Modelling asset return using multivariate asymmetric mixture models with applications to estimation of Value-at-Risk

Sharon X. Lee [a]**, Geoffrey J. McLachlan** [a]

[a]*School of Mathematics and Physics, University of Queensland, Brisbane, Australia*
Email: s.lee11@uq.edu.au

Abstract: Value-at-Risk (VaR) is a widely used statistical measure in financial risk management for quantifying the level of risk associated with a specific investment portfolio. It is well-known that historical return data exhibit non-normal features, such as heavy tails and skewness. Current analytical (parameteric) calculation of VaR typically assumes the distribution of the portfolio return to be a normal or log-normal distribution, which results in underestimation and overestimation of the VaR at high and low confidence levels, respectively, when a normal distribution is assumed.

This study develops a promising approach in modelling asset returns by fitting multivariate mixture models with asymmetric component densities, in particular, members of the skew-symmetric family and other non-elliptically contoured distributions. We focus on component densities with four or more parameters, including the multivariate skew t (MST) distribution, the multivariate normal-inverse-Gaussian (MNIG) distribution, and the multivariate generalized hyperbolic distribution (MGH) distribution. These distributions have proven to be effective in capturing heterogeneous data with asymmetric and heavy tail behaviour, and can flexibly adapt to a variety of distributional shapes. The fitting of these mixture models can be carried out via the Expectation–Maximization (EM) algorithm.

This approach has improved the accuracy of VaR estimation, and an example is demonstrated on a portfolio of three Australian stock returns. The asymmetric mixture models were fitted to the monthly returns of the shares Flight Centre Limited (FLT), Westpac Banking Corporation (WBC) and Australia and New Zealand Banking Corporation Group Limited (ANZ) for the period January 2000 to mid 2013. The VaR estimates predicted by the asymmetric mixture models for a range of significance levels of interest compare favourably to traditional methods based on symmetric models, with significant improvements observed in the accuracy of the estimates. Moreover, it is observed that models with relatively restrictive component densities (such as the normal mixtures, skew normal mixtures and shifted asymmetric Laplace mixtures) require a mixture of two components in order to accommodate the skewness and heavy-tails in the data, whereas the more flexible distributions can adequately capture the distributional shape of data with only a single component.

Keywords: mixture model, skew distributions, skew t-mixture model, generalized hyperbolic distribution, EM algorithm, Value-at-Risk

1 INTRODUCTION

Risk forecast is an important topic in finance. Numerical measures such as Value-at-Risk (VaR), expected shortfall (ES), and tail conditional expectation (TCE) rely on the assumption that the returns follow some distribution. It is well recognized in the financial literature that financial returns have non-normal distributions. They generally exhibit features such as heavy tails, skewness, and kurtosis. Many studies have shown that traditional approaches, which are typically based on the assumption of an underlying multivariate normal distribution (either explicitly or implicitly), have failed to perform satisfactorily when applied to real data. To allow for a greater flexibility in the shape of the underlying distribution, some authors have considered finite mixtures of normal distributions, and slight improvements were observed in the accuracy of risk measures (Venkataraman (1997)). Recently, Soltyk and Gupta (2011) reported that further improvements can be made by adopting component densities that incorporate a skewness parameter, that is, by fitting a multivariate skew normal mixture model. However, in addition to asymmetry, empirical evidence has indicated that the distribution of financial returns also exhibit other features such as heavy tails and kurtosis.

In this paper, we propose the use of more flexible parametric families in modelling portfolio returns. In particular, we consider finite mixtures of asymmetric distributions, where the component densities have four or more parameters to regulate the shape of distribution. This includes the skew t-distribution, the normal inverse Gaussian distribution, and a special case of the generalized hyperbolic distribution. The model parameters can be estimated efficiently using the EM algorithm. We illustrate the usefulness of this approach in modelling a portfolio of three Australian shares. VaR estimates are then calculated based on the fitted model parameters and the results are compared to methods based on other non-normal distributions.

2 MULTIVARIATE ASTMMETRIC MIXTURE MODELS

2.1 The multivariate skew normal and skew t-mixture models

The skew symmetric family of distributions, in particular, the skew normal and skew t-distributions, have received increasing attention in recent years. These families of distributions are extensions of the (univariate) skew normal distribution (Azzalini, 1985) with additional parameters to regulate skewness and other features. For a comprehensive survey of skew distributions, the reader is referred to the edited volume of Genton (2004) and the papers by Azzalini (2005), Arellano-Valle and Azzalini (2006), and Genton (2006).

To establish notation, a random vector $\boldsymbol{Y}$ has a restricted multivariate skew normal (rMSN) distribution (Pyne et al., 2009) with location vector $\boldsymbol{\mu}$, scale matrix $\boldsymbol{\Sigma}$, and skewness parameter $\boldsymbol{\delta}$, if its density is given by

$$f(\boldsymbol{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\delta}) = 2\phi_p(\boldsymbol{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) \, \Phi_1\left(\boldsymbol{\delta}^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{\mu}); 0, 1 - \boldsymbol{\delta}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\delta}\right), \tag{1}$$

where $\phi_p(.; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ is the density of the p-variate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, and $\Phi_p(.; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ is the corresponding distribution function. The density of a restricted multivariate skew t (rMST) distribution can be defined in a similar way, given by

$$f(\boldsymbol{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\delta}, \nu) = 2t_p(\boldsymbol{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu) \, T_1\left(\boldsymbol{\delta}^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{\mu})\sqrt{\frac{\nu + p}{\nu + d(\boldsymbol{y})}}; 0, \lambda, \nu + p\right), \tag{2}$$

where $\lambda = 1 - \boldsymbol{\delta}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\delta}$, and $d(\boldsymbol{y}) = (\boldsymbol{y} - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{\mu})$ is the squared Mahalanobis distance between $\boldsymbol{y}$ and $\boldsymbol{\mu}$ with respect to $\boldsymbol{\Sigma}$. Here, we let $t_p(.; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$ denote the p-dimensional t-density with location vector $\boldsymbol{\mu}$, scale matrix $\boldsymbol{\Sigma}$, and degrees of freedom ν, and $T_p(.; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$ is the corresponding distribution function. Note that (1) and (2) are known as a *restricted* form of the MSN and MST distributions, respectively, due to a restriction on the underlying latent skewing variable in the characterizations; see Lee and McLachlan (2013c) for details.

Sahu et al. (2003) proposed an alternative form of MSN and MST distribution, referred to as the *unrestricted* form by Lee and McLachlan (2013a), where the univariate distribution function in (1) and (2) is replaced by a multivariate version. The density of the unrestricted MST (uMST) distribution is given by

$$f(\boldsymbol{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\delta}, \nu) = 2^p t_p(\boldsymbol{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu) \, T_p\left(\boldsymbol{\Delta}\boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{\mu})\sqrt{\frac{\nu + p}{\nu + d(\boldsymbol{y})}}; 0, \boldsymbol{\Lambda}, \nu + p\right), \tag{3}$$

where $\boldsymbol{\Lambda} = \boldsymbol{I}_p - \boldsymbol{\Delta}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Delta}$. It is important to note that (3) is neither a superclass of (2) nor an equivalent form of it. The two distributions are equivalent only in the univariate case.

In this paper, we consider the modelling of stock returns using finite mixtures of MSN and MST distributions. The density of a g-component finite mixture model is defined as a convex linear combination of g densities, given by

$$f(\boldsymbol{y}; \boldsymbol{\Psi}) = \sum_{i=1}^{g} \pi_i f(\boldsymbol{y}; \boldsymbol{\theta}_i), \tag{4}$$

where $\boldsymbol{\theta}_i$ contains the unknown parameters of the ith component density, $\boldsymbol{\Psi} = (\pi_1, \ldots, \pi_{g-1}, \boldsymbol{\theta}_1^T, \ldots, \boldsymbol{\theta}_g^T)^T$ is the vector of all unknown parameters of the mixture model, and the mixing proportions π_i are non-negative and sum to one. When component densities $f(\boldsymbol{y}; \boldsymbol{\theta}_i)$ in (4) take the form of (1), (2), and (3), we obtain a finite mixture of rMSN (FM-rMSN) distributions, of rMST (FM-rMST) distributions, and of uMST (FM-uMST) distributions, respectively.

2.2 The family of multivariate generalized hyperbolic mixture models

The generalized hyperbolic (GH) distribution has been used repeatedly in the finance literature as a flexible alternative to the normal distribution. This family of distributions was introduced as a normal mean-variance mixture distribution in Barndorff-Nielsen (1977) and Barndorff-Nielsen (1981). It includes many popular distributions as subfamilies, special or limiting cases, including the normal, t, variance gamma, normal-inverse-Gaussian, hyperbolic, asymmetric Laplace, skew normal and skew t-distributions. However, the canonical characterization of the GH distribution suffers from an identifiability issue, and hence its use in mixture modelling has been limited. For our purpose, we follow the alternative parameterization considered in Browne and McNicholas (2013), where the scale parameter is fixed to be one. Under this parametrization, the multivariate GH (hereafter MGH) can be obtained as a convolution of a (univariate) generalized inverse Gaussian (GIG) random variable and a centered (multivariate) normal random vector, as follows

$$\boldsymbol{Y} = \boldsymbol{\mu} + V\boldsymbol{\delta} + \sqrt{V}\boldsymbol{X}, \tag{5}$$

where $V \sim GIG(\omega, \omega, \lambda)$, $\boldsymbol{X} \sim N_p(\boldsymbol{0}, \boldsymbol{\Sigma})$ and GIG denotes the generalized inverse Gaussian density (Bardorff-Nielsen and Halgreen, 1977). The density of this MGH distribution is given by

$$
\begin{aligned}
f(\boldsymbol{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\delta}, \lambda, \omega) \quad = \quad & \frac{e^{-\boldsymbol{\delta}^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})}}{(2\pi)^{\frac{p}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}}} \left(\frac{\omega + d(\boldsymbol{y})}{\omega + \boldsymbol{\delta}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\delta}} \right)^{\left(\frac{\lambda}{2} - \frac{p}{4}\right)} \\
& \frac{K_{\lambda - \frac{p}{2}}\left(\sqrt{(\omega + \boldsymbol{\delta}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\delta})(\omega + d(\boldsymbol{y}))} \right)}{K_\lambda(\omega)},
\end{aligned} \tag{6}
$$

where $d(\boldsymbol{y}) = (\boldsymbol{y} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{\mu})$, and K_λ denotes the λth order modified Bessel function of the third kind. The parameter ω is known as a concentration parameter and λ is an index parameter. It is worth noting that this restricted formulation (6) no longer encompasses the aforementioned subfamilies of commonly used distributions.

We consider also a special case of the canonical GH distribution. The multivariate normal inverse Gaussian (MNIG) distribution is obtained by taking the GIG mixing variable V to have parameter $\boldsymbol{\theta} = (\psi, \chi, -\frac{1}{2})^T$. The MNIG distribution, denoted by $MNIG(\boldsymbol{y}; \boldsymbol{\mu}, \boldsymbol{\delta}, \xi, \sigma)$, has a set of five parameters, where ξ affects the tails of the distribution and σ is a scale parameter. Its density is given by

$$
\begin{aligned}
f(\boldsymbol{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\delta}, \xi, \sigma) \quad = \quad & 2^{-\left(\frac{p-1}{2}\right)} \sigma \left(\frac{\sqrt{\xi^2 + \boldsymbol{\delta}^T \boldsymbol{\Sigma} \boldsymbol{\delta}}}{\pi \sqrt{\sigma^2 + d(\boldsymbol{y})}} \right)^{\frac{p+1}{2}} e^{\xi\sigma + \boldsymbol{\delta}^T(\boldsymbol{y}-\boldsymbol{\mu})} \\
& K_{\frac{p+1}{2}}\left(\sqrt{(\xi^2 + \boldsymbol{\delta}^T \boldsymbol{\Sigma} \boldsymbol{\delta})(\sigma^2 + d(\boldsymbol{y}))} \right).
\end{aligned} \tag{7}
$$

Note that the multivariate normal distribution can be obtained as a limiting case of the MNIG distribution with ξ and σ approaching infinity. The MGH and the MNIG distributions considered here belong to the restricted form of skew distributions, as the latent skewing variable V is a univariate random variable. In the example to follow, we consider also the fitting of mixtures of MGH and MNIG (hereafter FM-MGH and FM-MNIG, respectively) distributions to a portfolio of Australian stocks.

3 PARAMETER ESTIMATION AND THE EM ALGORITHM

Parameter estimation for the asymmetric mixture models can be effectively carried out via the Expectation–Maximization (EM) algorithm (Dempster et al., 1977). The EM algorithm has become the standard tool for the calculation of maximum likelihood (ML) estimates of the parameters of a mixture model. It is an iterative procedure that alternates the E- and M-steps until some convergence criterion is satisfied. A detailed description of the EM algorithm for the restricted and unrestricted skew normal and skew t-distributions can be found in Lee and McLachlan (2013a) and the references therein. Software implementation of the restricted and unrestricted mixture models is available from the R packages EMMIX-skew (Wang et al., 2009) and EMMIX-uskew (Lee and McLachlan, 2013b), respectively. For the generalized hyperbolic family and its subfamilies, the EM algorithms are given in Browne and McNicholas (2013), Karlis and Santourian (2009), and Franczak et al. (2012), for the FM-MGH, FM-MNIG, and FM-MSAL models, respectively.

4 ESTIMATION OF VALUE-AT-RISK AND PERFORMANCE EVLUATION

4.1 Definition of Value-at-Risk

Value-at-Risk (VaR) is one of the most commonly use risk measures for investment loss. It is given by the predicted maximum loss over a specified holding period given a specified confidence level. More formally, let $Y_1, \ldots, Y_p$ denote a portfolio of p assets returns and $Y_R = \sum_{j=1}^{p} Y_j$ be the aggregate return. Then the VaR is defined to be the negative of the largest value of y_α satisfying

$$\mathrm{pr}\{Y_R < y_\alpha\} \leq \alpha, \tag{8}$$

where α is the significance level. Thus the VaR can be interpreted as the negative of the quantile of order α of the portfolio's (hypothetical) return distribution F_R; that is, VaR can be expressed as $-F_R^{-1}(\alpha)$. Note that the negative sign preceding $F_R^{-1}(\alpha)$ ensures the VaR is a positive value, that is, a positive amount of 'losses'. For asset returns, typical values of α are small, ranging between 0.01 and 0.05. For example, if $\alpha = 1\%$ and the time period is one day, then a VaR of one million dollars can be interpreted as meaning that the probability of incurring a loss in excess of one million dollars for this portfolio over this day is bounded by 0.01.

4.2 Backtesting

Various techniques have been proposed for evaluating the accuracy of a VaR measure. We shall consider two of the most common statistical tests available, namely the unconditional coverage test (or backtesting) and the (Markov) independence test. The backtest (Kupiec, 1995) is a strategy for testing the unconditional coverage property of a VaR estimate model, that is, it is concerned with whether or not the observed and expected violation rate are statistically different. Specifically, the test examines the proportion of violations, given by $r = v/n$, where v denotes the number of violations in the data and n is the number of observations, and determine whether it is statistically different from the expected rate of $\alpha \times 100\%$ of the sample. A violation occurs if the actual loss exceeds the VAR as implied by the model for the given level of significance. Under the null hypothesis that the model is adequate, v follows a binomial distribution, leading to a likelihood ratio (LR) test statistic of the form

$$\mathrm{LR}_{bt} = 2 \log \left[\left(\frac{1-r}{1-\alpha} \right)^{n-v} \left(\frac{r}{\alpha} \right)^{v} \right], \tag{9}$$

which has a chi-squared distribution with one degree of freedom. Hence the test would reject a VaR model if it generates too many or too few violations. The backtest, however, cannot detect whether a VaR model satisfies the independence property.

4.3 Independence test

Christoffersen (1998) proposed a more elaborate test that examines whether or not a VaR violation process is serially dependent. The underlying assumption is that, for a good and accurate VaR model, the probability of observing a VaR violation should not depend on whether or not a violation has occurred for the previous observation, that is, the sequence of violations should be independently distributed. The LR test statistic for the Independence test can be expressed as

$$\mathrm{LR}_{idp} = -2 \log \left[\frac{(1-q)^{N_1+N_2} q^{N_3+N_4}}{(1-q_1)^{N_1} q_1^{N_3} (1-q_2^{N_2}) q_2^{N_4}} \right], \tag{10}$$

	FLT	**WBC**	**ANZ**
minimum	-47.13	-15.06	-16.73
maximum	39.08	12.24	16.87
mean	0.8562	16.87	1.102
std. dev.	11.48	5.496	5.858
skewness	-0.7944	-0.5709	-0.3418
kurtosis	5.8182	3.173	3.613

Table 1. Summary statistics of the monthly returns of three Australian stocks for period of early 2000 to mid-2013.

	FM-MN	FM-rMSN	FM-rMST	FM-uMST	FM-MNIG	FM-MGH
VaR	27.14	27.27	27.90	28.59	28.26	28.26
exceeding ratio	1.12	1.12	1.12	1.12	1.12	1.12
backtesting	0.74	0.74	0.74	0.74	0.74	0.74
independence	1.00	1.00	1.00	1.00	1.00	1.00

Table 2. Performance of various skew mixture models on estimating the 5% VaR of three Australian stocks. The backtesting and independence values refers to the P-value of the respective tests. The empirical VaR is $29.64.

where N_1 denotes the number of observations with no violation followed by no violation, N_2 denotes the number of observations with violation followed by no violation, N_3 denotes the number of observations with no violation followed by a violation, and N_4 denotes the number of observations in which a violation has occurred followed by another violation. The proportions q, q_1, and q_2 are defined, respectively, as the proportion of observations in which a violation has occurred, the proportion for which a violation is observed given no violation has occurred in the previous observation, and the proportion for which a violation is observed given a violation had occurred in the previous observation. More formally, they are given by $q = \frac{N_3+N_4}{N_1+N_2+N_3+N_4}$, $q_1 = \frac{N_3}{N_1+N_3}$, and $q_2 = \frac{N_4}{N_2+N_4}$. The null hypothesis of $q_1 = q_2$ can be tested against the alternative of first-order Markov independence, and the LR statistic (10) again has a χ_1^2 distribution.

There are existing strategies for handling financial return series that are serially dependent. One popular approach is to apply a GARCH filter to the returns before fitting a skew model, then perform de-filtering on the fitted density to obtain a conditional risk forecast for the return; see, for example Hu and Kercheval (2007).

4.4 Exceeding Ratio

We consider also the exceeding ratio (ER) (Choi and Min, 2011), defined as the ratio of the estimated number of violations over the expected number of violations, that is,

$$ER = \frac{\nu}{\alpha n}. \tag{11}$$

An ER value greater than one indicates the model is under-forecasting the VaR, and an ER value less than one indicates an over-forecast VaR estimate.

5 APPLICATION TO AUSTRALIAN STOCKS

5.1 The data

In this paper, we consider a portfolio of three shares listed on the Australian Stock exchange (ASX). The data consist of monthly returns for the shares of Flight Centre Limited (FLT), Westpac banking Corporation (WBC), and Australia and New Zealand Banking Group Limited (ANZ), for the period 1st January 2000 to 28rd June 2013. The monthly returns of each share was calculated on the adjusted closing prices, and results are recorded as a percentage. A plot of the data is given in Figure 1(a). The summary statistics of the stock returns (Table 1) suggest that the data do not satisfy the normality assumption, with mild skewness and excess kurtosis present in all three returns.
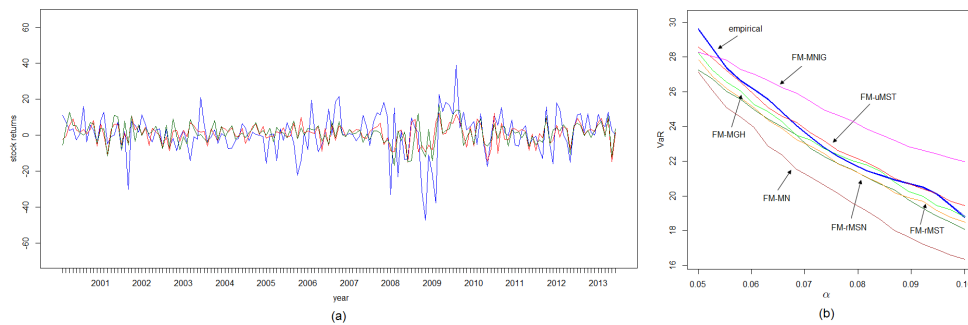
Figure 1. Estimation of VaR for a portfolio of Australian stocks. (a) the return series of three Australian stocks (blue:FLT; red: WBC; green: ANZ) and the aggregate return (dotted line) for the period of Jan 2000 to Jun 2013; (b) the estimated VaR models from various asymmetric mixture models.

5.2 Results

For this illustration, we compare the performance of the FM-MN (multivariate normal), FM-rMSN, FM-rMST, FM-uMST, FM-MGH, and FM-MNIG distributions in fitting this trivariate dataset. To determine the optimal number of components g, each mixture model is fitted with g ranging from 1 to 4, and the best model is selected on the basis of BIC. Estimation of VaR can be obtained via a simulation approach based on the fitted parameters of the mixture model (see Soltyk and Gupta (2011) for an example using the FM-uMSN model). Except for the FM-MN and FM-rMSN models for which BIC selected $g = 2$ number of components, the other mixture models require only a single component. The estimated 5% VaR value given by the FM-MN and other five mixture models is reported in the first row of Table 2. Given that the empirical VaR calculated from the data is \$29.64, it can be observed that the FM-uMST model gives the closest estimate. All the models considered did not fail the backtesting and independence test (rows 3 and 4 in Table 2). The ER indicates that all models over-estimated the VaR, but it is evident from Figure 1 that the FM-uMST and FM-MGH models predict risk more accurately. As expected, it can be observed from Figure 1(b) that the asymmetric mixture models follow more closely the curve of the empirical VaR (blue line). This support the results observed in Soltyk and Gupta (2011), the FM-uMSN model was shown to provide a more accurate fit to stock returns than the FM-MN model. We observe here that further improvements can be made by fitting more flexible parametric distributions.

6 CONCLUSIONS

In this paper, we have presented several finite mixtures of asymmetric distributions for modelling asset returns. The families of these distributions are natural extensions of the normal distribution, with additional parameters to accommodate a range of non-normal features. Results of the illustrative example suggest that the accuracy of risk measures such as VaR can be improved by adopting these flexible parametric families of distributions In particular, the fitting of mixtures of skew t and generalized hyperbolic distributions was shown to outperform competing models in this example. It is also observed that the more restrictive models (the FM-MN, FM-MSN and FM-MSAL models) have fitted a two-component mixture model in order to accommodate the skewness and heavy-tails in the data, whereas only a single component is required for the more flexible distributions.

REFERENCES

Arellano-Valle, R. B. and A. Azzalini (2006). On the unification of families of skew-normal distributions. *Scandinavian Journal of Statistics 33*, 561–574.

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics 12*, 171–178.

Azzalini, A. (2005). The skew-normal distribution and related multivariate families. *Scandinavian Journal of Statistics 32*, 159–188.

Bardorff-Nielsen, O. and C. Halgreen (1977). Infinite divisibility of the hyperbolic and generalized inverse gaussian distributions. *Z. Wahrscheinlichkeitstheorie Verw. Gebiete 38*, 309–311.

Barndorff-Nielsen, O. (1977). Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society of London A 353*, 401–419.

Barndorff-Nielsen, O. (1981). Hyperbolic distributions and distributions on hyperbolae. *Scandinavian Journal of Statistics 5*, 151–157.

Browne, R. P. and P. D. McNicholas (2013). A mixture of generalized hyperbolic distributions. *arXiv:1305.1036 [stat.ME]*.

Choi, P. and I. Min (2011). A comparison of conditional and unconditional approaches in value-at-risk estimation. *Journal of the Japanese Economic Association 62*, 99–115.

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

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of Royal Statistical Society B 39*, 1–38.

Franczak, B. C., R. P. Browne, and P. D. McNicholas (2012). Mixtures of shifted asymmetric laplace distributions. *arXiv:1207.1727 [stat.ME]*.

Genton, M. G. (Ed.) (2004). *Skew-elliptical Distributions and their Applications: a Journey beyond Normality*. Chapman & Hall, CRC.

Genton, M. G. (2006). Discussion of: The skew-normal distribution and related multivariate families, by a. azzalini. *Scandinavian Journal of Statistics 9*, 1–37.

Hu, W. and A. Kercheval (2007). Risk management with generalized hyperbolic distributions. *Proceedings of the Fourth IASTED International Conference on Financial Engineering and Applications*, 19–24.

Karlis, D. and A. Santourian (2009). Model-based clustering with non-elliptically contoured distributions. *Statistics and Computing 19*, 73–83.

Kupiec, P. (1995). Techniques for verifying the accuracy of risk management models. *Journal of Derivatives 3*, 73–84.

Lee, S. and G. J. McLachlan (2013a). Finite mixtures of multivariate skew t-distributions: some recent and new results. *Statistics and Computing*, DOI 10.1007/s11222–012–9362–4.

Lee, S. X. and G. J. McLachlan (2013b). EMMIX-uskew: An R package for fitting mixtures of multivariate skew t-distributions via the EM algorithm. *Journal of Statistical Software*. To appear. Preprint arXiv:1211.5290.

Lee, S. X. and G. J. McLachlan (2013c). On mixtures of skew-normal and skew t-distributions. *Advances in Data Analysis and Classification*, DOI 10.1007/s11634–013–0132–8.

Pyne, S., X. Hu, K. Wang, E. Rossin, T.-I. Lin, L. M. Maier, C. Baecher-Allan, G. J. McLachlan, P. Tamayo, D. A. Hafler, P. L. De Jager, and J. P. Mesirow (2009). Automated high-dimensional flow cytometric data analysis. *Proceedings of the National Academy of Sciences USA 106*, 8519–8524.

Sahu, S. K., D. K. Dey, and M. D. Branco (2003). A new class of multivariate skew distributions with applications to Bayesian regression models. *Canadian Journal of Statistics 31*, 129–150.

Soltyk, S. and R. Gupta (2011). Application of the multivariate skew normal mixture model with the EM algorithm to Value-at-Risk. *MODSIM 2011 - 19th International Congress on Modelling and Simulation, Perth, Australia, December 12-16, 2011*.

Venkataraman, S. (1997). Value at risk for a mixture of normal distributions: the use of quasi-Bayesian estimation techniques. *Economic Perspectives 21*, 2–13.

Wang, K., G. J. McLachlan, S. K. Ng, and D. Peel (2009). *EMMIX-skew: EM Algorithm for Mixture of Multivariate Skew Normal/t Distributions*. R package version 1.0-12.