

Value at Risk and its statistical estimation based on extreme value theory: a review.

Zhouman Wu. *

Instructor's Name: Wenqing Hu. †

MATH 4997W, Spring 2016, University of Minnesota.

1 Introduction.

Risk can be defined as the volatility of unexpected outcomes. Firms are exposed to various types of risks. Because of the potential for damage, risk should be monitored carefully. One of the measurements to measure and limit potential loss related to financial risk management is called Value at Risk (or VaR in short). Modern financial markets adopt three different major types of risk: credit risk, operational risk and market risk. Concerning the market risk, VaR is a point estimate of potential financial loss. Extension of VaR to the other two types of risks are also possible. As mentioned in Tsay's article [2], "VaR is a single estimate of the amount by which an institution's position in a risk category could decline due to general market movements during a given holding period." VaR can be used by financial institutions to assess risks or by a regulatory committee to set margin requirements. The original purpose of this measure is to ensure the survival of financial institutions during catastrophic events, such as the stock market crash on Wall Street in October 1987, or most recently, the bankruptcy of Lehman Brothers. Although mathematically the same notion, the definition of VaR can be interpreted in two different ways, corresponding to viewpoints from financial institution and from regulatory committee. From the viewpoint of financial institutions, VaR is defined as the maximal loss of a financial position subject to a given probability during the given period. From the viewpoint of regulatory committee, VaR is defined as the minimal loss under extreme market movements. In this paper, we will introduce a formal definition of VAR. Then we will turn to several methods available for market

*School of Mathematics, University of Minnesota, Twin Cities, Minneapolis, USA. Email: wuxx0687@umn.edu

†School of Mathematics, University of Minnesota, Twin Cities, Minneapolis, USA. Email: huxxx758@umn.edu

risks including RiskMetrics, Empirical Quantile and Extreme Value Theory (EVT) with R demonstration.

2 Mathematical definition of VaR.

We can now formally define the Value at Risk (VaR) of a portfolio with all the requisite tools in place in the previous section. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space. Suppose we are interested in the risk of a financial position for the next ℓ period at time t . Let $\Delta V(\ell)$ be a random variable and it is equal to the change in value of asset from the starting time t to ending time $t + \ell$. We also denote the cumulative distribution function of $\Delta V(\ell)$ by $F_\ell(\bullet)$. We can define the VaR of a long position over the horizontal target ℓ with probability p as

$$p = \mathbf{P}(\Delta V(\ell) \leq \text{VaR}) = F_\ell(\text{VaR}) . \quad (1)$$

The VaR defined in equation (1) usually assumes long position. Alternatively, we can interpret VaR from the viewpoint of a short position as

$$1 - p = \mathbf{P}(\Delta V(\ell) \geq \text{VaR}) = 1 - \mathbf{P}(\Delta V(\ell) \leq \text{VaR}) = 1 - F_\ell(\text{VaR}) . \quad (2)$$

The difference between these two equations is that they focus on the different sides of the tail behavior of the $F_\ell(\text{VaR})$. For a holder of a long position, his/her loss occurs when $\Delta V(\ell) < 0$. On the contrast, for a holder of a short position, his/her loss occurs when the value of asset increases, that is, $\Delta V(\ell) > 0$. Therefore we are focusing on the left tail of the $F_\ell(\text{VaR})$ for a long position while focusing on the right tail of $F_\ell(\text{VaR})$ for a shot position.

Calculating VaR usually involves the following factors ([2]):

1. The probability of interest p .
2. The horizontal time ℓ .
3. The frequency of the data.
4. The CDF $F_\ell(\text{VaR})$ or its quantiles.
5. The market-to-market value of the position.

Now we will turn to practical methods in computing VAR.

3 Risk Metrics.

One of the methodologies that is introduced to VaR calculation is called *Risk-Metrics*. This method is developed by J.P. Morgan ([3]). It assumes that the contin-

ously compounded daily return r_t follows a conditional normal distribution, that is $r_t|\mathcal{F}_{t-1} \sim \mathcal{N}(\mu_t, \sigma_t^2)$. Here \mathcal{F}_{t-1} is our information set at time $t-1$, μ_t is the conditional mean and σ_t^2 is the conditional variance. Also, this method assumes the following simple IGARCH(1,1) model when the two quantities evolving over time:

$$\mu_t = 0, \quad \sigma_t^2 = \beta\sigma_{t-1}^2 + (1-\beta)r_{t-1}^2, \quad 0 < \beta < 1.$$

Let me first give a brief introduction of this special random-walk IGARCH(1,1) model. IGARCH(1,1) model is defined as the unit-root GARCH models. One of the significant features of this model is that the impact of the past squared shocks on a_t^2 is persistent [1, page 122]. IGARCH(1,1) model can be written as

$$a_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = \alpha_0 + \beta_1 \sigma_{t-1}^2 + (1-\beta_1) a_{t-1}^2.$$

Here ε_t is a sequence of independent and identically distributed (i.i.d.) $\mathcal{N}(0, 1)$ random variables. In the RiskMetrics method, the model is a special case of IGARCH(1,1) model, that is, $\alpha_0 = 0$. In this method, we define the geometric rate of return r_t in terms of the logarithm of price ratio.

$$r_t = \ln(P_t + D_t) - \ln(P_{t-1})$$

For simplicity, we will assume the dividends D_t are zero, which can also be thought of as we reinvested all dividends. Thus

$$r_t = \ln(P_t) - \ln(P_{t-1})$$

where $r_t = \sigma_t \varepsilon_t$ as defined in the IGARCH(1,1) process above, $\ln(P_t)$ is the logarithm of the daily price, and the typical value of β is 0.94.

Suppose we are interested in the VaR of a portfolio under RiskMetrics for a k -period horizon. First we denote the k -horizon return from time $t+1$ to $t+k$ as $r_t[k] = r_{t+1} + \dots + r_{t+k-1} + r_{t+k}$. Under the special IGARCH(1,1) model we defined above, we can compute the variance $\sigma_t^2[k]$ as follows:

$$\sigma_t^2[k] = \mathbf{Var}(r_t[k]|\mathcal{F}_t) = \sum_{i=1}^k \mathbf{Var}(r_{t+i}|\mathcal{F}_t).$$

We can rewrite the volatility equation of the IGARCH(1,1) model by applying $r_{t-1} = \sigma_{t-1} \varepsilon_{t-1}$ as

$$\sigma_t^2 = \beta\sigma_{t-1}^2 + (1-\beta)\sigma_{t-1}^2 \varepsilon_{t-1}^2.$$

After rearranging, we get the following equation for all t :

$$\sigma_t^2 = \sigma_{t-1}^2 + (1 - \beta)\sigma_{t-1}^2(\varepsilon_{t-1}^2 - 1) .$$

Apply this volatility equation to time $t + i$ for $i = 2, \dots, k$, we have

$$\sigma_{t+i}^2 = \sigma_{t+i-1}^2 + (1 - \beta)\sigma_{t+i-1}^2(\varepsilon_{t+i-1}^2 - 1) .$$

Since $\varepsilon_t \sim \mathcal{N}(0, 1)$, $\mathbf{E}(\varepsilon_{t+i-1}^2 - 1 | \mathcal{F}_t) = 0$ for $i \geq 2$, after eliminating $(1 - \beta)\sigma_{t+i-1}^2(\varepsilon_{t+i-1}^2 - 1)$ from the previous equation, we have

$$\mathbf{E}(\sigma_{t+i}^2 | \mathcal{F}_t) = \mathbf{E}(\sigma_{t+i-1}^2 | \mathcal{F}_t) \text{ for } i = 2, \dots, k .$$

Since we obtained $\mathbf{Var}(r_{t+i} | \mathcal{F}_t) = \mathbf{E}(\sigma_{t+i}^2 | \mathcal{F}_t)$ recursively, and by the volatility equation we see that $\sigma_{t+1}^2 = \beta\sigma_t^2 + (1 - \beta)r_t^2$, we can apply these to the conditional expectation equation above as

$$\mathbf{E}(\sigma_{t+i}^2 | \mathcal{F}_t) = \mathbf{E}(\sigma_{t+i-1}^2 | \mathcal{F}_t) = \mathbf{E}(\sigma_{t+i-2}^2 | \mathcal{F}_t) = \dots = \mathbf{E}(\sigma_{t+1}^2 | \mathcal{F}_t) \text{ for } i \geq 1 .$$

Thus we obtained

$$\mathbf{Var}(r_{t+i} | \mathcal{F}_t) = \mathbf{E}(\sigma_{t+1}^2 | \mathcal{F}_t) = \sigma_{t+1}^2 \text{ for } i \geq 1 .$$

Hence the variance of $r_t[k] | \mathcal{F}_t$ is $\sigma_t^2[k] = k\sigma_{t+1}^2$. And the results above show that

$$r_t[k] | \mathcal{F}_t \sim \mathcal{N}(0, k\sigma_{t+1}^2) .$$

We can say that the conditional variance $\sigma_t^2[k]$ of $r_t[k]$ is proportional to the time horizon k and the conditional standard deviation

$$\sqrt{\sigma_t^2[k]} = \sqrt{k\sigma_{t+1}^2} = \sqrt{k}(\sigma_{t+1}) .$$

So under RiskMetrics, a k -period horizon can be calculate as

$$\text{VaR}(k) = \sqrt{k}\text{VaR}$$

where VaR is measured in daily.

R demonstration for Risk Metrics.

I use the data of daily log returns of IBM stock from 2001 to 2010 : d-ibm-0110.txt. ([1])

```

> getwd()
> setwd("/Users/wuzhouman1/Desktop")
> data = read.table("d-ibm-0110.txt",header=T)
> head(data)
> ibm = log(data$return+1)
> neg_ibm = 100*-ibm
### RiskMetrics ###
> source("Igarch.R")
> m1 = Igarch(xt)
Estimates: 0.942857
Maximized log-likelihood: 4501.827

Coefficient(s):
      Estimate Std. Error t value Pr(>|t|)
beta 0.942857 0.007172 131.464 < 2.22e-16 ***

> names(m1)
[1] "par" "volatility"
> beta = m1$par
> v1 = m1$volatility
> length(neg_ibm)
[1] 2515
### volatility prediction ###
> vt = beta*v1[2515]^2+(1-beta)*neg_ibm[2515]^2
> source("RMeasure.R")
> RMeasure(0,sqrt(vt))

Risk Measures for selected probabilities:
      prob      VaR      ES
[1,] 0.950 1.173279 1.471339
[2,] 0.990 1.659391 1.901106
[3,] 0.999 2.204272 2.401756
>

```

4 Quantile Estimation.

We can use quantiles of the distribution to measure the downside risks. By definition, *quantiles* are the cutoff values such that the area to their left or right presents a given probability.

4.1 Empirical Quantile.

We can use empirical quantile of return of an asset to calculate the VAR. Suppose returns of an asset are measured in a fixed time and are denoted by r_t . Assume the return sequence $\{r^{(1)}, \dots, r^{(n)}\}$ follows order statistic that are arranged in increasing order:

$$r^{(1)} \leq r^{(2)} \leq \dots \leq r^{(n)} ,$$

where $r^{(1)}$ is the minimum return and $r^{(n)}$ is the maximum return in this sample return sequence. Also, assume that $\{r_1, \dots, r_n\}$ are i.i.d random variables with a continuous distribution with pdf $f(x)$ and CDF $F(x)$. Then we can construct x_p as the p^{th} quantile of the CDF $F(x)$ and $f(x_p) \neq 0$. In this way, the order statistic $r^{(g)}$ follows

$$r^{(g)} \sim \mathcal{N} \left(x_p, \frac{(1-p)p}{[f(x_p)]^2 n} \right) ,$$

where $0 \leq p \leq 1$ and $g = np$. Then we can use $r^{(g)}$ to estimate x_p where $x_p = F^{-1}(p)$. For a given probability p , if $np = g$ is an integer, then the empirical quantile is $r^{(g)}$. If np is not an integer, we need to find two neighboring integers g_1 and g_2 such that $g_1 \leq np \leq g_2$. By interpolation, we obtain the empirical quantile

$$\hat{x}_p = \frac{p_2 - p}{p_2 - p_1} r^{(g_1)} + \frac{p - p_1}{p_2 - p_1} r^{(g_2)} .$$

VaR obtained by the empirical quantile can be seen as a lower bound for the actual VaR.

4.2 Quantile Regression.

We can use quantile regression to investigate VAR models, since it can be viewed as a conditional quantile function of the return sequence $\{r_1, \dots, r_n\}$. That is, we actually consider the distribution function $r_{t+1}|F_t$.

4.3 R Demonstration for Empirical Quantile.

We use the same data as in Risk Metrics method.

```

> data = read.table("d-ibm-0110.txt",header=T)
> ibm = log(data$return+1)
> neg_ibm = -ibm*100
### Empirical quantile ###
> quantile(neg_ibm,c(.95,.99))
      95%      99%
2.653783 5.013151
> length(neg_ibm)
[1] 2515
### Empirical expected shortfall ###
> idx = c(1:2515)[neg_ibm > 2.653783]
> ES = mean (neg_ibm[idx])
> ES
[1] 3.994857
> idx = c(1:2515)[neg_ibm > 5.013151]
> ES = mean(neg_ibm[idx])
> ES
[1] 6.074108
>

```

5 Extreme Value Theory.

To improve the estimation of extreme quantiles, we discuss the Extreme Value Theory (EVT) in this section. Extreme Value theory focuses on the tail behavior of the return of an asset. First, we define the returns as r_t that measured in a fixed time. And denote the return sequence by $\{r^{(1)} \leq \dots \leq r^{(n)}\}$ that follows order statistic. In this collection, $r^{(1)}$ is the minimum return that is relevant to a long position; $r^{(n)}$ is the maximum return that relevant to a short position. Specifically, $r^{(1)} = \min_{1 \leq i \leq n} \{r_i\}$ and $r^{(n)} = \max_{1 \leq i \leq n} \{r_i\}$. By doing a simple sign change, we obtain the minimum return of an asset $r^{(1)}$ from $r^{(n)}$:

$$r^{(1)} = - \max_{1 \leq i \leq n} \{-r_i\}.$$

5.1 Basic Concepts.

Assume the return sequence $\{r_t\}_{t \geq 1}$ is independent and identically distributed (i.i.d.) with a common cumulative distribution function (CDF) $F(x)$. Denote the CDF

of the maximum return $r^{(n)}$ by $F_{n,n}(x)$. By definition of maximum and independence, we have

$$F_{n,n}(x) = \mathbf{P}(r^{(n)} \leq x) = [F(x)]^n.$$

The extreme value theory is concerned with finding the position sequence $\{\beta_n\}$ and scale position sequence $\{\alpha_n\}$ such that the distribution of this properly normalized minimum $r_n^* \equiv \frac{r^{(n)} - \beta_n}{\alpha_n}$ converges to $F_*(x)$ as n goes to infinity.

$$F_*(x) = \begin{cases} \exp[-(-(1 + \xi x)^{\frac{1}{\xi}})] & \text{if } \xi \neq 0 ; \\ \exp[-\exp(-x)] & \text{if } \xi = 0 . \end{cases}$$

If $\xi < 0$, the above is true for $x < 1/\xi$; and if $\xi > 0$, the above is true for $x > 1/\xi$. Here ξ is the *shape parameter*, and $\alpha = -1/\xi$ is the *tail index* of the distribution.

This limiting distribution is so called *generalized extreme value distribution* (GEV). The GEV includes three extreme value distributions as special cases, the Gumbel distribution ($\xi = 0$), the Fréchet distribution ($\xi > 0$), and the Weibull distribution ($\xi < 0$).

Case I: $\xi = 0$, the Gumbel family which consists of thin-tailed distributions.

$$F_*(x) = \exp[-\exp(-x)] \quad \text{if } -\infty \leq x \leq \infty.$$

Case II: $\xi > 0$, the Fréchet family with CDF as

$$F_*(x) = \begin{cases} \exp[-(-(1 + \xi x)^{\frac{1}{\xi}})] & \text{if } x > -1/\xi; \\ 0 & \text{otherwise.} \end{cases}$$

Case III: $\xi < 0$, the Weibull family with CDF as

$$F_*(x) = \begin{cases} \exp[-(-(1 + \xi x)^{\frac{1}{\xi}})] & \text{if } x < -1/\xi; \\ 1 & \text{otherwise.} \end{cases}$$

From here we obtain the probability density function (pdf) of the generalized limiting distribution $f_*(x)$:

$$f_*(x) = \begin{cases} (1 + kx)^{1/k-1} \exp[-(-(1 + kx)^{\frac{1}{k}})] & \text{if } k \neq 0; \\ 0 & \text{otherwise.} \end{cases}$$

5.2 Statistical Methodology.

After introducing basic EVT distribution, we can now compute the quantile with location parameter β_n , scale parameter α_n and shape parameter ξ . Suppose there are N data points, we divide the sample into n non-overlapping subsamples.

$$\{r_1, \dots, r_n | r_{n+1}, \dots, r_{2n} | r_{2n+1}, \dots, r_{3n} | \dots | r_{(g-1)n+1}, \dots, r_{ng}\}.$$

where

$$r_{n,i} = \max_{1 \leq j \leq n} \{r_{(i-1)n+j}\}$$

for $i = 1, \dots, g$.

The idea of EVT is to find the minimum of each subgroup. Then we use these maxima to estimate the three parameters we mentioned above. Several estimation methods are available. Here, we use maximum likelihood estimates method and regression method.

Maximum likelihood

Suppose the sub-period maxima $r_{n,i}$ of n sub-groups follows EVT distribution with parameter (ξ, β, α) . We can calculate the pdf of $r_{n,i}$:

$$f_*(r_{n,i}) = \begin{cases} \frac{1}{\alpha_n} \left[\left(1 + \frac{\xi_n(r_{n,i} - \beta_n)}{\alpha_n}\right)^{-\frac{1+\xi_n}{\xi_n}} \right] \exp\left[\left(1 + \frac{\xi_n(r_{n,i} - \beta_n)}{\alpha_n}\right)^{-\frac{1}{\xi_n}}\right] & \text{if } \xi_n \neq 0 ; \\ \frac{1}{\alpha_n} \exp\left[\left(-\frac{r_{n,i} - \beta_n}{\alpha_n}\right) - \exp\left(-\frac{r_{n,i} - \beta_n}{\alpha_n}\right)\right] & \text{if } \xi_n = 0 . \end{cases}$$

To calculate VaR for r_t , we use a two-step procedure because data is divided into subgroups.

$$\text{VaR} = \begin{cases} \beta_n - \frac{\alpha_n}{-\xi_n} \{1 - [-n \ln(1-p)]^{-\xi_n}\} & \text{if } \xi_n \neq 0 ; \\ \beta_n + \alpha_n \ln[-[-n \ln(1-p)]] & \text{if } \xi_n = 0 . \end{cases}$$

Regression method

Now we will introduce another method called regression method that use least squares estimates. It is consistent but less efficient than the likelihood estimates.

Suppose $r_{n,i}$ follows EVT distribution and has properties of order statistics, we denote the order statistics of the sub-period maxima $\{r_{n,i}\}_{i=1}^g$ as $r_{n(1)} \leq r_{n(2)} \leq \dots \leq r_{n(g)}$.

First we have the expectation of $r_{n(i)}$:

$$\mathbf{E}[F_*[r_{n(i)}]] = \frac{i}{g+1}, \quad i = 1, \dots, g .$$

We can transform

$$F_* = \exp[-(-(1 + \xi x)^{\frac{1}{\xi}})]$$

for $\xi \neq 0$ into

$$F_*[r_{n(i)}] = \exp\left[-\left(1 + \frac{\xi_n(r_{n,i} - \beta_n)}{\alpha_n}\right)^{-\frac{1}{\xi_n}}\right] .$$

By doing simple logarithm transformation, we then obtain

$$\frac{i}{1+g} = \exp\left[-\left(1 + \frac{\xi_n(r_{n,i} - \beta_n)}{\alpha_n}\right)^{-\frac{1}{\xi_n}}\right], \quad i = 1, \dots, g .$$

Finally, when $\xi_n = 0$, we can reduce the equation to

$$\ln\left[-\ln \frac{i}{1+g}\right] = -\frac{1}{\alpha_n} r_{n(i)} + \frac{\beta_n}{\alpha_n} + e_i, \quad i = 1, \dots, g.$$

5.3 R demonstration for Extreme Value Approach.

```
### Traditional EVT Approach ###
> library(evir)
> m2 = gev(neg_ibm,block=21)
> m2
$n.all
[1] 2515
$n
[1] 120
$data
 [1] 4.0335654 4.6038703 6.9818569 4.6158081 2.9868660 1.8749682\
 [7] 5.1245927 3.6876663 3.3919828 2.3115108 2.0342513 2.4236342\
[13] 4.8790114 4.7451211 10.6674712 7.3298915 3.1040820 6.3058018\
[19] 4.9832257 4.1026182 7.0740111 5.3694069 3.2279420 3.3725360\
[25] 5.6782014 2.8647448 3.1710505 1.7130900 4.1022015 2.0702832\
[31] 4.0106625 1.2247698 2.1467793 3.8022791 1.2162666 1.2266932\
[37] 2.0414975 1.9473382 2.0080266 1.8147677 1.4965425 2.0119037\
[43] 1.8123238 1.7842230 1.6585787 1.1930891 1.6205604 1.2909976\
[49] 1.9149181 1.5308581 1.0340277 8.6620544 2.2660828 2.3848119\
[55] 1.1244989 1.7883969 1.4590932 1.2093837 0.6296783 2.6523663\
[61] 2.1040814 1.4236865 1.5996261 1.8530637 2.7150255 2.1617994\
[67] 1.6564436 1.1022525 2.1935841 0.8557511 2.0563997 0.9094227\
[73] 3.3537135 3.0913950 1.4987757 2.4177945 3.2943735 2.0471111\
[79] 3.4377182 1.7755704 1.6698650 3.2461210 5.6767196 3.6600708\
[85] 2.7332143 2.5186535 1.4527008 2.3683247 2.7861553 2.8494129\
[91] 1.7898222 2.8629954 6.0170475 6.1018324 5.9323344 2.7901656\
[97] 3.5234509 5.1061742 4.8006072 1.9594729 2.0101692 3.0112876\
[103] 1.2639544 1.4527008 1.4558462 5.0808149 1.2491697 1.2677015\
[109] 2.9428811 0.8985247 0.7097125 2.2935008 4.0058744 3.0624170\
[115] 2.5279860 2.0077205 0.9940241 3.4183660 1.5317720 0.7572600\

$block
[1] 21
```

```

$par.ests
      xi      sigma      mu
0.251353 1.028910 1.965850

$par.ses
      xi      sigma      mu
0.08847742 0.09013351 0.10932034

$varcov
      [,1]      [,2]      [,3]
[1,] 0.007828254 -0.001080741 -0.003453668
[2,] -0.001080741 0.008124049 0.006145413
[3,] -0.003453668 0.006145413 0.011950936

$converged
[1] 0

$llh.final
[1] 210.1579

attr("class")
[1] "gev"
> plot(m2)
Make a plot selection (or 0 to exit):
1: plot: Scatterplot of Residuals
2: plot: QQplot of Residuals
Selection: 1

Make a plot selection (or 0 to exit):
1: plot: Scatterplot of Residuals
2: plot: QQplot of Residuals
Selection: 2

> source("evtVAR.R")
> evtVaR(0.251353, 1.028910, 1.965850, 21, 0.01)
[1] 3.924483
>

```

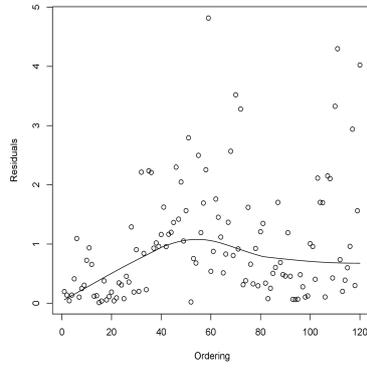


Fig. 1: Residual plots from fitting GEV distribution to daily negative IBM log returns for data from Jan 2, 2001, to December 31, 2010. Original Data.

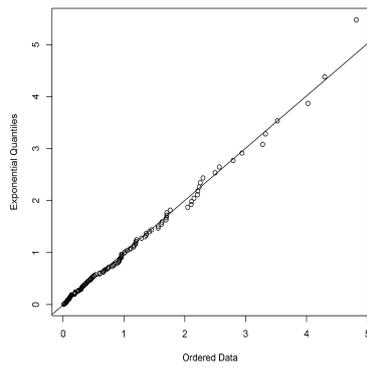


Fig. 2: Residual plots from fitting GEV distribution to daily negative IBM log returns for data from Jan 2, 2001, to December 31, 2010. After regression.

6 Summary.

In this review article we discussed the mathematical definition of Value at Risk (VaR) and its statistical estimation methodologies. Three different methods are discussed about the estimation of Value at Risk: the RiskMetrics, the quantile estimation and the extreme value theory. In particular, the extreme value theory provides us with a powerful technique in the estimation of VaR. We used the R software and some sample data to test the methodologies and we provide the results.

References

- [1] Tsay, R.S., *Analysis of financial time series*, Second edition, Wiley Series in Probability and Statistics, Wiley–Interscience, 2005.
- [2] Tsay, R.S., Extreme values and their applications in finance. online paper.
- [3] Longestaey, J., More, L., Introduction to RiskMetricsTM, 4th edition. *Morgan Guaranty Trust Company*, New York, 1995.