Mixture normal Value at Risk models of some European market portfolios[1]

Jiří Valecký[2]

Abstract

The assumption of normal probability distribution belongs to the biggest imperfections of estimating Value at Risk. In point of fact, the returns of financial time series are rather distributed leptokurtic than normally. Moreover, the empirical distributions are often skewed. In these cases, the assumption of normal distribution results in over- or underestimation of VaR especially when the quantiles are very high/low. Therefore it is necessary to put emphasis on respecting the leptokurtic and skewed return distribution. In this paper, we interpret the one out of the method how to estimate VaR with respect to the empirical distributions. We describe the analytical solution of VaR under mixture and Markov-Switching normal distribution condition and we compare the estimates according to both approaches. We also present the estimation method of distribution parameters. Thus, we briefly describe and derive the maximum likelihood method based on the iterative EM-algorithm. Using the four selected European market portfolios (DAX, FTSE 100, PX and ATX) we estimate the parametric and Monte-Carlo VaR and we give evidence that the estimates are very inaccurate when the normal distribution is assumed.

Keywords

Value at Risk, mixture normal distribution, maximum likelihood, EM-algorithm, Monte-Carlo simulation, Markov-Switching.

JEL Classification: C1, C13, C52, G1

1 Introduction

The Value at Risk (VaR) is a risk measure representing value of loss at some significance level. There is also a possibility to look on VaR like on the methodology of managing risks. Value at Risk models find the utilization in modelling credit, operational and market risk which contains various risks such as equity risk, forex risk, commodity risk and option risk. A good introduction to Value at Risk methodology is provided by the technical document from (Morgan, 1996) or by many follow-up books such as (Danielsson, 2007), (Jorion, 2007), (Alexander, 2008) and others.

The methods for estimating VaR are generally three. It can be estimated analytically (variance-covariance method), on the basis of historical simulation (bootstrap method) or Monte-Carlo simulation. The historical simulation is mostly preferable because it does not impose any assumption on probability distribution and it is relative easy to use as far as we believe that the future will be the same like the history. On the contrary, to estimate the analytical VaR (or parametric VaR) and MC VaR, it is necessary to impose the assumption of probability distribution, mostly normal. But then, the normal distribution is unsuitable in the

[1] This paper is solved within the support by the Grant Agency of the Czech Republic, GACR: P403/12/P692.
[2] Ing. Jiří Valecký, Ph.D., VŠB – Technical University of Ostrava, Faculty of Economics, Department of Finance, Sokolska tr. 33, 701 21 Ostrava 1; e-mail: jiri.valecky@vsb.cz.

most cases because the empirical distributions of financial time series are leptokurtic and often skewed. Here, the normality assumption results in the overestimation or underestimation of VaR at the given significance level.

Therefore, there are several approaches how to respect the leptokurtic and skewed empirical distributions which result from the existence of outliers and extreme observations. For instance, (Duffie & Pan, 2001) used the jump-diffusion models, (Giannopoulos, 2003) calculated the VaR strictly by the filtered historical simulation to obtain the empirical distribution. Next approach employs general Lévy processes, for instance (Tichý, 2009), (Tichý, 2010), and some authors make experiment with fuzzy-stochastic approach, (Zmeškal, 2005a), (Zmeškal, 2005b). Another approach respecting the empirical distribution is to impose the assumption of another probability distribution instead of normal, for example Student distribution. Nevertheless also in this case, the Student distribution does not suffice and therefore some mixture distributions are considered. This method presumes that the distribution is compound from many individual distributions (components) which parameters are estimated. The most preferable estimator is maximum likelihood estimator which is based on the iterative EM algorithm, see (Dempster, et al., 1977) or (McLachlan & Krishnan, 1997) (1997) for further details.

Thus, we focus on the estimation of VaR respecting the empirical distributions of four selected European market portfolios (represented by four market indexes) in that way that we presume mixture normal and Markov-Switching normal distributions and thus we highlight the overestimation/underestimation of VaR under normality condition. Respecting the fact that the probability assumption is not needed if the historical VaR is estimated, we focus on parametric and Monte Carlo VaR only. In the purpose of obtaining all parameter estimates of individual distribution, we employ the EM-algorithm.

The paper is organized as follows: first, parametric and Monte-Carlo VaR are described under normal, mixture and Markov-Switching (MS) normal distribution condition. Next section is devoted to description and derivation of EM-algorithm and then in following section we estimate the VaR analytically and on the basis of Monte-Carlo simulation. The results conditioned by all mentioned probability distributions are compared. The last section concludes the paper.

2 Value at Risk

As it was mentioned in introduction part, Value at Risk can be calculated analytically, on the basis historical simulation or Monte-Carlo simulation. We pay attention here to the analytical solution and MC simulation only.

Value at Risk can be formulated mathematically as a value of loss at given confidence level and it equals to lower quantile of probability distribution of random variable, thus

$$P\left(\tilde{X} < x_\alpha\right) = \alpha \qquad (1)$$

where $\tilde{X}_t \overset{iid}{\sim} N\left(\mu, \sigma^2\right)$ and $x_\alpha = -VaR$.

2.1 Analytical VaR

Analytical solution of VaR consists in normalization of Equation (1). After this operation, we can derive the equation in the form of

$$P\left(\tilde{X} < x_\alpha\right) = P\left(\frac{\tilde{X}-\mu}{\sigma} < \frac{x_\alpha-\mu}{\sigma}\right) = P\left(Z < \frac{x_\alpha-\mu}{\sigma}\right) = \alpha, \qquad (2)$$

where μ and σ is mean value and standard deviation and $Z \sim N(0,1)$. Respecting the equality

$$\frac{x_\alpha - \mu}{\sigma} = \Phi^{-1}(\alpha), \tag{3}$$

where Φ^{-1} is inverse standard normal distribution function, and respecting the fact that the equality $-\Phi^{-1}(\alpha) = \Phi^{-1}(1-\alpha)$ holds for the symmetric distribution the VaR can be estimated analytically by the following formula

$$VaR = -x_\alpha = \Phi^{-1}(1-\alpha)\sigma - \mu\,. \tag{4}$$

The analytical formula under condition of mixture normal distribution can be derived along the similar lines. The idea of the normal mixture probability distribution consists in the assumption that the distribution of random variable is compound from the K normal distributions (components). It is observable generally that the combination of two (high and low volatility period) or three components (bear, bull market and stagnant market) is sufficient in financial applications.

Let´s consider time series following the k-th regime with unconditional probability π_k, then the normal mixture distribution is formulated as weighted sum of each component, thus

$$G(x) = \sum_{k=1}^{K} \pi_k F_k\left(x; \mu_k, \sigma_k^2\right), \tag{5}$$

where $G(x)$ is a mixture distribution function, $F\left(x; \mu_k, \sigma_k^2\right)$ is distribution function of k-th regime (component) with the mean value μ_k and variance σ_k^2 and where K is number of components. The Value at Risk can be derived from the formula

$$P\left(\tilde{X} < x_\alpha\right) = G(x_\alpha) = \sum_{k=1}^{K} \pi_k F_k\left(x_\alpha; \mu_k, \sigma_k^2\right) = \sum_{k=1}^{K} \pi_k P\left(Z < \frac{x_\alpha - \mu_k}{\sigma_k}\right) = \alpha\,. \tag{6}$$

Under the assumption that the relationships of (3) and (4) hold, the Value at Risk is found by setting $VaR = -x_\alpha$ and solving the Equation (6) by goal programming.

Finally, let’s have a state variable s_n following the first-order Markov chain with transition probability $P\left(s_n = k \middle| s_{n-1} = j\right) = p_{jk}$ representing the probability of switching from state j at time $n-1$ into state k at t. Considering predicted conditional probability $P\left(s_n = k \middle| \Omega_{n-1}\right)$, given the information set Ω_{n-1} at time $n-1$, instead of π_k in the Equation (6), we obtain the Markov-switching normal VaR.

2.2 Monte-Carlo VaR

We do not focus on the details of generating random and pseudo-random numbers, we refer to some detailed literature such as (Glasserman, 2004) or (Chan & Wong, 2006). We concentrate here only on the Monte-Carlo VaR estimates.

Both simulated risk measures are estimated empirically. The basis consists in the simulation of distribution for the portfolio´s return. The MC-VaR is calculated as -1 times the α-quantile of this distribution. The portfolio´s return distributed normally can be obtained easily by transforming generated uniform variable

$$x = \Phi^{-1}(u)\sigma + \mu\,, \tag{7}$$

where u is uniform random variable and Φ is standard normal distribution function, μ and σ is mean value and standard deviation of required normal distribution.

The general mixture normal distribution compound from K components (normal distributions) in the form of Equation (5) is generated in two stages: (1) firstly, a random variable from categorical distribution of size K and probabilities π_k for $k = 1, \ldots, K$ is generated; (2) then, the uniform random variable of size $\pi_k N$ for all k is drawn and then transformed into the individual normal distribution F_k by

$$x = F_k^{-1}(u) = \Phi^{-1}(u)\sigma_k + \mu_k \,, \tag{8}$$

where u is uniform random variable, Φ and F_k is standard normal and normal distribution function of *k*-th component with mean μ_k and standard deviation σ_k.

To obtain a Markov-Switching normal distributed random variable, we are needed to respect the existence of transition matrix

$$\mathbf{P} = \begin{pmatrix} p_{11} & \cdots & p_{K1} \\ \vdots & \ddots & \vdots \\ p_{1K} & \cdots & p_{KK} \end{pmatrix}, \tag{9}$$

where the row k, column j element of $\mathbf{P}$ is the transition probability p_{jk}. The random variable, which follows the mixture Markov-Switching normal distributed process, is obtained in the following steps:

(1) generate uniform random variable u of size N;

(2) for given $s_{n-1} = j$ compute $x = F_k^{-1}(u) = \Phi^{-1}(u)\sigma_k + \mu_k$ if $u \in P_{jk}$ for $k = 1,\ldots,K$, where $\{P_{jk}\}$ forms a partition of $(0,1)$ in the sense that $\bigcup_{k=1}^{K} P_{jk} = (0,1)$, $P_{jk} \cap P_{ji} = \varnothing$ for all $k \neq i$ and $P_{jk} = \left(\sum_{i=0}^{k-1} p_{ji}, \sum_{i=0}^{k} p_{ji}\right]$ with $0 = p_{j0} < \sum_{i=0}^{k=1} p_{ji} < \sum_{i=0}^{k=2} p_{ji} < \ldots. < \sum_{i=0}^{k=K} p_{ji} = 1$;

(3) set $s_n = k$.

2.3 Maximum likelihood estimator of mixture models

Obtaining the parameter estimates $\theta = (\pi_k, \mu_k, \sigma_k)$ is difficult because of unknown realizations of the process $\{Z\}$ which would indicate the k-th regime (component) generating the realizations of our process $\{X\}$. However, it is possible for the purpose of parameter estimation to employ maximum likelihood method based on the iterative EM algorithm.

The EM algorithm (expectation-maximization) consists in two steps. In the first one (*E-step*) for given initial values of parameters $\theta^{(t)}$, the realization of unknown process $\{Z\}$ indicating the regime which generates the realization of process $\{X\}$ is estimated and the expectation logarithm of likelihood function is constructed. In the second step (*M-step*), this function is maximized and new parameter estimates $\theta^{(t+1)}$ are obtained. Both steps are repeated until the convergency criterion is satisfied.

2.4 EM-algorithm

Let´s define the general principle of maximum likelihood estimator. The likelihood function of one observation is assumed to be in the form of

$$f(x;\theta) = \sum_{k=1}^{K} \pi_k f(x;\mu_k,\sigma_k), \tag{10}$$

where $\theta = (\pi_k, \mu_k, \sigma_k)$ are estimated parameters, K is number of regimes (components) of normal mixture distribution, π_k is probability of k-th regime and

$$f\left(x;\mu_k,\sigma_k\right)=\frac{1}{\left(\sqrt{2\pi}\sigma_k\right)}\exp\left(-\frac{\left(x-\mu_k\right)^2}{2\sigma_k^2}\right) \tag{11}$$

is normal distribution density function with mean value μ_k and variance σ_k^2.

The joint likelihood function of normal mixture distribution is defined by

$$L\left(X;\theta\right)=\prod_{n=1}^{N}\sum_{k=1}^{K}\pi_k f\left(x;\mu_k,\sigma_k\right). \tag{12}$$

E-step

The conditional probability of regime $p^{(t)}\left(Z;x,\theta^{(t)}\right)$ generating the realization x and given by estimates $\theta^{(t)}$ can be rewritten in terms of conditional probability law

$$P\left(A\middle|B\right)=\frac{P\left(A\cap B\right)}{P\left(B\right)}$$

in the form of

$$p^{(t)}\left(Z;x,\theta^{(t)}\right)=\frac{\pi_k^{(t)} f\left(x;\mu_k^{(t)},\sigma_k^{(t)}\right)}{\sum_{k=1}^{K}\pi_k^{(t)} f\left(x;\mu_k^{(t)},\sigma_k^{(t)}\right)}, \tag{13}$$

where $\pi_k^{(t)}$ is estimated probability of k-th regime and $f\left(x;\mu_k^{(t)},\sigma_k^{(t)}\right)$ is normal density function given by mean value $\mu_k^{(t)}$ and standard deviation $\sigma_k^{(t)}$.

By rearrangering of likelihood function (10) and applying Jensen inequality for convex functions according to

$$\log\sum_{k=1}^{K}c_k=\log\sum_{k=1}^{K}c_k\frac{\gamma_k}{\gamma_k}\geq\sum_{k=1}^{K}\gamma_k\log\frac{c_k}{\gamma_k} \tag{14}$$

we obtain the form of expectation logarithm of likelihood function for given parameters $\theta^{(t)}$, thus

$$\begin{aligned}\ell\left(X;\theta\right)=\log L\left(X;\theta\right)&=\sum_{n=1}^{N}\log\sum_{k=1}^{K}\pi_k^{(t)} f\left(x;\mu_k^{(t)},\sigma_k^{(t)}\right)\\&=\sum_{n=1}^{N}\log\sum_{k=1}^{K}\pi_k^{(t)} f\left(x;\mu_k^{(t)},\sigma_k^{(t)}\right)\frac{p^{(t)}\left(Z;x,\theta^{(t)}\right)}{p^{(t)}\left(Z;x,\theta^{(t)}\right)}\\&\geq\sum_{n=1}^{N}\sum_{k=1}^{K}p^{(t)}\left(Z;x,\theta^{(t)}\right)\log\frac{\pi_k^{(t)} f\left(x;\mu_k^{(t)},\sigma_k^{(t)}\right)}{p^{(t)}\left(Z;x,\theta^{(t)}\right)}\\&=E\left[\log L\left(X;\theta^{(t)}\right)\right].\end{aligned} \tag{15}$$

M-step

In the second step, the function (15) is maximized with respect of parameters $\theta^{(t)}$, i.e.

$$\hat{\theta}^{(t+1)}=\arg\max_{\theta}\ell\left(X;\theta^{(t)}\right) \tag{16}$$

and both steps are repeated until the convergence.

2.5 Parameter estimates of Markov-Switching models

Firstly and foremost, the difference between the estimating the general mixture model and Markov-switching model consists in the number of estimated parameters because the transition probabilities p_{jk} are also needed to be estimated.

The general likelihood function (12) can be rewritten in the form of

$$\begin{aligned} L(X;\theta) &= \prod_{n=1}^{N}\sum_{k=1}^{K}\pi_k\, f\left(x;\mu_k,\sigma_k\right) \\ &= \prod_{n=1}^{N}\sum_{k=1}^{K}P\left(s_n=k\middle|\Omega_{n-1}\right) f\left(x\middle|s_n=k,\Omega_{n-1};\mu_k,\sigma_k\right), \end{aligned} \tag{17}$$

where $P\left(s_n=k\middle|\Omega_{n-1}\right)=\hat{\boldsymbol{\xi}}_{n|n-1}$ represents predicted probability that $s_n=k$ given by all information available till $n-1$. These predicted probabilities are obtained on the basis of two following equation

$$\hat{\boldsymbol{\xi}}_{n|n} = \frac{\hat{\boldsymbol{\xi}}_{n|n-1} \odot \boldsymbol{\eta}_n}{\mathbf{1}'\left(\hat{\boldsymbol{\xi}}_{n|n-1} \odot \boldsymbol{\eta}_n\right)}, \tag{18}$$

$$\hat{\boldsymbol{\xi}}_{n+1|n} = \mathbf{P}\cdot\hat{\boldsymbol{\xi}}_{n|n-1}, \tag{19}$$

where $\hat{\boldsymbol{\xi}}_{n+1|n}$ and $\hat{\boldsymbol{\xi}}_{n|n}$ is a vector of predicted and filtered probabilities, $\boldsymbol{\eta}_n$ is vector of particular density function (11) with mean μ_k and volatility σ_k and $\odot$ is element-by-element multiplication.

The parameter estimates are obtained by the following iterative procedure. For given values of parameters $\theta^{(t)}=\left(\mu_k^{(t)},\sigma_k^{(t)},p_{jk}^{(t)}\right)$ the predicted probabilities are calculated according to (18) and (19). Then, the expectation of likelihood function is constructed according to the procedure explained above in accordance of Equations (13)-(15) and new parameters $\mu_k^{(t+1)},\sigma_k^{(t+1)}$ are estimated in M-step.

(Hamilton, 1990) showed that the maximum likelihood estimates of transition probabilities $p_{jk}^{(t+1)}$ are calculated according to the equation in the form of

$$p_{jk}^{(t+1)} = \frac{\sum_{n=2}^{N} P\left(s_n=k,s_{n-1}=j\middle|\Omega_N;\mu_k^{(t)},\sigma_k^{(t)}\right)}{\sum_{n=2}^{N} P\left(s_{n-1}=j\middle|\Omega_N;\mu_k^{(t)},\sigma_k^{(t)}\right)}, \tag{20}$$

where $P\left(s_n=k,s_{n-1}=j\middle|\Omega_N;\mu_k^{(t)},\sigma_k^{(t)}\right)=P\left(s_n=k\middle|\Omega_N;\mu_k^{(t)},\sigma_k^{(t)}\right)\cdot P\left(s_{n-1}=j\middle|\Omega_N;\mu_k^{(t)},\sigma_k^{(t)}\right)$ and $P\left(s_{n-1}=j\middle|\Omega_N;\mu_k^{(t)},\sigma_k^{(t)}\right)$, $P\left(s_n=k\middle|\Omega_N;\mu_k^{(t)},\sigma_k^{(t)}\right)$ represent the smoothed probabilities $\hat{\boldsymbol{\xi}}_{n|N}$ obtained using an algorithm developed by (Kim, 1994) on the basis of iterating on

$$\hat{\boldsymbol{\xi}}_{n|N} = \hat{\boldsymbol{\xi}}_{n|n} \odot \left[\mathbf{P}'\cdot\left(\hat{\boldsymbol{\xi}}_{n+1|N} \div \hat{\boldsymbol{\xi}}_{n+1|n}\right)\right] \tag{21}$$

backward for $n=N-1,N-2...$, where $\mathbf{P}$ is matrix of transition probabilities and $\div$ is element-by-element division. The procedure starts at $\hat{\boldsymbol{\xi}}_{N|N}$ corresponding to the filtered probabilities at time N.

From the Equation (20) the new transition matrix $\mathbf{P}$ is obtained and setting $\hat{\xi}_{1|0} = \hat{\xi}_{0|N}$ the whole procedure is repeated until convergency. The whole iterative algorithm starts for given initial values of $\theta^{(0)} = \left(\mu_k^{(0)}, \sigma_k^{(0)}, p_{jk}^{(0)}\right)$ and by setting the starting values $\hat{\xi}_{1|0} = p_{jk}^{(0)}$ for $k = j$.

3 Estimation and results

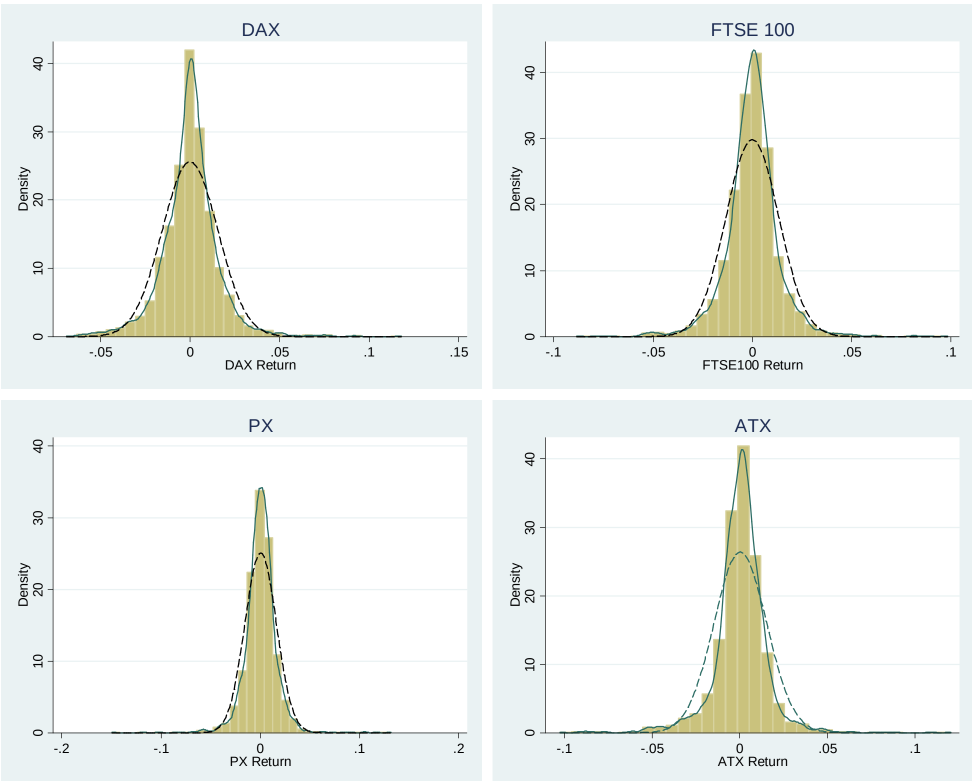
In this section, we demonstrate the procedure of estimating VaR analytically and on the basis of MC simulation. We work with the assumption of normal, mixture and Markov-Switching (MS) normal probability distribution and we compare VaR estimated under all conditions.

First of all, we describe the application data and we conduct some preliminary verification tests of normality. Then we estimate the parameters of mixture distributions employing the maximum likelihood method and we verify the appropriateness of estimated and parameterized probability distribution via two-sample KS test. Finally, we calculate both risk measure considering normal, mixture and MS normal distribution.

3.1 Input data

For the purpose of comparing estimated VaR under the condition of normal, mixture and MS normal probability distribution, we chose the time series of four European market portfolios represented by market indexes, namely DAX, FTSE 100, PX and ATX. We worked with daily returns from the period from 2000 to 2009. In the next Figure 1 we depict the histograms of individual time series and also recorded empirical and normal density functions.

Figure 1: Histograms and empirical density function of time series



Even it is already apparent from the figure above that the empirical distribution are leptokurtic and some even skewed, in spite of it we verify the non-normality by conducting some statistical tests. This step is very important for the suitability of presuming mixture distributions generally.

Firstly and foremost, we tested the empirical skewness and kurtosis against the values of normal distributions (i.e. 0 for skewness and 3 for kurtosis) and we also performed the Shapiro-Wilk test of normality. In the following table there are recorded the empirical characteristics of location and variability. There are also *p*-values of relevant test in brackets.

Table 1: Empirical characteristics and p-values of relevant statistical tests

	DAX		FTSE 100		PX		ATX	
Mean	0.000		0.000		0.000		0.000	
St.dev	0.016		0.013		0.016		0.015	
Skew	0.348	(0.000)	0.058	(0.234)	-0.161	(0.001)	-0.385	(0.000)
Kurtosis	7.820	(0.000)	9.309	(0.000)	14.395	(0.000)	8.716	(0.000)
S-W test		(0.000)		(0.000)		(0.000)		(0.000)

It follows from the results above that empirical distribution of selected time series are non-normal. The distribution of DAX, PX and ATX are even skewed (DAX positive, PX and ATX negative); the kurtosis is in all cases higher than it corresponds to the normal distribution. Therefore we can conclude that it is really reasonable assuming the mixture or MS normal distribution in our application.

3.2 Parameter estimates

The parameter estimates were obtained by maximum likelihood method according to the EM algorithm described in Section 3. We considered only 2 components (normal distributions) and after parameters estimation we verify whether the parameterized mixture and MS normal distribution compound from 2 components fit the empirical distribution adequately. The estimated parameters are recorded in the next table.

Table 2: Parameter estimates of mixture and MS normal distribution

		μ_1	μ_2	σ_1	σ_2	π_1	π_2	p_{11}	p_{12}	p_{22}	p_{21}
Mixture normal	DAX	0.06%	-0.24%	1.10%	2.87%	0.775	0.225	n/a	n/a	n/a	n/a
	FTSE 100	0.04%	-0.16%	0.81%	2.34%	0.763	0.237	n/a	n/a	n/a	n/a
	PX	0.08%	-0.51%	1.16%	3.89%	0.915	0.085	n/a	n/a	n/a	n/a
	ATX	0.13%	-0.41%	0.85%	2.92%	0.807	0.193	n/a	n/a	n/a	n/a
Markov-Switching	DAX	0.08%	-0.18%	0.94%	2.60%	0.673	0.327	0.715	0.285	0.414	0.586
	FTSE 100	0.05%	-0.13%	0.67%	2.12%	0.663	0.337	0.726	0.274	0.461	0.539
	PX	0.09%	-0.22%	0.98%	3.05%	0.814	0.186	0.851	0.149	0.349	0.651
	ATX	0.14%	-0.33%	0.74%	2.76%	0.757	0.243	0.849	0.151	0.530	0.470

As we can see in the Table 2, the parameter estimates are very similar. Only the difference of estimated unconditional probabilities π can be termed as significant. Next, it is necessary to verify whether the parameterized mixture and MS normal distribution fit the empirical distribution well. In other words, we verify whether we can generate (using our estimates) a random variable with the empirical probability distribution corresponding to empirical distribution of time series. For that reason and also for the purpose of simulation, it is totally necessary to verify that the estimated mixture distributions represent the empirical distributions adequately.

Thus, we conducted two-sample Kolmogorov-Smirnov test. Firstly, we generated the sample of size 10,000 values and consequently we conducted two-sample KS test. We tested the H_0: distribution of generated random variable corresponds to the empirical distribution against the H_1: non H_0. The results in the form of *p*-values are introduced in the next table.

Table 3: P-values of two-sample KS test

Sample size	DAX	FTSE 100	PX	ATX
Normal	0.000	0.000	0.000	0.000
Mixture normal	0.108	0.482	0.219	0.139
MS - normal	0.141	0.717	0.904	0.820

In accordance to the results, we confirm the H_0 in the case of mixture distributions only (mixture normal and Markov-Switching) and we conclude that our parameter estimates fit the empirical distribution well at the 95% confidence interval. It can also be noted that estimated MS normal distributions fit the data even better than the parameterized mixture normal distributions.

3.3 Calculating and comparing estimates of risk measure

Finally at 5 and 1 % significance level, we estimate 1, 5 and 10-day analytical VaR and ES according to the Equation (4) and (6) considering the unconditional and conditional probability of each regime (i.e. mixture and MS normal distribution). We also estimate the VaR on the basis of Monte-Carlo simulation. For the simulation we generated random sample of size 10,000 values over 10-day horizon. The empirical risk measures are recorded in the Table 4 and the results are summarized in the Table 5.

Table 4: Empirical Value at Risk at 5 and 1% significance level

	1-day		5-day		10-day	
	5%	1%	5%	1%	5%	1%
DAX	2.68%	5.05%	5.95%	11.13%	8.54%	16.62%
FTSE 100	2.13%	4.01%	4.45%	8.16%	6.24%	12.29%
PX	2.30%	4.24%	5.64%	10.79%	8.18%	16.46%
ATX	2.37%	4.88%	5.37%	11.86%	7.62%	16.55%

Table 5: Analytical and simulated Value at Risk estimates at 5 and 1% significance level

		analytical						simulated					
		1-day		5-day		10-day		1-day		5-day		10-day	
		5%	1%	5%	1%	5%	1%	5%	1%	5%	1%	5%	1%
Normal	DAX	2.76%	**3.90%**	6.19%	**8.74%**	8.77%	**12.38%**	2.71%	**3.86%**	6.08%	**8.41%**	**8.48%**	**12.06%**
	FTSE 100	2.22%	**3.13%**	4.99%	**7.03%**	7.08%	**9.97%**	2.25%	**3.12%**	4.90%	**6.85%**	6.87%	**9.89%**
	PX	2.32%	**3.30%**	**5.11%**	**7.29%**	**7.13%**	**10.22%**	2.45%	**3.34%**	**5.37%**	**7.47%**	**7.77%**	**10.57%**
	ATX	2.45%	**3.48%**	5.40%	**7.70%**	**7.55%**	**10.81%**	2.52%	**3.59%**	5.69%	**7.89%**	7.98%	**11.21%**
Mixture normal	DAX	**2.66%**	5.12%	6.36%	12.12%	9.54%	17.85%	**2.42%**	**4.61%**	6.02%	**8.80%**	**8.41%**	**12.12%**
	FTSE 100	2.13%	4.19%	5.10%	9.81%	7.60%	14.35%	**2.02%**	4.16%	4.85%	**7.42%**	6.87%	**9.88%**
	PX	**2.21%**	5.12%	**4.98%**	12.84%	**7.16%**	19.63%	**2.17%**	**4.16%**	5.77%	**9.10%**	**8.15%**	**11.83%**
	ATX	**2.36%**	5.17%	6.26%	12.67%	10.03%	19.11%	**2.01%**	**4.26%**	5.48%	**8.19%**	8.02%	**11.57%**
Markov-Switching	DAX	2.86%	**5.04%**	6.85%	11.76%	10.19%	17.15%	**2.62%**	**4.75%**	6.91%	**10.40%**	10.58%	**15.42%**
	FTSE 100	2.34%	4.12%	5.59%	9.57%	8.29%	13.92%	**2.08%**	3.79%	5.56%	8.57%	8.54%	12.77%
	PX	2.36%	5.13%	**5.57%**	12.08%	8.33%	17.73%	2.31%	4.64%	7.53%	11.54%	12.39%	18.14%
	ATX	2.60%	5.12%	6.70%	12.36%	10.43%	18.43%	**2.01%**	**4.41%**	5.79%	**9.86%**	9.47%	**14.80%**

According to the results in the table above, the normal distribution condition underestimates the empirical VaR at 1% significance level (bold numbers) and overestimates the true values at 5% significance level. This conclusion holds for both analytical and simulated VaR. The similar conclusion holds for simulated VaR conditioned by normal mixture distribution in the most cases. On the other hand, we can say in general that the parametric VaR under normal mixture distribution overestimates the true values at both significance level and the VaR under the MS normal distribution condition overestimates mostly the empirical values regardless the estimation methods. As far as estimation accuracy is concerned (grey fields), we can conclude that the normal VaR estimates are the most precise estimates at 5 % significant level when the analytical computation is considered. In overall comparison the mixture normal distribution gives better results in a several cases (in the most of them compared to the normal VaR). Finally, Value at Risk under Markov-Switching normal distribution condition outperforms the estimates under the other conditions, especially at lower significance level.

4 Conclusions

The paper was devoted to parametric and Monte-Carlo Value at Risk estimation under mixture and Markov-Switching normal distribution condition. Firstly, we defined the analytical and Monte-Carlo VaR estimates under condition of normal, mixture and Markov-Switching normal distribution and we shortly derived the maximum likelihood estimator for obtaining estimates of parameters of both kind of mixture distributions. Thus we briefly described the iterative EM-algorithm.

Using four selected European market portfolios represented by market indexes (DAX, FTSE 100, PX and ATX), we estimated VaR in accordance to all approaches and we compared all estimates. We gave the evidence on our applications that the normal distribution condition overestimates VaR at given significance level at the same significance level regardless the estimation method. On the other hand, we highlighted that the VaR estimates using the MS normal distribution outperform the other distributions when the lower significant level is considered.

Thus on our results, we highlighted the advantage of using mixture and MS normal distribution (or another approach respecting the empirical distribution). Considering only two components (normal distributions), they fit completely the empirical distribution of selected time series. Especially, the Markov-Switching normal distribution can be highly recommended in accordance with the results of Kolmogorov-Smirnov test. It gives more precise Value at Risk estimates in the most of cases, especially at 99 % confidence level it outperforms the estimates subjected to the normal distribution at all.

References

[1] Alexander, C., 2008. *Value at Risk models.* Chichester: Wiley.

[2] Danielsson, J., 2007. *The Value at Risk: Key Issues in the Implementation of Market Risk.* London: Risk Books.

[3] Dempster, A., Laird, N. & Rubin, D., 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society* , Volume 39, pp. 1-38.

[4] Duffie, D. & Pan, J., 2001. Analytical value-at-risk with jumps and credit risk. *Finance and Stochastics*, Volume 5, pp. 155-180.

[5] Giannopoulos, K., 2003. VaR Modelling on Long Run Horizons. *Automation and Remote Control*, Volume 64, pp. 1094-1100.

[6] Glasserman, P., 2004. *Monte Carlo Methods in Financial Engineering.* New York: Springer Verlag.

[7] Hamilton, J. D., 1990. Analysis of Time Series Subject to Changes in Regime. *Journal of Econometrics,* Volume 45, pp. 39-70.

[8] Chan, N. H. & Wong, H. Y., 2006. *Simulation Techniques in Financial Risk Management.* Hoboken: Wiley.

[9] Jorion, P., 2007. *Value at Risk: The New Benchmark for Managing Financial Risk.* New York: McGraw-Hill.

[10] Kim, C. J., 1994. Dynamic Linear Models with Markov-Switching. *Journal of Econometrics*, Volume 60, pp. 1-22.

[11] McLachlan, G. & Krishnan, T., 1997. *The EM algorithm and extensions.* Chichester: Wiley.

[12] Morgan, J., 1996. *RiskMetricsTM Technical Document.* New York: s.n.

[13] Tichý, T., 2009. *Risk estimation of FX rate sensitive portfolio: subordinated Lévy models and copula approach.* San Jose, s.n.

[14] Tichý, T., 2010. Examination of portfolio currency risk estimation by means of Lévy models. *Politická ekonomie,* Volume 4, pp. 504-521.

[15] Zmeškal, Z., 2005a. Value at Risk methodology of international index portfolio under soft conditions (fuzzy-stochastic approach). *International Review of Financial Analysis,* pp. 263-275.

[16] Zmeškal, Z., 2005b. Value at Risk methodology under soft conditions approach (fuzzy-stochastic approach). *European Journal of Operational Research,* pp. 337-347.