Extreme Value Approach for estimating Value at Risk metrics with respect to Basel II

Lubor Homolka

Abstract— A large number of articles have been written about methods designed to assess easily interpretable value reflecting risk taken from a (not exclusively financial) process. In the financial environment, prevailing concepts include Value at Risk (VaR) and its derivatives, such as Conditional Value at Risk. The purpose of this paper is to describe appropriateness of the VaR metrics under Basel II legislative framework and to stress VaR estimation techniques. A relatively new approach titled Extreme Value Theory and methods allowed by Basel II are compared on illustrative example of a skewed distribution with presence of outliers. Our findings suggest alternative methods assess higher VaR than the classical ones (historical simulation, mean-variance model and Monte Carlo simulation) and are more precise in terms of variance.

Keywords—Value at Risk, Extreme Value Theory, Historical simulation, Bootstrap, Basel II

I. INTRODUCTION

Banks and insurance institutions hold decisive position in financial system as well as in overall macroeconomic environment. The industry faces both internal (operational) and external (market and credit) risks. To make the systems stable and safe, banks have to respect strict regulations. Almost 24-year development of banking legislative under the Bank for International Settlement (starting with Basel I) has evolved into general framework grounded in a three-pillar system – capital framework; risk management and supervision pillar and market discipline pillar. Because of the nature of such highly stochastic environment, contemporary regulation philosophy is closely tied to advanced mathematical and statistical procedures. One of the most widely used (and criticized) approach used in required capital determination within Basel II [1] legislative is a family of Value at Risk (VaR) methods.

Critique of the VaR methodology is twofold. The first argues the methodological properties itself and following misleading interpretation. The second highlights technical problems such as parameter estimation and distribution approximation. This paper focuses mainly on the VaR methodology from the latter perspective and extends author’s previous work [2] for other risk metrics and their EVT counterparts. We suggest Extreme Value Theory concept to be more appropriate method for modelling skewed and non-normal tailed distribution of losses in a sense of safety (overestimating risk rather than underestimating) than methods allowed under Basel II.

First section introduces VaR methodology and its connection to regulatory framework and economic capital. A description of historical simulation, Monte Carlo simulation and EVT approach follows. Finally, comparable results are presented.

II. METHODS

A. Value at Risk

Let the X be a random variable (e.g. returns or losses) with the distribution function $F_X$. The VaR $\alpha$ at probability level $\alpha$ in $0 < \alpha < 1$ is its $\alpha$ quantile. Formally, this can be written as

$$ F_X^{-1}(\alpha) := \inf \{ x \in \mathbb{R} : F_X(x) \geq \alpha \} $$

(1)

VaR represents overall portfolio’s absolute risk measure, generally defined as (when the losses are with positive sign)

$$ \alpha = \mathbb{P}(x \leq \text{VaR}) = \int_{\text{VaR}}^{\infty} f(x) \, dx = F(\text{VaR}) $$

(2)

where $f(x)$ is (usually the empirical) probability density function of a variable (i.e., losses over some time period from portfolio’s value changes) and $F(x)$ its corresponding cumulative distributive function, thus loss value higher or equal than $x$ will occur only with probability $\alpha$. The key issue is how the distribution function should be assessed with respect to a fitting accuracy and allowance for computation.

VaR suffers from some conceptual deficiencies; one of them is subadditivity problem (detailed discussion in [3]). Subadditivity arises when the risk of the portfolio (of X and Y) is estimated by overall VaR. It can be shown

Lubor Homolka, Tomas Bata University in Zlin, FAMAE CZ, Department of Statistics and Quantitative Methods Department of Statistics and Quantitative Methods (e-mail: homolka@fame.utb.cz).
that $\text{VaR}(X + Y) \leq \text{VaR}(X) + \text{VaR}(Y)$. This implies that VaR is not coherent risk metric. Although numerous articles were written on VaR (for a historical review see [4]), some topics remain uncovered. In probability theory, two fundamental approaches coexist each of which tackles the problem in a different way.

While the majority prefers the classical frequentist approach, the Bayesian one is more promising in incorporating genuine expert information [5] and handling non-linear systems [6].

The scope of this article is concerned about the general techniques which differ from the Bayesian in that sense they do not allow impingement of expert’s own experiences thus this kind of analysis is omitted here.

Crucial point in the VaR estimation process lies in identifying appropriate distribution which generates the data. To create it, several approaches were designed. Starting from the simplest historical simulation form through mean-variance to simulation models, all of them suffer from the main idea of VaR – estimating overall distribution which may underestimate the real risk hidden in the tail(s). Several metrics stem from standard VaR, such a limited VaR [7]

B. Conditional VaR

The problem VaR deals with is the unexpressed uncertainty about the losses beyond the cut-off VaR value. The remedy to this situation can be found in conditional VaR, in literature also named Expected Shortfall (ES). This risk indicator is defined as:

$$\text{CVaR}_\alpha(X) = \frac{1}{1-\alpha} \int_0^{1-\alpha} \text{VaR}_\varepsilon(X) \, d\varepsilon$$  \hspace{1cm} (3)

where $\varepsilon$ stands for the tail probability of VaR which satisfies the $\alpha \leq \varepsilon$ thus VaR is always smaller than CVaR. [8]

Moreover, CVaR is a coherent risk metrics, as is proved in McNeil [3].

C. Classical approaches to determining VaR

Under Basel II three approaches to VaR estimation for regulatory capital purposes are allowed:

Variance–Covariance methods

Based on the portfolio volatility, descriptive statistics (such a central moments) are derived to be used as parameters in parametric probability model. The simplest model refers to normal distribution but more proficient methods were introduced to handle non-normal distributions as well (variance is not even needed to be computed). Although many computationally extensive methods (kernel estimators) were introduced and are available in statistical packages such R, SPSS, Matlab, we restrict our paper to mathematical estimator using Cornish-Fisher Expansion (CF). This expansion starts with assumption about underlying distribution $z = N(0,1)$ but ends with transformed non-normal $z_\varepsilon$ distribution

$$z_\varepsilon = z + \left(\frac{\varepsilon^2 - 1}{6} b_1 + \frac{\varepsilon^3 - 3\varepsilon}{24} b_2 - \frac{(2\varepsilon^3 - 5\varepsilon)}{36} S^2\right)$$  \hspace{1cm} (4)

$S$ represents skewness and $K$ kurtosis of variable $z$. The assumption about the underlying distribution of VaR is crucial for interval estimate. If we assume VaR follows normal or student distribution with known parameters $\mu$ and $\sigma^2$, the confidence interval (CI) is derived as

$$P \left( b\left(\frac{N - 1}{\chi^2_{\nu, 0.5}}\right) < \text{VaR} < b\left(\frac{N - 1}{\chi^2_{\nu, 0.5}}\right) \right) = 1 - \alpha$$  \hspace{1cm} (5)

Where $b$ is quantile of normal or student distribution, $N$ is a sample size and $\chi^2$ is sample variance. If the parameters of location and variance are not known, it is reasonable to use simulated results from random sampling of the parameters $\mu$ and $\sigma^2$ domains. The previous formulae can be rearranged to the form of new variable $S$, where $\bar{t}$ is sample mean value.

$$S = b\left(\frac{N - 1}{\chi^2_{\nu}}\right) - N\left(\frac{(N - 1) \bar{t}^2}{n \chi^2_{\nu}}\right)$$  \hspace{1cm} (6)

After random sampling from $S$, the appropriate quantiles are computed.

Historical simulation

This approach presumes existence of some underlying repeating structure such as time-invariant probabilistic model. Only the corresponding quantile of loss distribution is computed. This procedure can be used under Basel II only if data sample is sufficiently large and proved over specified time horizon.

According to work of Pétrignon [9] (sample size of 50 large international commercial banks in 2005) historical simulation is the most frequently used method with the share at least of 47.4%. Monte Carlo methods follow with 14% and only 3.5% banks use “other” methods which combines the latter approaches; 35.1% of asked financial institutes didn’t provide an answer.

Monte Carlo simulation

Monte Carlo procedures (including bootstrap) are sampling procedures which draw random samples from the initial sample to estimate value of predefined quantity. We propose bootstrap method to be suitable for deriving VaR because of its properties (sampling with replacement, computationally efficient, no strict assumptions). After bootstrapped values are known, graphical analysis providing useful information about
quantity variation and sensitivity with respect to initial settings follows.

D. Extreme Value Theory

Observed data and their approximation by any classical distribution may underestimate the risk. Although the largest part of a distribution can fit perfectly, tail values may cause real damage. Financial data tend to have “fat” tails and thus need to be treated carefully. Extreme Value Theory deals with this extreme part of distribution as with another property of the phenomena being investigated.

If any observation in a dataset would be considered as a iid sample taken from its random variable, than the overall distribution can be divided into two parts – predictable behaving part with high density and tail part described by tail quantile function. Similarly to central limit theorem when some distribution converges to normal distribution, we focus on convergence of distribution \( X_{n,n} = \max\{X_1, X_2, \ldots, X_n\} \) to some distribution \( G \).

\[
\lim_{n \to \infty} \left\{ \frac{\left( X_{n,n} - b_n \right)}{a_n} \right\} \leq x \rightarrow G(x)
\]

(7)

Variables \( a \) and \( b \) (sequence of \( n \) numbers, \( n \) represents number of samples) standardize the initial distribution. Based on these numbers we estimate the distribution of \( X \). This is only one part of the problem where distribution is in literature named domain of attraction. The second step involves finding the \( G(x) \) distribution. Fisher-Tippet theorem [1] shows that \( G(X) \) underlying distribution converge to one of three distributions of the family extreme value distribution.

\[
G_\gamma (x) = \exp \left[ - \left( 1 + \gamma x \right)^{-1/\gamma} \right]
\]

(8)

In the expression \((1 + \gamma x) > 0\) and the Extreme Value Index (EVI) \( \gamma \in R \). If the \( \gamma < 0 \) we are talking about Frechet-Pareto distribution, whose domain distributions include Burro distribution, log-gamma and Generalized Pareto distribution. In case of \( \gamma > 0 \) the extreme distribution is Weibull with its domain distributions Reversed Burro or Beta. The last one Gumball distribution \( \gamma = 0 \) comprises exponential, logistic or log-normal distribution. Only possible distributions which satisfy the limiting assumptions are the extreme distributions. According to Beirlant [10] a general limit distribution combining the previously mentioned exists.

The EVI index value in (8) can be estimated from general limit distribution, which is named GEV – Generalized Value distribution (GEV) where \( \gamma \) is a tail index, \( \mu \) location and \( \sigma \) scale parameter. Larger \( \gamma \) produces fatter distribution. Parameters \( \sigma, \mu, \gamma \) can be estimated by maximum likelihood (ML) method, method of probability weighted means, and as described in Lye [12] by Bayes estimator too. Two essential models are recognized from the point of determining extreme values. So far, we were interested in maximal values of \( X_n \). This approach, considered as the simpler one, Block maxima, estimates the extremes in fixed time period or within \( n \) logically justified blocks of variables. For more detailed discussion see [13,14]

Computationally-intensive peak over threshold (POT) method is the second approach. Extreme values are those exceeding a (sufficiently high) threshold value \( u \), which is constant over the sample. Limit distribution used in POT method is Generalised Pareto Distribution (GPD).

\[
\text{GPD}\beta (x) = \begin{cases} 1 - \left( 1 + \xi x / \beta \right)^{-1/\xi} & \xi \neq 0 \\ 1 - e^{-x/\beta} & \xi = 0 \end{cases}
\]

(10)

When \( \xi > 0 \), GPD becomes ordinary Pareto distribution, also known as distribution of large losses in actuarial statistics. In case of \( \xi = 0 \), GPD is identical to exponential distribution while with \( \xi < 0 \), GPD is known as Pareto type II distribution with short tails [13,14] and the parameters usually estimated by ML methods.

According to Coles [15] standard approach is to set the threshold level as lower as the model provides reasonable approximation. Threshold value can be estimated by informed guess, but several mathematical approaches were developed. First one is based on semiparametric statistics and is known as Hill estimator [16]. Hill estimator estimates parameter \( \xi \) as a slope of an exponential Q-Q plot, which should properly rescaled tail values follow. When the points form a convex shape compared to the line of expected exponential distribution quantities, the distribution is thin tailed, if concave, heavy tailed. When \( \xi \) is zero, than the distribution is exponential, otherwise \( \xi \) is parameter of GPD. Hill estimate suffers from tendency to provide biased results when small sample is analysed. Remedy to this can be found in application wavelet analysis and kernel estimate of the tail distribution [17]. The second one is Mean Excess Function \( M(t) \) which computes mean value of values exceeding threshold.

\[
M(t) = E \left[ (X - t) | X > u \right]
\]

(11)
Within the “normal” part of distribution and moving to extremes (to the larger numbers) mean value should rise steadily because extreme values does not play such an important part. Fluctuating of M function suggest change in data structure thus extreme part break.

**EVT and metrics**

When the assumption of EVT are met (the threshold \( u \) is sufficiently high and the data belongs to the maximum domain of attraction), the ratio between number of values larger than \( u \) to the sample size \( \frac{N_u}{N} \) is directly computed and distribution parameters \( \beta \) and \( \xi \) are estimated. Those values allows for direct VaR computation using the equation [18]:

\[
\text{VaR}_\alpha(X) = u + \frac{\beta}{\xi} \left[ \frac{N_u}{N} (1 - \alpha) \right]^{-\frac{\xi}{\beta}} - 1
\]  

(12)

In case of CVaR we rearrange the previous as follows:

\[
\text{ES}_\alpha(X) = \frac{1}{1 - \alpha} \int \text{VaR}_\alpha(X) dx = \frac{\text{VaR}_\alpha(X)}{1 - \beta} + \frac{\beta - \xi u}{1 - \frac{\xi}{\beta}}
\]

(13)

When using Block maxima approach (underlying GEV) explicit form of VaR is [18]:

\[
\text{VaR}_\alpha = \left\{ \begin{array}{ll}
\mu - \sigma \left( 1 - \left( -\log \left( 1 - \frac{1}{p} \right) \right)^{-\gamma} \right) & \text{if } \gamma \neq 0 \\
\mu - \sigma \log \left( -\log \left( 1 - \frac{1}{p} \right) \right) & \text{if } \gamma = 0
\end{array} \right.
\]

(14)

The ES for block maxima is derived by re-parameterizing of the GPD estimation [19].

**E. Field of application**

It can be distinguished between two kinds of capital. *Regulatory capital (RC)* is the minimal amount of risk capital to be hold to meet regulatory rules and application guidance. In 1999, Basel Committee for Banking Supervision released the New Basel Capital accord. According to this legislature (also known as Basel II, [1]) VaR value can be used (after certain requirements are met) to determine capital adequacy matching to unexpected losses. Figure 1 shows the application of VaR as a complementary part of (i.e credit) losses to Expected Loss (EL), which is an amount of regular losses and in the long term period rapidly predictable, and to Unexpected Loss (UL). UL summarizes the irregular and highly improbable losses. UL can be estimated implicitly as a product of stress probability of default, stress loss given default and stress exposure at default for each transaction [20]. If the bank is allowed to use Internal ratings-based (IRB) approach for credit risk and to create stochastic credit risk portfolio model, some quantity of losses with small and fixed probability \( \alpha \) can be estimated in order to derive RC (for discussion about benefits resulting from IRB see [21]). RC is then set as a difference between \( \text{VaR}(\alpha) \) and UL.

![Fig 1 – Value at Risk and regulatory capital [22].](Image)

The previous part described UL only. In the process of estimating EL, VaR is involved as well. Exposures to credit risk\(^2\) (\( E^* \)) are calculated as

\[
E^* = \max \left\{ 0, \left[ \sum E_i - \sum C_i \right] + \text{VaR} \right\}
\]

(15)

where \( E_i \) is current value of exposure \( i \) and \( C_i \) the received collateral value. The VaR is also promoted as a market risk metric since the Market Risk Amendment was released in 1996, where the banks are encouraged to measure risk by internal models rather than using external agencies services.

Value at Risk serves as a metric for underestimating of risk which bank faced to in predefined time period. The process of stressing actually needed capital with the predicted is called backtesting. Backtesting under Basel II is focused on number of exceedances over predicted VaR not on total volume of these exceedances, which results in lower statistical power [23,24]. The observed number of exceedances then affects the regulatory capital for market risk by adding number from 0 to 1(in case of 10 exceedances) to multiplicative factor, which is set for the three months period as follows:

\[
CR_t = \text{mf}_t \cdot \text{VaR}_\alpha(\alpha)
\]

(16)

where \( \text{mf}_t \) is multiplicative factor for time \( t \), which is set to 250 days. The ground level of multiplicative (scaling) factor is three. [25]

**Economic capital** is a result of shareholder’s trade-off between solvency and profitability. The capital size optimization process must reflect not only maximization of performance indicators, but also all foreseeable risks accruing from specific portfolio structure, long term planning objectives and capital’s signal function of stability.

Although both of them measure very similar-meaning  

\(^2\) Credit risk is considered for demonstration purposes only, although Basel (2006) allows using VaR for market and operational risks, as well.
variable, final figures very often differs.

According to Saita [26], performance indicators such as RAROC are usually based on economic capital while the amount of regulatory is omitted in the performance indicators.

Several approaches for economic capital stipulation were introduced (i.e., CreditMetrics or Moody's KMV). Basel II motivates bank’s management to measure different sources of risks by implementing these internal models to be more accurate in the risk-evaluation processes. If banks use internal-rating-based approach these metrics are tightened with regulatory capital. Own risk measures are transformed into risk weights specified by the Basel Committee. [27]

### III. RESULTS

Assume a portfolio consisting of X entities. Histogram on the Fig.1 provides information about distribution of losses (with positive sign) in the last 200 days and estimated probability density function.

#### A. Historical simulation and variance-covariance estimate

To set Value at Risk according to the standard methods, we proceed with estimating density function of the gathered data through rescaled histogram. This allows computing historical VaR (HS). After the shape of distribution is estimated using Cornish-Fisher, kernel estimation component risk contribution with no weighting preferences (authors Epperlein and Smillie, implementation in [28]) is provided as well.

Table 1 – Classic methods (Own processing, [28])

<table>
<thead>
<tr>
<th>Historical simulation</th>
<th>Cornish-Fisher approximation</th>
<th>Normal distribution approximation</th>
<th>Kernel estimator</th>
</tr>
</thead>
<tbody>
<tr>
<td>41.59</td>
<td>41.27</td>
<td>39.14</td>
<td>41.59</td>
</tr>
</tbody>
</table>

Using normal distribution for estimating VaR at 5% level showed the lowest value. Difference between HS, kernel and CF estimates is only 0.3 in favour of HS (from the point of conservativeness).

#### B. Bootstrapped estimate

If the previous approach fails, sampling methods known as Monte Carlo may be used instead. In the paper, original data were replicated using bootstrapped sampling procedure of 1000 replicates. To make data replicable, appropriate distribution and parameterisation have to be set. Gamma distribution was used and parameters were estimated using maximum likelihood estimator (MLE). After the distribution is estimated, data can be generated randomly from it.

![Fig. 3 – Bootstrapped VaR growth and corresponding 0.95 confidence intervals [2,13].](image)

Fig. 3 shows estimated VaR for \(1 - \alpha \in (80,99)\) at \(x\) axis and corresponding 95% bootstrapped confidence interval. This interval adds information about the volatility. With growing accent on precision (lowering \(\alpha\)), VaR risk grows but the interval width has the same tendency as can be seen at Fig. 4.

![Fig. 4 – Volatility vs. precision [2]](image)

Values constituting the confidence intervals and mean values are presented at Table 2.

Table 2 – Bootstrap confidence intervals [2]

<table>
<thead>
<tr>
<th>VaR((\alpha))</th>
<th>0.9</th>
<th>0.95</th>
<th>0.99</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower CI</td>
<td>31.835</td>
<td>35.271</td>
<td>41.700</td>
</tr>
<tr>
<td>Mean</td>
<td>35.052</td>
<td>40.825</td>
<td>46.025</td>
</tr>
<tr>
<td>Upper CI</td>
<td>39.392</td>
<td>42.687</td>
<td>56.957</td>
</tr>
</tbody>
</table>

![Fig. 2 – Density estimate (Own processing, [28])](image)
C. EVT estimate

If our data is periodical variable, we would choose as extreme values those which are the highest on every subperiod. Our data do not follow any similar pattern, thus we will not be concerned about Block maxima approach.

Using the statistically naïve POT method and selecting the 90% largest value as a threshold, our cut-off value would be 35.39. Graphical analysis for appropriate threshold selection can be provided through mean-excess plot. When the data’s trend remains stable, no extreme values are present. After data become more jittered, they should be considered as inconsistent with the preceding data. Fig. 5 shows more dispersion when $x$ exceeds 40. [2]

![Fig 5 – Mean-excess plot [2, 30]](image)

After the threshold is set, the remaining data can be approximated using a distribution described in equation 4. Estimate was provided using Maximum Likelihood function.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\xi$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimate</td>
<td>-0.1012</td>
<td>6.6263</td>
</tr>
<tr>
<td>Standard error</td>
<td>0.3069</td>
<td>2.6516</td>
</tr>
</tbody>
</table>

Estimated value for the shape parameter $\xi$ is small relative to its standard error. Because the parameter determines type of definition used for approximation, the information should be treated carefully. QQplot below supports our choice of distribution and parameterization.

![Fig 6 – QQ plot for GPD [30,2]](image)

In Fig. 7 the estimate is denoted as a solid line. Vertical dashed line shows .95 quantile considered at VaR at $\alpha = 0.05$. Horizontal line defines confidence interval on a level scaled on the second y axis.

![Fig 7 – VaR(0.95) Confidence interval [2,31]](image)

At the predefined level the confidence interval is asymmetric which is to be expected due to underlying distribution’s skewness.

<table>
<thead>
<tr>
<th>Table 4 EVT estimates [13]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower CI</td>
</tr>
<tr>
<td>40.593</td>
</tr>
</tbody>
</table>

Results of ES comparison show the same pattern, EVT returns more conservative results. Figure 8 provides comparison of both estimates; chain-dotted line at 47.46 represents ES calculated by historical simulation (Cornish-Fisher Expansion results in 46.60) and EVT estimate 47.67. The point estimates are almost the same, but the confidence interval in case of EVT estimate is highly biased to large losses.

<table>
<thead>
<tr>
<th>Table 5 ES estimates (Own processing)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower CI</td>
</tr>
<tr>
<td>EVT</td>
</tr>
<tr>
<td>His. simul.</td>
</tr>
</tbody>
</table>
In 1996, Basel Committee released an amendment to the Capital Accord which allowed banks to use VaR variance-covariance method, historical and Monte Carlo simulations. EVT was neither allowed nor mentioned. [1] This paper compared VaR quantified by all previously mentioned methods with following results. The lowest VaR estimate was provided by approximation using normal distribution (39.14). Direct quantile estimate within historical simulation and kernel procedures follows (41.58). Other estimates were accompanied by confidence intervals in which the worst case scenario (upper CI value) was considered as VaR value. Using Monte Carlo (bootstrapped) method VaR is smaller (42.687, overall interval width = 7.416) than the EVT’s (44.607, width = 4.01).

These findings suggest that officially approved methods are inappropriate (when the loss distribution has similar positive skew shape as ours) in terms of identifying potential risk value. The size of VaR consequently influences the capital requirement which allows banks to hold less capital reserves when standard methods are used. [2]

In May 2012 consultative document [32], fundamental review of the trading book were published by the Bank for International Settlement. This document extends impact of Basel 2.5 rules. This European legislative primarily aims at banks’ trading books and suggests to complete rejection of simple VaR in favour of ES and other robust risk metrics. This step can be interpreted as conceptual framework abandonment which might be partly caused by inappropriate computing methods used.

ACKNOWLEDGMENT

This paper was supported by Project No. IGA/FaME/2012/12: Optimization of internal rating model parameters of commercial banks in the small and medium enterprises.

REFERENCES


