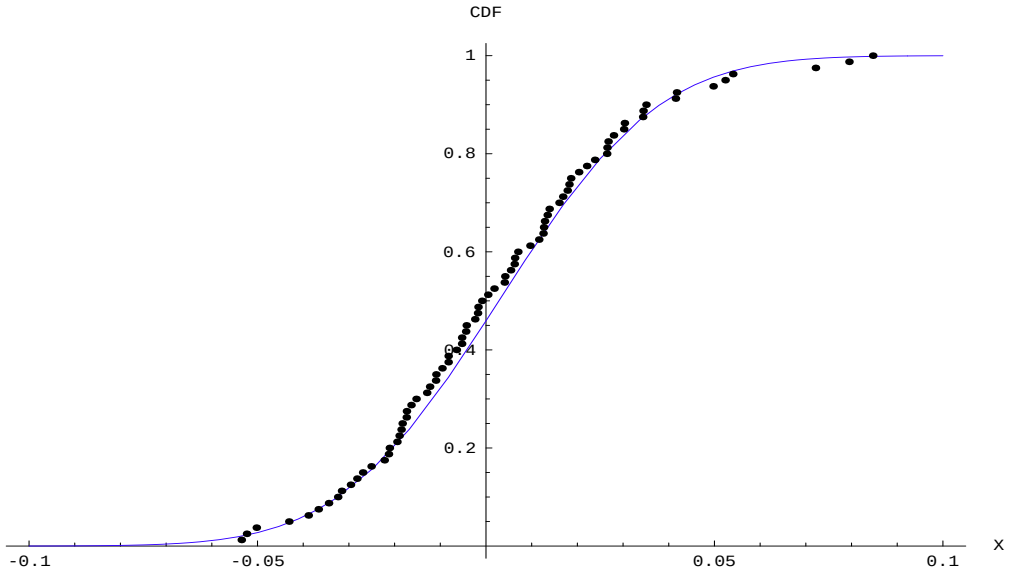


Figure 4: Distribution function for daily increments of VWAP logarithms, obtained from top of the book on Nasdaq for Intel.

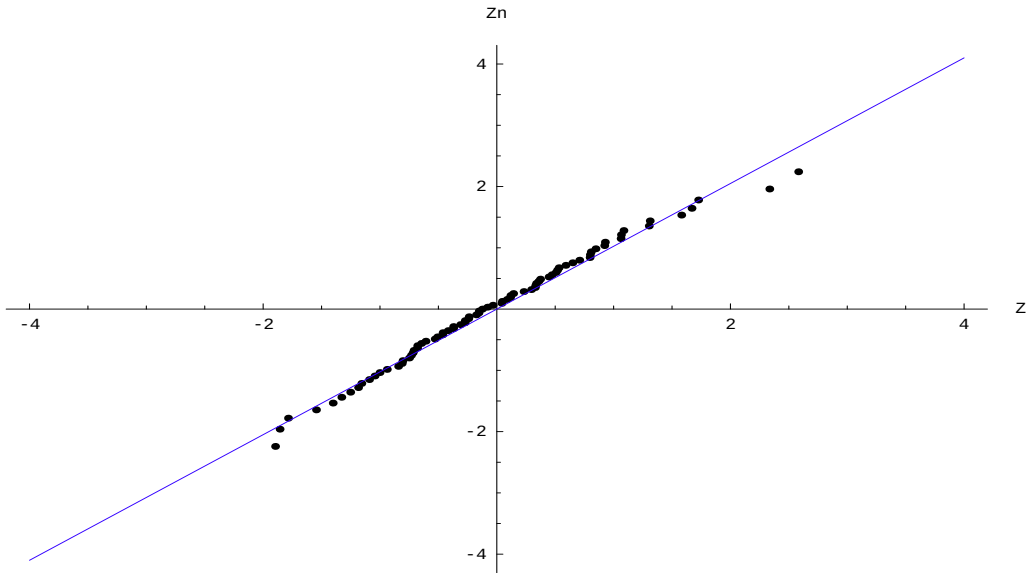


Figure 5: Normal probability paper for daily increments of VWAP logarithms: Intel on Nasdaq.

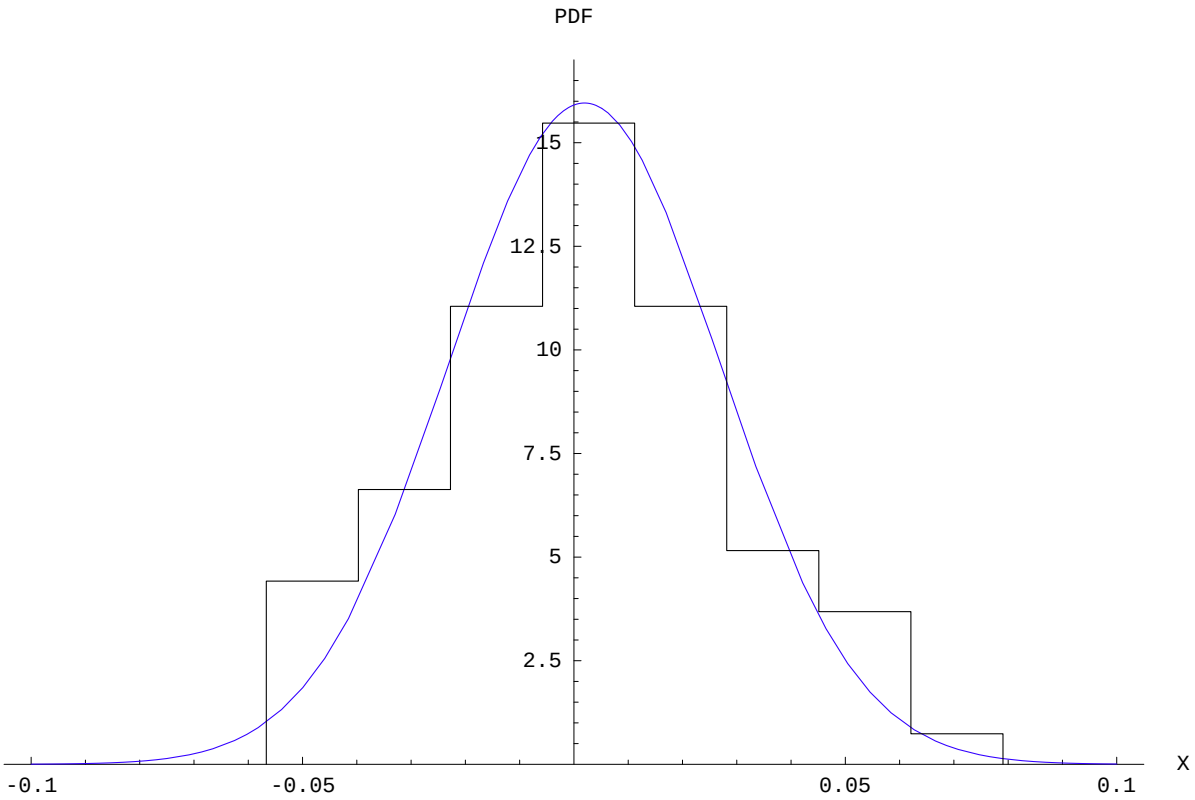


Figure 6: Empirical density function for daily increments of VWAP logarithms in comparison with normal density: Intel on Nasdaq.

Appendix 2

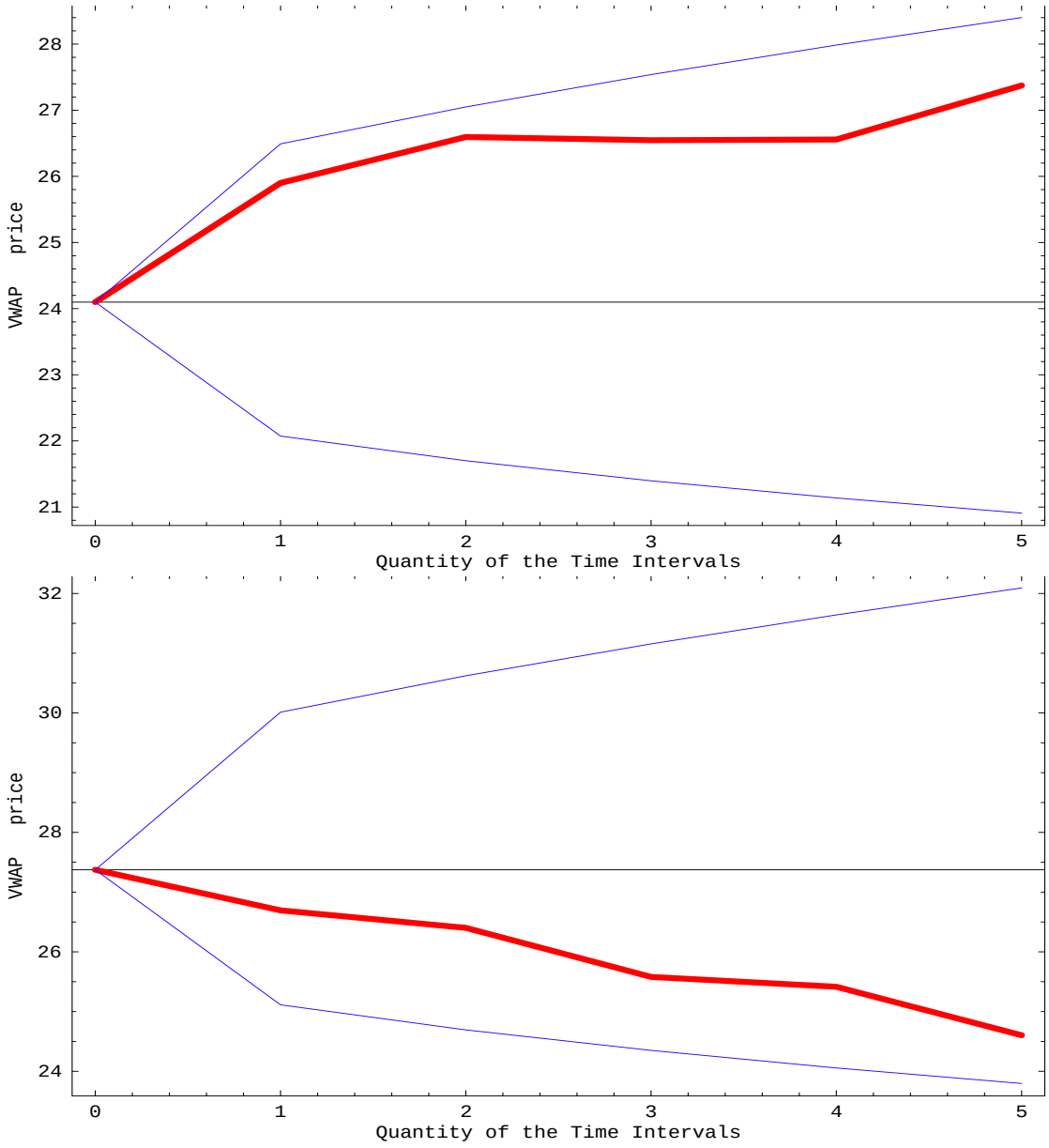


Figure 7: Actual movement of VWAP through the period of risk estimation. Lower and upper bounds of risk estimation are represented with dark blue color.

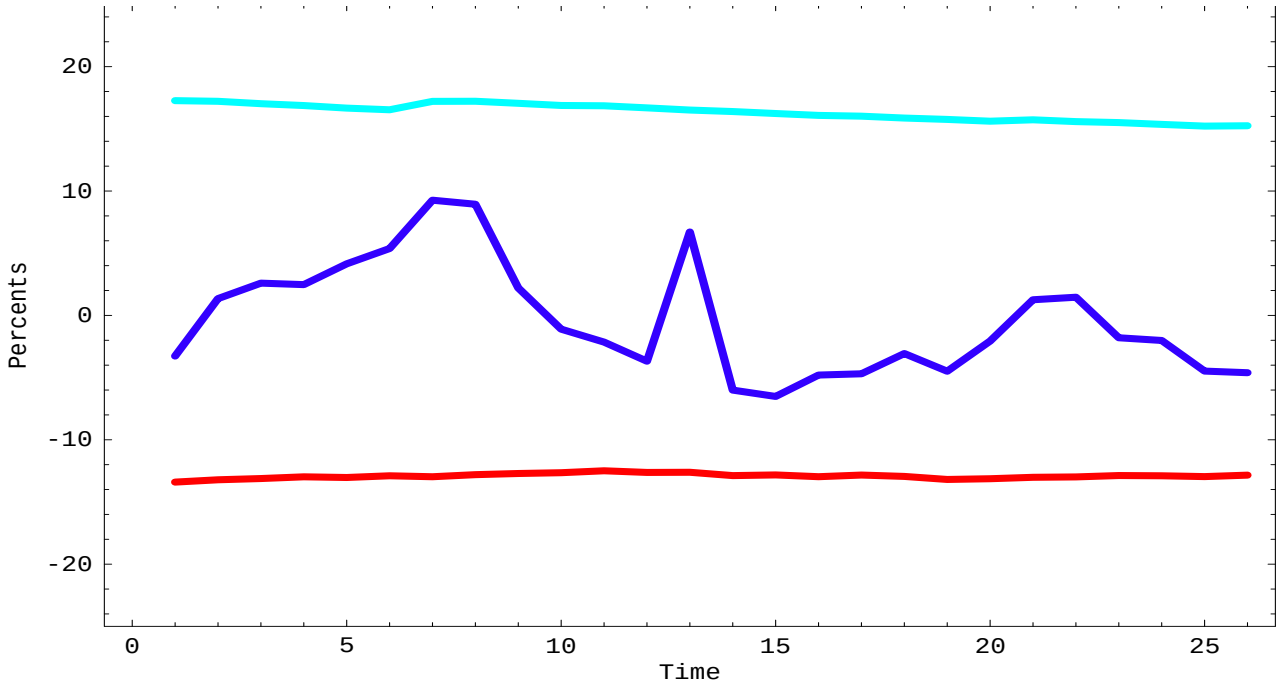


Figure 8: Risk estimation in dynamics. Each point on the time axe has three values: the lowest is the lower bound of risk estimation for this moment of time with period of estimation equal to five days; the highest point is the upper bound of the profit (with regard to the initial cost of the portfolio) which can be earned in the same period of time; the middle point shows the actual change of the portfolio cost (in percents) to the end of the estimation period. The risk was reestimated each day during a 25-day period.

Working Paper Series in Economics

recent issues

No. 52 *Mher Safarian:* On portfolio risk estimation, December 2013

No. 51 *Klaus Nehring, Marcus Pivato, Clemens Puppe:* The Condorcet set: majority voting over interconnected propositions, December 2013

No. 50 *Klaus Nehring, Marcus Pivato, Clemens Puppe:* Unanimity overruled: majority voting and the burden of history, December 2013

No. 49 *Andranik S. Tangian:* Decision making in politics and economics: 5. 2013 election to German Bundestag and direct democracy, December 2013

No. 48 *Marten Hillebrand, Tomoo Kikuchi, Masaya Sakuragawa:* Bubbles and crowding-in of capital via a savings glut, November 2013

No. 47 *Dominik Rothenhäusler, Nikolaus Schweizer, Nora Szech:* Institutions, shared guilt, and moral transgression, October 2013

No. 46 *Marten Hillebrand:* Uniqueness of Markov equilibrium in stochastic OLG models with nonclassical production, November 2012

No. 45 *Philipp Schuster and Marliese Uhrig-Homburg:* The term structure of bond market liquidity conditional on the economic environment: an analysis of government guaranteed bonds, November 2012

No. 44 *Young Shin Kim, Rosella Giacometti, Svetlozar T. Rachev, Frank J. Fabozzi, Domenico Mignacca:* Measuring financial risk and portfolio optimization with a non-Gaussian multivariate model, August 2012

No. 43 *Zuodong Lin, Svetlozar T. Rachev, Young Shin Kim, Frank J. Fabozzi:* Option pricing with regime switching tempered stable processes, August 2012

No. 42 *Siegfried K. Berninghaus, Werner Güth, Stephan Schosser:* Backward induction or forward reasoning? An experiment of stochastic alternating offer bargaining, July 2012