Application of the multivariate skew normal mixture model with the EM Algorithm to Value-at-Risk

S. Soltyk [a]**, R. Gupta**[a]

[a]*Department of Mathematics and Statistics, Curtin University, GPO BOX U1987, Perth, Western Australia, 6845*
Email: sylviasoltyk@gmail.com

Abstract: Since returns of financial assets generally exhibit skewness and kurtosis, modelling returns using a distribution with the ability to capture both of these statistical aspects will increase the accuracy of risk forecasts based on these distributions. The authors propose the use of the multivariate skew normal (MVSN) mixture model to fit asset returns in order to increase the accuracy of Value-at-Risk (VaR) estimates. This paper presents a novel application of the MVSN mixture model to estimate VaR.

There is generally no explicit analytical solution for the parameters of the MVSN mixture model via maximum likelihood estimation (MLE), therefore the use of the Expectation Maximization (EM) Algorithm is proposed in order to find the parameter estimates of the model.

The example provided in this paper consists of a portfolio of monthly returns of six shares listed on the Australian Securities Exchange (ASX). The shares are BHP Billiton Limited (BHP), Commonwealth Bank of Australia (CBA), Cochlear Limited (COH), News Corporation (NWS), Origin Energy (ORG), and Wesfarmers Limited (WES). Hence, the dimensionality, p, of this portfolio is six. The period of analysis for the data is 01/01/1998 - 01/04/2011.

This paper models the MVSN mixture model with a number of mixtures ranging from one to four. A mixture of multivariate normal densities is modelled for comparison to the MVSN mixture model. We find that for one to three mixtures, the MVSN mixture model provides an improved fit. The improved fit of the MVSN mixture model is translated to the performance of the VaR models, where the results show that for one to three numbers of mixtures, the VaR model using the MVSN mixture model assumption indicates improved risk forecasts when compared to the mixture of multivariate normal densities. Furthermore, for the example examined, we find that the model which incorporates the skewness parameter (MVSN mixture model) requires a fewer number of mixtures when compared to a mixture of normal densities. This is an interesting result as reduced model complexity requires less computational ability, computation time, and will results in decreased computational anomalies.

Keywords: Value-at-Risk, Mixture models, Skew normal, Skew normal mixtures, EM Algorithm

1 INTRODUCTION

Returns of financial assets generally exhibit skewness and kurtosis, rendering the normal distribution assumption for estimating risk, such as Value-at-Risk (VaR), a weak tool. The normal distribution has been found to underestimate risk for returns (See, Zangari [1996], Li [1999]). However, the use of mixtures of normal densities have been shown to increase the accuracy of VaR forecasts as the assumption of this distribution allows for higher kurtosis when compared to the normal distribution alone (See, for example, Venkataraman [1997]). If indeed financial returns are skewed, the incorporation of a skewness parameter will further increase the accuracy of risk forecasts. The multivariate skew normal (MVSN) distribution is a natural extension of the multivariate normal distribution, which can accommodate for asymmetry. Therefore, we propose the use of a mixture of MVSN densities (MVSN mixture model) to model asset returns and estimate VaR. This density allows for the inclusion of a skewness parameter, therefore it is more flexible than a mixture of normal densities. There is generally no explicit analytical solution for the parameter estimates of the MVSN mixture model via maximum likelihood estimation (MLE); hence the EM algorithm is used to find these estimates iteratively. Lin [2009] proposes a concise framework for implementing the EM Algorithm to find the parameter estimates of the MVSN mixture model. The use of the MVSN mixture model will be demonstrated by modelling a portfolio which consists of monthly returns of six shares listed on the Australian Securities Exchange (ASX). Therefore, the portfolio corresponds to a dimensionality of six. The multivariate normal mixture model will also be fitted to this portfolio for comparison to the MVSN mixture model. VaR will be calculated via a simulation based on the parameter estimates of the two models, and the results will be compared via a range of tests.

2 THE MULTIVARIATE SKEW NORMAL MIXTURE MODEL

The univariate skew normal density was first considered in depth by Azzalini [1985]. Various forms of multivariate extensions have been considered by Azzalini, A. and A. Dalla Valle [1996], Azzalini, A. and A. Capitanio [1999], Arnold and Beaver [2002], and Sahu, Dey, and Branco [2003].

For a random vector $\mathbf{X}$, the MVSN density with $p \times 1$ location vector ξ, $p \times p$ positive definite covariance matrix $\boldsymbol{\Sigma}$, and $p \times p$ skewness matrix $\boldsymbol{\Lambda}$, is defined as

$$f(\mathbf{x}|\xi, \boldsymbol{\Sigma}, \boldsymbol{\Lambda}) = 2^p \phi_p(\mathbf{x}|\xi, \boldsymbol{\Omega}) \Phi_p\left(\boldsymbol{\Lambda}^T \boldsymbol{\Omega}^{-1}(\mathbf{x} - \xi)|\boldsymbol{\Delta}\right) \tag{1}$$

where $\boldsymbol{\Omega} = \boldsymbol{\Sigma} + \boldsymbol{\Lambda}^T \boldsymbol{\Lambda}$, $\boldsymbol{\Delta} = (\mathbf{I}_p + \boldsymbol{\Lambda}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\Lambda})^{-1}$, and $\phi_p(.|\mu, \boldsymbol{\Sigma})$ and $\Phi_p(.|\boldsymbol{\Sigma})$ denote the pdf of $N_p(\mu, \boldsymbol{\Sigma})$ and cdf of $N_p(\mathbf{0}, \boldsymbol{\Sigma})$, respectively [Lin, 2009].

Now consider a set of random variables $X_1, \ldots, X_n$ that follow a mixture of MVSN densities. The probability density function (pdf) can be defined as

$$\mathbf{X}_j \sim \sum_{i=1}^g w_i f(\xi_i, \boldsymbol{\Sigma}_i, \boldsymbol{\Lambda}_i), \quad w_i \geq 0, \quad \sum_{i=1}^g w_i = 1 \tag{2}$$

where $f(\xi_i, \boldsymbol{\Sigma}_i, \boldsymbol{\Lambda}_i)$ is defined as per equation (1) and, $\boldsymbol{\Theta} = (\theta_1, \ldots, \theta_g)$, where $\theta_i = (w_i, \xi_i, \boldsymbol{\Sigma}_i, \boldsymbol{\Lambda}_i)$ are the unknown parameters of mixture i, and w_i mixing probabilities.

Then, the log-likelihood function of the MVSN mixture model of equation (2) is

$$\ell(\boldsymbol{\Theta}|y) = \sum_{j=1}^n \log\left(\sum_{i=1}^g w_i f(\xi_i, \Sigma_i, \Lambda_i)\right). \tag{3}$$

There is generally no explicit analytical solution for the log-likelihood function of equation (3) via MLE. Therefore, Lin [2009] suggests the use of the EM Algorithm, in order to solve for the maximum likelihood (ML) estimates of the MVSN mixture model. It is considered more efficient when compared to multidimensional optimisation.

3 EM ALGORITHM

The EM Algorithm is an iterative procedure which solves for the parameter estimates via simulation. There are two steps which comprise the algorithm, the E-step and the M-step. Firstly, $\mathbf{x}_1, \ldots, \mathbf{x}_n$ is considered the incomplete or observed data. For the E-step, an unobservable vector of zero-one indicator variables, $\mathbf{Z_j} = (Z_{1j}, \ldots, Z_{gj})^T$, for $j = 1, \ldots, n$, must be introduced. Where $\mathbf{Z_j} \sim multinomial(1; w_1, \ldots, w_g)$, and $Z_{rj} = 1$ if $\mathbf{X_j}$ comes from mixture r. The associated complete-data log-likelihood is $\ell_c(\mathbf{\Theta}|\mathbf{x}, \mathbf{Z}, \tau)$, where $\tau \sim HalfNormal_p(\mathbf{0}, \mathbf{I}_p)$. $\mathbf{\Theta}$ cannot be estimated via the complete-data log-likelihood since it contains unobservable components, however, the E-step overcomes this problem by compensating for the unobservable data by averaging the complete-data log-likelihood over its conditional distribution, given the observed data. Therefore, the Q-function is defined as $Q(\mathbf{\Theta}|\hat{\mathbf{\Theta}}^{(k)}) = E(\ell_c(\mathbf{\Theta}|\mathbf{x}, \tau, \mathbf{Z})|\mathbf{x}, \hat{\mathbf{\Theta}}^{(k)})$, where $\hat{\mathbf{\Theta}}^{(k)}$ is the current parameter estimates. The calculation of the Q-function involves the computation of $E(Z_{ij}|\mathbf{x}_j, \hat{\mathbf{\Theta}})$, $E(Z_{ij}\tau_j|\mathbf{y}_j, \hat{\mathbf{\Theta}}^{(k)})$, and $E(Z_{ij}\tau_j\tau_j^T|\mathbf{x}_j, \hat{\mathbf{\Theta}}^{(k)})$.

Next, the M-step maximizes the Q-function over the current parameters estimates. The E-step and M-step are repeated, or iterated, until convergence. The algorithm converges when the difference between the successive log-likelihoods is small enough, usually a difference of 10^{-5} is considered sufficient.

The starting values of the EM Algorithm can be chosen by performing a k-means clustering algorithm initialised with random start, and by specifying the indicator variable zero-one membership according to the results of the k-means clustering. The initial values of $w_i, \xi_i, \mathbf{\Sigma}_i$ should be chosen as $\hat{w}_i^{(0)} = \sum_{j=1}^n \hat{Z}_{ij}^{(0)} n^{-1}, \hat{\xi}_i^{(0)} = (\sum_{j=1}^n \hat{Z}_{ij}^{(0)} \mathbf{x}_j)(\sum_{j=1}^n \hat{Z}_{ij}^{(0)})^{-1}$, and $\hat{\mathbf{\Sigma}}_i^{(0)} = (\sum_{j=1}^n \hat{Z}_{ij}^{(0)}(\mathbf{x}_j - \hat{\xi}_i^{(0)})(\mathbf{x}_j - \hat{\xi}_i^{(0)})^T)(\sum_{j=1}^n \hat{Z}_{ij}^{(0)})^{-1}$, respectively. The initial values of $\mathbf{\Lambda}_i$, can be chosen as diagonal. The diagonal imposition of the skewness matrix ensures that the correlation structure of the data is unaffected [Sahu, Dey, and Branco, 2003]. Therefore, $\hat{\mathbf{\Lambda}}_i^{(0)} = Diag\{\hat{\Lambda}_{i1}^{(0)}, \ldots, \hat{\Lambda}_{ip}^{(0)}\}$, with values slightly deviated from zero, whose sign corresponds to the sign of the sample skewness of the k-means clustering observations [Lin, 2009].

The computations are performed in the program R (available at http://www.r-project.org/). Even with existing high computational power there are computational limitations, as there are complexities which exist within the methods of using iterative procedures such as the EM Algorithm. The EM Algorithm is an efficient method; however there are computational limitations in some cases. For example, increasing the portfolio size (dimensionality) coupled with a larger number of mixtures (e.g., $g \gg 4$) has been found to drastically increase computational time, and computational anomalies.

4 MODEL COMPARISON

Let SNmix model denote the MVSN mixture model as defined in equation (2). Setting the skewness parameter, $\mathbf{\Lambda}_i = 0 \; \forall i = 1, 2, \ldots, g$ in equation (2) results in a mixture of normal densities. Let this model be denoted NORMmix. The two models will be compared by the construction of Likelihood Ratio (LR) tests, with reference to the skewness parameter, $\mathbf{\Lambda}$. The hypotheses tested for the LR test are $H_0 : \mathbf{\Lambda_1} = \ldots = \mathbf{\Lambda_g} = \mathbf{0}$ versus $H_1 : \mathbf{\Lambda_i} \neq \mathbf{0}$, for at least one $i = 1, \ldots, g$. The associated LR test statistic is

$$LR = -2\{\ell(\hat{\mathbf{\Theta}}|\mathbf{x})_0 - \ell(\hat{\mathbf{\Theta}}|\mathbf{x})_1\} \tag{4}$$

which is asymptotically distributed as χ_r^2, where r is the number of restrictions. Hence, the LR tests will show whether of the introduction of the parameter, $\mathbf{\Lambda}$, which represents the skewness, is significant. The LR test will be conducted for each $g = 1, 2, 3, 4$.

If m and n are the number of parameters, and the number of observations, respectively, then the AIC the BIC can be defined as $AIC = -2(\ell(\hat{\mathbf{\Theta}}|\mathbf{y}) - m)$, and $BIC = -2(\ell(\hat{\mathbf{\Theta}}|\mathbf{y}) - 0.5m \log(n))$. Accounting for over parameterisation, the model with the lower AIC and BIC is preferred.

Assuming asymptotic normality, the parameter significance (i.e., the parameters are different from zero) of the SNmix model are assessed by analysing the standard errors. The standard errors are found via the inverse of the information matrix [Lin, 2009]. The parameter estimates and standard errors of the parameter estimates are available upon request from the authors.

5 VALUE-AT-RISK AND MODEL EVALUATION

Although VaR has been criticised for not satisfying the axioms to be a coherent risk measure by Artzner et al. [1999], VaR is consistently used by banking institutions as a measure of risk.

Consider $\mathbf{X}$ as the aggregate return, such that, $\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2 + \ldots + \mathbf{X}_p$, where each $\mathbf{X}_i$ for $i = 1, 2, \ldots, p$ are the return of each of the shares in the portfolio. If one is interested in possible losses, or negative returns, then, VaR can be defined as the smallest value satisfying

$$P(\mathbf{X} < x_q) = \alpha \tag{5}$$

where α is the significance level and x_q is the qth quantile of $\mathbf{X}$. Hence, VaR can be understood as the quantile of $\mathbf{X}$ at a given significance level.

Assuming that $(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_p)$ follows the SNmix model and NORMmix model, respectively, and the corresponding parameter estimates are found via the EM Algorithm, a large set of observations is simulated based on the parameters of the fitted distribution. Next, the simulated value of $\mathbf{X}$ is obtained, and VaR is calculated by finding the appropriate quantile. These two models will be denoted SNmix VaR model and NORMmix VaR model, respectively.

In order to verify if the VaR models are adequate, Kupiec [1995] proposes a test, based on backtesting, to determine if the losses projected by the VaR models accurately predict actual losses. The number of violations follows a binomial distribution [Jorion, 1995]. Let the proportion of violations, $\nu = \frac{\omega}{n}$, where ω is the number of violations of VaR in the actual data, and n is the sample size. The hypotheses of this test are $H_0 : \alpha = \alpha^*$ versus $H_1 : \alpha \neq \alpha^*$, where α is the probability of violation as defined by the VaR model, and α^* is the target value of α. The LR statistic is

$$LR_{backtesting} = 2\{\ln\left(\nu^\omega(1-\nu)^{n-\omega}\right) - \ln\left(\alpha^{*\omega}(1-\alpha^*)^{n-\omega}\right)\} \tag{6}$$

which is χ_1^2 distributed.

Furthermore, it can be seen that as $\alpha^* \to \alpha$, then $\frac{\omega}{\alpha \times n} \to 1$. The exceeding ratio, as defined by Choi and Min [2011], is therefore defined as the number of violations of VaR, ω, divided by the expected number of violations, $\alpha \times n$. It can be seen that if the exceeding ratio is greater than 1, then the VaR model is under-forecasting VaR, and if the exceeding ratio is less than 1, the VaR model is over-forecasting VaR. This result can be used to rank the VaR models, and will be used as a means to compare the SNmix VaR models to the NORMmix VaR models in order to judge which more accurately forecasts risk. Moreover, the exceeding ratio can be used as a method for identifying which number of mixtures, g, of the SNmix VaR models, provides the best risk estimate.

Christoffersen [1998] proposes a test for serial independence of VaR violations. This tests the null hypothesis of serial independence against the alternative of first-order Markov dependence. The LR statistic is

$$LR_{independence} = (1-\pi)^{(n_{00}+n_{10})}\pi^{(n_{01}+n_{11})} \tag{7}$$

where n_{00} denotes the number of observations in which a violation followed consecutively by another violation, n_{01} denotes the number of observations in which a violation followed consecutively by no violation, n_{10} denotes the number of observations in which no violation is followed consecutively by a violation, n_{11} denotes the number of observations where no violation is followed consecutively by no violation, and $\pi = \frac{n_{01}+n_{11}}{n_{00}+n_{10}+n_{01}+n_{11}}$. The test statistic is χ_1^2 distributed.

6 EXAMPLE

The example presented consists of a portfolio of the monthly returns of six shares listed on the Australian Securities Exchange (ASX). The data is sourced from Datastream. The shares are BHP Billiton Limited (BHP), Commonwealth Bank of Australia (CBA), Cochlear Limited (COH), News Corporation (NWS), Origin Energy (ORG), and Wesfarmers Limited (WES). Hence, the dimensionality, p, is 6. The return is calculated by $r_t = 100 \times \log\left(\frac{P_t}{P_{t-1}}\right)$, where r_t is the return at time t, P_t and P_{t-1} correspond to the price of the share at time t and $t-1$, respectively. The period of analysis for the data is 01/01/1998 -

01/04/2011. Table 1 shows the summary statistics for the returns of each share, and the return of the total portfolio which is the sum of all the returns for each share. There is skewness and excess kurtosis present in all share returns and the return of the total portfolio.

Table 1. Summary statistics

	BHP	CBA	COH	NWS	ORG	WES	Total
Minimum	-22.662	-24.5296	-29.972	-30.1022	-37.9751	-25.1258	-78.020
Maximum	14.3187	16.0597	16.9658	23.8883	29.0528	21.7674	77.710
Mean	0.2844	0.1511	0.4050	-0.0040	0.2519	0.1600	1.248
Std Dev	4.3210	3.2624	4.2065	5.0084	3.9754	3.6951	26.4785
Skewness	-0.5212	-0.2092	-0.3296	-0.0137	0.8513	-0.1845	-0.4152
Kurtosis	5.3142	7.0802	6.9871	6.4571	11.0189	7.8920	6.3987

The LR statistic for $g = 1$ number of mixtures is 14.796 (p-value = 0.0219), $g = 2$ number of mixtures is 24.484 (p-value = 0.0175), $g = 3$ number of mixtures is 40.058 (p-value = 0.0021), and $g = 4$ number of mixtures is 14.036 (p-value = 0.9458). Thus, there is sufficient evidence to conclude that the skewness parameter, Λ, is significantly different from zero for $g = 1, 2, 3$ number of mixtures at the 5% level of significance. Therefore, these results suggest that for $g = 1, 2, 3$ number of mixtures, the SNmix model is favoured over the NORMmix model. The results of the AIC suggest that the SNmix model is statistically favoured over the NORMmix models for $g = 1, 2, 3$ number of mixtures, however the NORMmix model is favoured for $g = 4$ number of mixtures. The results of the more stringent BIC suggest that the SNmix model is favoured over the NORMmix model for only $g = 1$ number of mixtures, with the NORMmix model being favoured for all other number of mixtures.

Table 2. Summary of the VaR model results

number of mixtures, g		SNmix	NORMmix
1	VaR	-55.83	-53.90
	exceeding ratio	1.2579	2.5157
	$LR_{Backtesting}$	0.0987	2.5976
	$LR_{Independence}$	$2.1962e^{-05}$	$3.8660e^{-07}$
2	VaR	-57.19	-52.16
	exceeding ratio	1.2579	3.1447
	$LR_{Backtesting}$	0.0987	4.7115$\star\star$
	$LR_{Independence}$	$2.1962e^{-05}$	$9.5141e^{-09}$
3	VaR	-79.30	-39.84
	exceeding ratio	0.6289	5.0314
	$LR_{Backtesting}$	0.2547	13.2960$\star\star\star$
	$LR_{Independence}$	0.0023	$1.6908e^{-14}$
4	VaR	-78.48	-56.77
	exceeding ratio	0.6289	1.2579
	$LR_{Backtesting}$	0.2547	0.0987
	$LR_{Independence}$	0.0023	$2.1962e^{-05}$

$\star \star \star$ significant at the 0.01 level

$\star\star$ significant at the 0.05 level

$\star$ significant at the 0.10 level

The results of the VaR, with $\alpha = 0.01$, are presented in Table 2. The results of the in-sample backtesting show that the SNmix VaR model accurately predicts risk for all number of mixtures. Moreover, the NORMmix VaR model inaccurately predicts risk for $g = 2, 3$ number of mixtures. Furthermore, the results of the exceeding ratio show that the SNmix VaR model out-performs the NORMmix VaR model for $g = 1, 2, 3$ number of mixtures. The results of the serial independence test show that for all number of mixtures, for both the SNmix and NORMmix VaR models, the violations are not serially correlated. Furthermore, since the exceeding ratio can be used to rank VaR models, the models which perform equally best are the SNmix VaR model with $g = 1, 2$ number of mixtures, and the NORMmix VaR model with $g = 4$ number. The associated VaR is \$55.83, \$57.19, and \$56.77 respectively, corresponding to a 1% chance of a loss equal to these amounts.

7 CONCLUSIONS AND RECOMMENDATIONS

We propose the use of the MVSN mixture model to fit asset returns since financial data tends to exhibit both skewness and excess kurtosis. The EM Algorithm is used to find the parameter estimates of the MVSN mixture model, since there is generally no explicit analytical solution for the parameter estimates via MLE. The example presented within this paper consists of monthly log returns of shares listed on the ASX. The dimensionality of the portfolio presented is six. For this example, the results of the model fit indicates that the SNmix model with $g = 1, 2, 3$ number of mixtures provides a more accurate fit to the data when compared to the NORMmix model. Hence, the incorporation of a skewness parameter results in an improved fit for the data. The improved fit of the SNmix model for these numbers of mixtures is translated to the estimation of VaR, where the model validation shows that the SNmix VaR model with these numbers of mixtures more accurately forecasts risk when compared to the NORMmix VaR model. Furthermore, with respect to the accuracy of forecasting risk, the SNmix VaR model with $g = 1, 2$ number of mixtures, and the NORMmix VaR model with $g = 4$ number of mixtures provide a comparable fit. In terms of estimation, the SNmix models require 34 and 67 parameters estimated, for $g = 1, 2$ number of mixtures, respectively. Whereas, the NORMmix model with $g = 4$ number of mixtures requires the estimation of 111 parameters. Therefore, for the given example, modelling share returns using a distribution which can capture skewness reduces the number of mixtures required to accurately forecast risk. This result is important as with reduced model complexity comes reduced computational time, and computational requirement. Moreover, computational anomalies are less prevalent when model complexity is reduced.

This paper presents the estimation of VaR with the assumption that the data analysed are independently and identically distributed (iid); the authors accept that there is empirical evidence which suggests that an underlying non-iid structure may be present with financial time series data. However, this paper demonstrates that, even for the iid case, fitting the MVSN mixture model to financial returns data increases the accuracy of VaR forecasts when compared to a more traditional method (mixture of normal densities). Thus, the authors suggest that the methodology presented within this paper can be extended and applied to data which has first been fitted with a conditional volatility model (GARCH-type model), to capture the possible time-varying relationships which may exist in financial returns data.

REFERENCES

Arnold, B. and R. Beaver (2002). Skewed multivariate models related to hidden truncation and/or selective reporting. *Sociedad de Esfadistica e Investigacion Operativa 11*, 7–54.

Artzner, P., F. Delbaen, J. Eber, and D. Heath (1999). Coherent Measures of Risk. *Mathematical Finance 9*, 208–228.

Azzalini, A. (1985). A Class of Distributions Which Includes the Normal Ones. *Scandinavian Journal of Statistics 12*, 171–178.

Azzalini, A. and A. Capitanio (1999). Statistical applications of the multivariate skew normal distribution. *Journal of the Royal Statistical Society 61*, 579–602.

Azzalini, A. and A. Dalla Valle (1996). The Multivariate Skew-Normal Distribution. *Biometrika 83*, 715–726.

Choi, P. and I. Min (2011). A Comparison of Conditional and Unconditional Approaches in Value-at-Risk Estimation. *The Journal of the Japanese Economic Association 62*, 99–115.

S. Soltyk and R.Gupta, Application of the multivariate skew normal mixture model ...

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review 39*(4), 841 – 862.

Jorion, P. (1995). *Value at Risk*. New York: The McGraw-Hill Companies Inc.

Kupiec, P. (1995). Techniques for Verifying the Accuracy of Risk Measurement Models. *The Journal of Derivatives 3*, 73–84.

Li, D. X. (1999). Value at risk based on the volatility, skewness and kurtosis. *Riskmertrics Group*.

Lin, T. (2009). Maximum likelihood estimation for multivariate skew normal mixture models. *The Journal of Multivariate Analysis 100*, 257–265.

Sahu, K., D. K. Dey, and M. D. Branco (2003). A New Class of Multivariate Skew Distributions with Applications to Bayesian Regression Models. *The Canadian Journal of Statistics 31*, 129–150.

Venkataraman, S. (1997). Value at risk for a mixture of normal distributions: The use of quasi-Bayesian estimation techniques. *Federal Reserve Bank of Chicago*.

Zangari, P. (1996). An improved methodology for measuring VaR. *Risk Metrics Monitor, Reuters/J.P. Morgan 2nd quarter*, 7–25.