APPLICATION OF K-RECORDS IN THE INTERVAL ESTIMATION OF THE VALUE AT RISK MEASURE (VaR)

Marcin Dudziński, Ewa Wasilewska

Warsaw University of Life Sciences – SGGW

ABSTRACT

Value at Risk, or shorter – VaR, is a major tool used in the processes related to the risk management of banks and other monetary institutions, as well as in the tasks connected with financial supervision and scrutiny. The VaR measure may be interpreted as the minimum amount of equity that the company should own in order to be able to cover its potential losses. Although many methods leading to VaR estimation have been established so far, there is still no universal and faultless approach of VaR calculation. In our work, the method of VaR estimation consisting in determination of confidence intervals for VaR in terms of the so-called k-records has been described and used. The proposed approach is illustrated with use of an example from banking sector, concerning the stock prices of PKO BP Bank in the period between 13.01.2017 and 22.03.2018.

Key words: risk measures, value at risk, k-records, quantiles, interval estimation

INTRODUCTION

Value at Risk (VaR) is considered – both by practitioners and theoreticians of financial markets – to be one of the most popular measures of the financial risk assessment. It is mostly used in the assessment of the market risk, i.e. the risk arising from changes in values of the owned assets, but it is also applied in order to measure other kinds of risk, like the credit and operational risks [Chlebus 2014].

Value at Risk denotes the maximal financial loss that some bank or other financial institution may incur by investing in some sets of securities in the given time horizon with a certain probability called the confidence level [Jajuga 2001]. In other words, VaR is a statistical risk measure, which determines the potential limit loss in a share portfolio that may occur with a given confidence level $1 - \alpha$ in a fixed time interval ($\Delta t$) [Łukaszewski and Kostur 2013]. In practice, provided that the confidence level is $1 - \alpha$, VaR is treated as the maximal value of loss that may be incurred in the given time period in $1 - \alpha \cdot 100\%$ cases. Instead of the coefficient level $1 - \alpha$ (which is close to 1), the so-called tolerance level $\alpha$ (which is close to 0) may be considered. It is worthwhile to mention that the lower tolerance level is or the longer time horizon is, the higher VaR is [Kuziak 2003].

As a risk measure, VaR has gained in importance at the turn of the 1980s and 1990s, when the Basle Committee on Banking Supervision stated that banks should be able to cover losses on their trading portfolios over a ten-day time horizon, 99% of the time, and simultaneously it recommended to apply the VaR measure for calculating those losses. All of the corresponding regulations stem from the two Basel Accords – the so-called Basel I and Basel II Accords [Balcerowicz 2008].

Although VaR is a commonly used risk measure, its estimation still remains a vital practical challenge,
as the homogenous and reliable method of its calculation has not been worked out yet. To the major drawbacks of the so far applied methods of VaR calculation belong: the necessity of making an assumption on the distribution of the considered rates of return (it is often assumed that they have normal distribution, which is usually not in accordance with the real facts), as well as the necessity of the parameter estimations based on long series of historical data (with an assumption that the rates of return of the given financial tool are invariable, which may also differ from reality).

We should emphasize that the methods of VaR estimation, which are nowadays used by banking institutions, usually tend to overestimate its value. It leads to overestimation of capital requirements in the context of the market risk evaluation [Chlebus 2014]. Application of too conservative models for the purpose of VaR estimation results in significant reassessment of the possessed portfolio of assets, which may lead to a non-effective implementation of the entrusted capital by banks, or may even cause a considerable slowdown in economic growth [Perignon et al. 2007, cited by Chlebus 2014].

The objective of our study is to propose and describe a method of interval estimation of the Value at Risk (VaR) measure with use of the so-called k-records and its applications in the evaluation of risk in the chosen bank institution. It is a distribution-free method, i.e. it does not require any knowledge regarding the distribution of gains/losses of the given financial tool, which is very useful in analysis of time series consisting of daily stock prices or rates of return.

The presented method uses the idea of interval estimation of quantile, introduced by Ahmadi and Balakrishnan [2009] in the form of a result that is cited as Theorem 1 in the further part of our paper. To the best of our knowledge, this result has not been overly exploited in empirical studies devoted to the analysis of risk measures so far, and from this perspective we may treat it as an implementation of the already existing, but – to some extent – new method.

Additionally, among other advantages of the proposed method – except for the earlier mentioned fact that it allows to obtain distribution-free confidence intervals for VaR – we mention the fact that this method is computationally tractable and it can be easily used in practice. We hope that the presented research will make a significant contribution to further developments and discussions in the area of VaR estimation.

The introduced method of determination of the confidence intervals for VaR in terms of k-records will be shown by using data consisting of the stock exchange quotations of Poland’s largest bank PKO BP Bank. We justify such a choice of data by the fact that PKO BP Bank is one of the largest stock exchange quoted companies in Poland, as well as by the fact that its shares form a significant part of many investment portfolios of the country.

The considered dataset has been collected from the data contained on the website BiznesRadar.pl and is comprised of 300 closing stock prices of PKO BP Bank from the period between 13.01.2017 and 22.03.2018. The decisive factors for a selection of such a period were: the requirement that the size of data should be sufficiently large (to obtain sufficiently precise estimated value), and the condition that these data should be possibly most up-to-date. The assumed time horizon fulfills both of these criteria. Additionally, in order to avoid the New Year’s effect, the closing prices from the beginning of 2017 have been omitted.

All of the conducted computations were carried out with use of R, which is an open source software environment for statistical computing and graphics, freely available under the GNU General Public License.

**Review of the literature devoted to the issue of VaR estimation**

To the most commonly used methods of VaR estimation belong: the variance-covariance approach, historical simulations, the Monte Carlo method, the quantile estimation, the extreme value approach, the method applying the so-called Archimedean copulas. Among the publications in which the subject matter of VaR evaluation is considered, the works by: Jorion [1997], Pritsker [1997], Dowd [1998], Artzner et al. [1999], McNeil and Frey [2000], Grabowska [2000], Jajuga [2001], Fernandez [2003], Wüthrich [2003], Cotter and Dowd [2006], Danielsson and Zigrand [2006], Dudziński and Furmańczyk [2009], McAleer et al. [2011], Iskra [2012], Chang et al. [2013], Luszyn [2013], Furmańczyk [2016] are especially worthwhile to mention.
A comprehensive methodology of VaR calculation has been given by Jorion [1997]. An author claims there that VaR is an excellent tool that allows – with help of easily interpretable figures – to estimate a diverse number of financial risk categories, and which simultaneously provides a measure that is very supportive in making strategic financial decisions. Additionally, it has been pointed out in the paper by Dowd [1998] that VaR may find widespread applications in various areas of risk management, in particular in financial forecasting. Furthermore, the works by Danielsson and Zigrand [2006] and McAleer et al. [2011] also contain thorough reviews of the methods of VaR evaluation. In turn, Pristker [1997] regards the approach based on Monte Carlo simulations as the primary method of VaR estimation on the financial market, whereas Cotter and Dowd [2006] advise to use the Monte Carlo approach as one of the most accurate methods of the risk measures calculations based on the quantile estimation, including the estimation of VaR. On the other hand, in the paper by Arztner et al. [1999], the notion of a coherent risk measure has been introduced and in particular, the conditions under which the corresponding measure is coherent have been determined. The justification that VaR is not such a coherent measure has been given there as well. In addition, in the works just cited paper of Wüthrich [2003] to the asymptotic VaR for sums of certain classes of dependent random variables has been depicted. Moreover, in Lusztyn [2013] the historic verification (backtesting) of VaR, obtained with use of some chosen econometric models applied to some data from the Polish stock, bond, interest rates and currency markets has been carried out. Finally, quite recently Furmańczyk [2016] generalized the method of VaR estimation introduced in the already cited paper of Wüthrich [2003] to the method in which no limit theorems had to be used.

**Definition of the upper k-records**

Before we introduce the notion of the upper k-records, we need to give the definitions of the k-th order statistics and record times. Let $X_1, X_2, \ldots, X_n$ be a sequence of random variables that determines prices of the considered financial instrument in the given moments of time.

**Definition 1.** The k-th order statistic $X_{k:n} = 1, 2, \ldots, n$, related to the random variables $X_1, X_2, \ldots, X_n$, is a random variable which is a function of the random vector $(X_1, X_2, \ldots, X_n)$, defined as follows: for any elementary event $\omega$, we put a sequence of realizations:

$$X_1(\omega) = x_1, X_2(\omega) = x_2, \ldots, X_n(\omega) = x_n$$

in a non-decreasing order and obtain a sequence of the form: $x_1 \leq x_2 \leq \ldots \leq x_n$. The k-th value in this sequence $(x_{k:n})$ is a realization of the random variable $X_{k:n}$, i.e. $X_{k:n}(\omega) = x_k$.

**Definition 2.** We determine the k-th upper record times $T(n, k)$ in the following manner:

$$T^*(1, k) := k, \quad T^*(n, k) = \min \left\{ j > T^*(n-1, k) : X_j > X_{T^*(n-1, k) + 1: T^*(n-1, k)} \right\}$$

(1)

This definition allows us to introduce the definition of the so-called upper k-records $U(n, k), n \geq 1$ [see also Dzubdziela and Kopociński 1976].

**Definition 3.** The upper k-records are determined as follows:

$$U(1, k) := X_{1:k}, \quad U(n, k) = X_{T^*(n, k) - k+1: T^*(n, k)}$$

(2)

**Remark:** For $k = 1$, the upper k-record is an ordinary upper record (recall that $X_{j, k}$ is an ordinary upper record if $X_j > X_i$ for all $j > i$).

Thus, the k-records provide the generalization of ordinary records. Practical application of ordinary records is limited due to the sparsity of these records. Indeed, it may namely be stated that the expected waiting time for every ordinary record is infinite after the first one and a sample of size $n$ will only give log $n$ records. This problem does not exist if we consider a sample of k-records instead [Ahmadi and Balakrishnan 2009].
VaR measure and quantiles

Assuming the continuity of distribution of the considered data, the VaR measure can be treated as an appropriate negative quantile of this distribution. In our investigations, we do not need to know the underlying distribution function $F$ of distribution of the corresponding data, since we shall consider distribution-free confidence intervals for VaR, which means that the derivation of formulas on these intervals will not demand the knowledge of $F$.

In the present part of our paper, we wish to describe the relationship between the notion of quantile and the definition of VaR. This relationship is the starting point of our further study leading to the VaR measure estimation with use of the confidence intervals for quantiles, which are expressed in terms of $k$-records.

**Definition 4.** The quantile of rank $p \in (0; 1)$ of the probability distribution of a random variable $Y$ with a distribution function $F_Y$ is a number $q \in \mathbb{R}$ satisfying the following property:

$$P(Y < q) \leq p \leq P(Y \leq q)$$

If a random variable $Y$ in the definition above is of a continuous type, then the quantile $q$ of rank $p$ is uniquely determined and given by the relation:

$$F_Y(q) = P(Y \leq q) = P(Y < q) = p$$

Formula (4) can be equivalently rewritten as $F_Y^{-1}(p) = q$, where $F_Y^{-1}$ stands for the inverse function of the distribution function $F_Y$.

It can be checked that the concept of VaR is strictly associated with the notion of quantile. It shows the following definition.

**Definition 5.** Let $X(t)$ denote the value of financial instrument in moment $t$, which generates a random gain or loss, $\alpha$ stand for a significance level, $1 - \alpha$ denote the confidence level. The Value at Risk measure for $X(t)$ on the confidence level $1 - \alpha$ (or the tolerance – significance – level $\alpha$) is defined as [see Uniejewski 2004]:

$$VaR(1 - \alpha; t) = -\inf \{x \in R : F_{X(t)}(x) > 1 - \alpha\} =$$

$$\sup \{-x \in R : F_{X(t)}(-x) < \alpha\}$$

or equivalently as:

$$VaR(1 - \alpha; t) = \inf \{x \in R : F_{X(t)}(x) \geq \alpha\} =$$

$$\sup \{x \in R : F_{X(t)}(x) \leq 1 - \alpha\}$$

It may be easily seen that in the case when the distribution of $X(t)$ is of a continuous type, the relations above imply that

$$VaR(1 - \alpha; t) = -q_{1-\alpha}^{X(t)}$$

where: $-q_{1-\alpha}^{X(t)}$ denotes the quantile of rank $1 - \alpha$ of the distribution of $X(t)$.

Definition 5 means, for example, that if for the confidence level $1 - \alpha = 0.95$ (i.e., for the tolerance (significance) level $\alpha = 0.05$) the quantity of VaR equals PLN –10,000, then the probability of the event that in the time horizon $(0; t]$ the loss in value of the given financial asset will not exceed PLN 10,000 amounts to 0.95. In other words, we may say that with probability 0.05 the considered loss will not exceed the value of PLN 10,000.

**Application of the upper $k$-records in construction of the distribution-free confidence intervals for quantiles**

The result below illustrates a capability of application of the upper $k$-records in the determination of confidence intervals for quantiles. The obtained formula will be later used for the interval estimation of the VaR measure.

**Theorem 1.** Let $\{X_n, n \geq 1\}$ be a sequence of independent and identically distributed (iid) random variables with a distribution function $F$ (thus, it is such that $F(x) = P(X_n \leq x)$ for any $n \geq 1$). Then, an interval $(U(i, k), U(j, k))$, $1 \leq i \leq j$, where $(U(n, k))$ denotes a sequence of the upper $k$-records, is a two-sided confidence interval for the population quantile $q_p = F^{-1}(p)$,
of rank $p \in (0; 1)$, whose confidence coefficient is free of $F$ and determined by [Ahmadi and Balakrishnan 2009]:

$$
\gamma(i, j; k, p) = (1 - p)^k \sum_{i=0}^{j-1} \frac{(-k \log(1 - p))^i}{s!},
$$

where $\log(x) = \ln\{\max(x, e)\}$

The thesis of the assertion above means that:

$$
P[U(i, k) \leq q_p \leq U(j, k)] = \gamma(i, j; k, p)
$$

Let us notice that provided the rank of a quantile $(p)$ and a desired level of confidence $(\gamma_0)$ are specified, then the mentioned two-sided confidence interval for a quantile $q_p$ exists on the condition that:

$$
P[U(1, k) \leq q_p \leq U(n, k)] \geq \gamma_0
$$

for sufficiently large $n$

It may be shown that the last relation is equivalent to:

$$
\max \gamma(i, j; k, p) = 1 - (1 - p)^k \geq \gamma_0,
$$

where $\gamma(i, j; k, p)$ are such as in (8),

which implies the condition:

$$
k \geq \frac{\log(1 - \gamma_0)}{\log(1 - p)}
$$

We should emphasize that in construction of confidence intervals, we aim to find a confidence interval with a possibly minimal width. Thus, we seek such $i$, $j$, in the formula for a confidence level $(U(i, k), U(j, k))$, such that:

$$
j - i \rightarrow \min, \text{ under the condition that } \gamma(i, j; k, p) \geq \gamma_0
$$

An algorithm that enables to determine $i$, $j$ for which an interval $(U(i, k), U(j, k))$, reaches its minimal width is presented in the paper by Ahmadi and Balakrishnan [2009].

In Table 1, we placed optimal values of $k$, $m$ and $(i, j)$, obtained on the basis of the mentioned algorithm, for $\gamma_0 = 0.95$ and some selected levels of $p$.

<table>
<thead>
<tr>
<th>$p$</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_{opt}$</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>$m_{opt}$</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>$(i, j)_{opt}$</td>
<td>(1.9)</td>
<td>(1.9)</td>
<td>(1.9)</td>
<td>(1.9)</td>
<td>(1.10)</td>
</tr>
</tbody>
</table>

Source: Ahmadi and Balakrishnan [2009].

**EMPIRICAL STUDY**

The basic inconvenience of our empirical research was the fact that our data, comprising of the share prices of PKO BP Bank, were not stochastically independent, whereas Theorem 1 can be used only if the considered observations are the realizations of independent random variables. In order to overcome this problem, we transformed a sequence of stock prices from our sample into a sequence of their logarithmic rates of return by acting as follows:

$$
\ln \frac{X_t}{X(t-1)}. \quad t = 2, \ldots, 300
$$

The corresponding time series, obtained from our original dataset through transformation (14), is depicted in the figure.

In order to check, whether our new sample of logarithmic rates of return comes from a sequence of independent random variables, we carried out the so-called Runs test. We conducted the Runs test with help of the R environment, in particular with use of the runs.test function, available after loading the tseries package of R. The $p$-value of the corresponding Runs test was equal to 0.4862 and exceeded the significance level 0.05. This means that, there is no reason to reject the null hypothesis that the logarithmic rates of return from our sample are the realizations of independent random variables.
As, in view of the conducted Runs test, we may assume that the calculated logarithmic rates of return of the share prices of PKO BP Bank in the period 13.01.2017–22.03.2018 are the realizations of some sequence of independent random variables, we may apply Theorem 1 in order to determine the confidence intervals for quantiles of the distribution of logarithmic rates of return of the considered stock prices. For this purpose, we provided a table containing the values of the upper $k$-records $(U(n, k))$ for our sample of logarithmic rates of return and some selected $k \in \{2, 3, 4, 5\}$.

Based on the values of the upper $k$-records listed in Table 2, we obtained the realizations of confidence intervals for the quantiles of rank $p$ of the logarithmic rates of return of the share prices of PKO BP Bank in the selected time horizon.

In particular, it may be seen from Table 3 below that the quantile of rate 0.9 of the considered logarithmic rates of return is with probability 0.951 a certain value from the interval $(-0.033918; 0.027558)$1. It means that with probability 0.9, on average in 95 out of 100 days, the value of daily logarithmic rate of return belongs to the interval $(-0.033918; 0.027558)$.

Thus, we may write that with probability 0.951:

$$F_{Y\mid X}^{-1}(0.9) \in (-0.033918; 0.027558),$$

where $Y_t = \ln \left( \frac{X(t)}{X(t-1)} \right)$, $t = 2, ..., 300$.  (15)

Due to (15), we have with probability 0.951 that:

$$P \left( \ln \left( \frac{X(t)}{X(t-1)} \right) \leq q \right) = 0.9$$

for some $q \in (-0.033918; 0.027558)$, $t = 2, ..., 300$, and consequently that:

$$P \left( \frac{X(t)}{X(t-1)} \leq q \right) = 0.9$$

for some $q \in \left[ \exp(-0.033918); \exp(0.027558) \right]$.  (17)

---

1 Logarithmic rate of return is a unitless quantity and it can take an arbitrary real value. Furthermore, it is positive if there is a growth in share prices and negative if otherwise.
Hence, with probability 0.951:

\[ P \left( \frac{X(t)}{X(t-1)} \leq q \right) = P \left[ X(t) \leq X(t-1) q \right] = 0.9 \]  

for some \( q \in (0.967; 1.028) \)  

Putting into a place of random variable \( X(t-1) \) in (18) its realization \( x(t-1) \) (which is an appropriate historical value from our data), we may write that with probability 0.951:

\[ P \left[ X(t) \leq x(t-1) q \right] = 0.9 \]  

for some \( q \in (0.967; 1.028) \)
Thus, we obtain that with probability 0.951 the value \( x(t-1)q \), where \( q \) denotes some number from the interval (0.967; 1.028), is the quantile of rank 0.9 of the distribution of \( X(t) \). The relations in (7) and (19) imply that \( \text{VaR}(t) \), i.e. a desired VaR for the share prices of PKO BP Bank in the considered period, is with probability 0.951 given by:

\[
\text{VaR}(t) = -x(t-1)q \\
\text{for some } q \in (0.967; 1.028)
\]  

(20)

The equality above implies that with confidence 95.1%, we may claim that probability of the event that the loss in stock prices in a time horizon \((0; t]\) will not exceed \( x(t-1)q \) (PLN), where \( q \) stands for a (unitless) number from the interval (0.967; 1.028), is equal to 0.9.

In case of the considered share prices of PKO BP Bank, we have that \( x(t-1) = \text{PLN 29.96 for } t = 300 \). Thus, we may state, with confidence 95.1%, that with probability 0.9 the value of potential loss will not exceed some number between PLN 28.97 and 30.80. In other words, on average within 95 out of 100 days, probability of the event that the value of potential loss will not exceed some number between PLN 28.97 and 30.80 is equal to 0.9. Relating the level of loss to an average price of the given financial tool, it is possible to quantify the risk level of investor. In our study, an average share price of PKO BP Bank in the assumed time period amounted to PLN 37.33 and its median to PLN 35.85. It means that, on average, during 95 out of 100 days, we may say that, with probability 0.9, the corresponding loss will not exceed a percentage from the range between 77.6 and 82.5% of the prices arithmetic mean, as well as it will not exceed a percentage from the range between 80.8 and 85.9% of the prices median. Such a rather large percentage contribution of the VaR measure both with relation to the mean price and with relation to its median gives an evidence that the financial risk of investing in the PKO BP Bank stocks is relatively high.

It is important to underline that the proposed approach is a method of interval estimation, which allows to establish lower and upper limits of the potential loss with a fixed confidence coefficient. This is not possible in case of the point estimation, where we only determine one specified value of estimator.

**SUMMARY**

Value at Risk is widely recommended by various authorities of financial control as a tool of measurement of the potential financial loss that may be incurred by banks and other financial institutions due to their operational activities. The correct estimation of the VaR measure is relevant, as its precise evaluation may have a substantial impact on the process of efficient use of the gathered capital and may even influence solvency of banks and other financial institutions.

In our work, we proposed the method of interval estimation of the VaR measure that used the notion of the upper \( k \)-records. Its application resulted in the conclusion that with confidence level 95.1% (i.e. on average, in 95 out of 100 days) and with probability 0.9 the potential loss of investing in the considered stocks will not exceed some value from the interval (PLN 28.97; 30.80). With regard to an average share price and its median, it may be interpreted that the corresponding loss will not exceed the range (77.6; 82.5%) of the average stock price and will not exceed the range (80.8; 85.9%) of the stock price median. The conducted research indicates that the risk of investing in the given shares is relatively high.

Although we have not conducted any comparisons between the proposed method of estimation and other approaches, it is worthwhile to mention that the presented method of VaR evaluation provides a non-parametric, distribution-free approach, which in addition does not require neither imposing on any normality or consistency assumptions nor employing big datasets. Furthermore, this method is easy for calculation and interpretation. Thus, we obtain attractive and plausible estimation procedure that stands out from many other existing methods of risk estimation in the sense of its feasibility and computational tractability.
Obviously, it would be interesting to compare risk obtained with use of the proposed approach for the PKO BP Bank with identically calculated risks for some other banking institutions, as well as to compare the presented method of risk estimation with some other distribution-free methods of its evaluation. However, such a comparative analysis goes beyond the scope of our paper. Simultaneously, it should be emphasized that the mentioned comparisons are necessary in the further research devoted to banking risk management.

REFERENCES


ZASTOSOWANIE $k$-TYCH REKORDÓW W ESTYMACJI PRZEDZIAŁOWEJ WARTOŚCI ZAGROŻONEJ RYZYKIEM (VaR)

STRESZCZENIE

Wartość zagrożona ryzykiem, lub krócej – wartość zagrożona (Value at Risk, VaR), jest podstawowym narzędziem wykorzystywanym zarówno w procesach zarządzania ryzykiem banków i innych instytucji finansowych, jak i w zadaniach związanych z jego nadzorem i kontrolą. Miarę VaR można interpretować jako minimalną wartość kapitału, którą dany podmiot powinien posiadać w formie zabezpieczenia na poczet ewentualnych strat. Choć metody szacowania wartości zagrożonej jest wiele, dotychczas nie wskazano uniwersalnej i pozbawionej wad metody jej kalkulacji. W niniejszej pracy zaproponowano oraz przedstawiono możliwość oszacowania miary VaR za pomocą przedziałów ufności wyrażonych w terminach tzw. $k$-tych rekordów. Zaproponowaną metodę zilustrowano na przykładzie Banku PKO BP. Wyznaczono realizacje przedziałów ufności dla wartości zagrożonej na podstawie danych dotyczących notowań cen akcji banku w okresie od 13.01.2017 r. do 22.03.2018 r.

Słowa kluczowe: miary ryzyka, wartość zagrożona, $k$-te rekordy, kwantyle, estymacja przedziałowa