Journal of Applied Finance & Banking, vol. 4, no. 6, 2014, 29-45
ISSN: 1792-6580 (print version), 1792-6599 (online)
Scienpress Ltd, 2014

Market Risk Measures using Finite Gaussian Mixtures

Jorge Rosales Contreras[1]

Abstract

Value at Risk (VaR) is the most popular market risk measure as it summarizes in one figure the exposure to different risk factors. It had been around for over a decade when Expected Shortfall (ES) emerged to correct its shortcomings. Both risk measures can be estimated under several models. We explore the application of a parametric model to fit the joint distribution of risk factor returns based on multivariate finite Gaussian Mixtures, derive a closed-form expression for ES under this model and estimate risk measures for a multi-asset portfolio over an extended period. We then compare results versus benchmark models (Historical Simulation and Normal) through back-testing all of them at several confidence levels. Evidence shows that the proposed model is a competitive one for the estimation of VaR and ES.

JEL classification numbers: C46, G17
Keywords: Value at Risk, Expected Shortfall, Finite Gaussian Mixture, Historical Simulation, Delta-Normal, Backtesting.

1 Introduction

According to the Basel Committee, *failure to capture major on- and off-balance sheet risks ... was a key destabalising factor during the crisis*. In response to the detected shortcomings in capital requirements, {*the enhanced treatment introduces a stressed Value at Risk (VaR) capital requirement* (see BCBS (2011), paragraphs 11 and 12).
VaR, the most used market risk measure to estimate daily potential losses in either trading or investment books, was not able to grasp the extent of the sub-prime mortgage market collapse in the United States that triggered aggregated losses in market value over 130 billion (from February 2007) for firms such as Citigroup, Merryl Linch, Morgan Stanley, UBS, among many others.
This was mainly due to calculations based on historical simulations (heavily dependent on sample window) or debatable assumptions whose validity was often not even verified.

---

[1]EGADE Business School, ITESM CCM.

Article Info: *Received :* July 31, 2014. *Revised :* August 28, 2014.
Published online : November 1, 2014

In spite of the above, VaR is still the most favored metric by institutions and regulators to monitor and control market risk (see, for instance, CNBV 2005) and the Basel Committee uses it to set minimum capital requirements. This Committee, however, has recently agreed to move from VaR(99%) to ES(97.5%) (see BCBS 2013).

In the context of risk management, Behr and Poetter (2009) model ten European daily stock indexes returns using hyperbolic, logF and mixtures of Gaussian distributions and conclude that the fit of the latter is slightly superior for all countries. Tan and Chu (TanChu2012) model the returns of an investment portfolio using a Gaussian Mixture and estimate Value at Risk. Kamaruzzaman et.al. (2012) fit a two-component Gaussian Mixture to univariate monthly log-returns of three Malaysian stock indexes. In a different work (2013) they estimate VaR and ES (using an expression that is a particular case of equation (8) below) for monthly and weekly returns of and index and find, through backtesting that GM is an appropriate model. Zhang and Cheng (2005) use Gaussian Mixtures with different number of components to estimate VaR of Chinese market indexes, bound it with the VaR of the components and link it to the behaviour of price movements and psychologies of investors.

Alexander and Lazar (2006) use the normal mixture GARCH(1,1) model for exchange rates. They find that a two-component model performs better than those with three or more components and better than Student's *t*-GARCH models.

Haas et al (2004) introduce a general class of normal mixture GARCH(p,q) models for a stock exchange index. Their models have very flexible individual variance processes but at the cost of parsimony: their best models require from 17 to 22 parameters to model the returns of only one index.

Hardy (2001) fits a regime-switching lognormal model to monthly returns of two equity indexes and estimates VaR and ES using the payoff function of a European put option written on an index.

Several other distributions have been used to model risk factors returns, such as non-symmetric *t* distribution (Yoon and Kang 2007) or Generalized Error Distribution (see Theodossiou 2000).

We propose the family of finite Gaussian Mixtures (GM) as an alternative model to fit risk factors returns distributions and estimate risk metrics. The GM family preserves parsimony of the usual parametric models while explicitly capturing high volatility episodes through at least one of the components. We fit the portfolio profit and loss distribution and then estimate VaR and ES at several confidence levels using three models: a non-parametric one based on the empirical distribution of the risk factors returns (Historical Simulation: HS) and two parametric models; one based on the Normal distribution (Delta-Normal) and another one based on the GM family (Delta-GM).

This paper is organized as follows. In Section 2 we introduce finite Mixture distributions in general and finite Gaussian Mixtures in particular and review some of their properties. In Section 3 we construct the portfolio loss random variable and its distribution as a linear function of risk factor[2] returns. We formally define VaR and ES and introduce their estimators under the three candidate models. A description of backtesting procedures for each metric closes that section. In Section 4 we propose a trial portfolio, estimate VaR and ES at different confidence levels for several years and back-test models under study.

[2]By *risk factors* we understand the variables that determine the market value of the asset, specified through a valuation model.

In Section 5 we outline conclusions and potential future work. The Appendix contains proof and derivation of expressions used in Section 3.

2 Finite Gaussian Mixtures

In this section we introduce the family of mixture distributions and review some properties of finite Gaussian Mixtures in both the univariate and multivariate cases.

Definition 2.1 *Let $\boldsymbol{X} \in \mathbb{R}^d$ be a random vector. We say that it follows a finite (g-component) mixture distribution if its density function can be written as.*

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \sum_{j=1}^{g} \pi_j f_j(\boldsymbol{x})$$

where f_j: $\mathbb{R}^d \to \mathbb{R}^+$, $j=1, \dots, g$ are density functions and π_j, $j=1, \dots, g$ are positive constants such that

$$\sum_{j=1}^{g} \pi_j = 1$$

Let us assume that the random vector $\boldsymbol{X}$, is defined over a sample space Ω and follows a g-component mixture distribution. An intuitive interpretation is that there exists a partition $\{\Omega_1, \Omega_2, \dots, \Omega_k\}$ of the sample space Ω, where $\pi_j = \Pr[\Omega_j]$, $j=1, \dots, g$. Densities in the mixture (f_j, $j \in \{1, \dots, g\}$) correspond to conditional probability densities of $\boldsymbol{X}$ given Ω_j, $j \in \{1, \dots, g\}$ respectively. In this case, the posterior probability of Ω_j given a realization $\boldsymbol{x}$ of $\boldsymbol{X}$, is

$$P(\Omega_j | X = \boldsymbol{x}) = \frac{\pi_j f_j(\boldsymbol{x})}{\sum_{i=1}^{g} \pi_i f_i(\boldsymbol{x})}$$

Definition 2.2 *We say that a random vector $\boldsymbol{X} \in \mathbb{R}^d$ follows a finite Gaussian Mixture distribution if its density function is a mixture of d-variate normal densities:*

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \sum_{j=1}^{g} \pi_j \frac{1}{(2\pi)^{d/2} |\Sigma_j|^{1/2}} exp\left\{-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu}_j)' \Sigma_j^{-1} (\boldsymbol{x} - \boldsymbol{\mu}_j)\right\}$$

where π_j, $j=1, \dots, g$ are as in the previous definition, $\mu_j \in \mathbb{R}^d$ and $\Sigma_j \in \mathbb{R}^{dxd}$ are positive definite matrices for each $j=1, \dots, g$.

Due to linearity of the integral, Definitions 2.1 and 2.2 may be written in terms of cumulative distributions functions, instead of densities. Besides that, the family of finite Gaussian Mixture distributions displays the following properties:

- it encompasses the Normal distribution (with $g=1$),
- it is very flexible: a g-component univariate Gaussian Mixture distribution can be defined using up to $3g-1$ parameters, and it can be used to model a continuous

distortion of the normal -skewness, leptokurtosis, contamination models, multi-modality, etc- often with $g=2$ only (see McLachlan and Peel 2000).

- it is not difficult to simulate, so it can be used in Monte Carlo or bootstrap processes.
- it matches financial stylized facts (as opposed to other distributions like Student *t* or hyperbolic), markedly market volatility regimes.
- it is closed under convolution.

The last property is very important and will be used in Section 3.2 to obtain aggregated risk measures. Since it inherits this property from the Normal distribution, we state it for both distributions and assign them a number for later reference. The proof makes use of characteristic functions (see McNeil, et. al. 2005).

Property 2.3 *(Normal case) If* $\boldsymbol{X} \sim N_d(\mu,\Sigma)$ *and* $l(\boldsymbol{x}) = -(c+\boldsymbol{w}'\boldsymbol{x})$, *then* $l(\boldsymbol{X}) \sim N(\mu_l,\sigma_l^2)$, *with* $\mu_l = -(c+\boldsymbol{w}'\mu)$ *and* $\sigma_l^2 = \boldsymbol{w}'\Sigma\boldsymbol{w}$.

Property 2.4 *(Gaussian Mixture case) If* $\boldsymbol{X} \sim GM_d\,(\pi,\{\mu_j\}_{j=1,\ldots,g}, \{\Sigma_j\}_{j=1,\ldots,g})$, π, $\mu_j \in \mathbb{R}^d$, $\Sigma_j \in \mathbb{R}^{dxd}$, $j=1,\ldots,g$ *and* $l(\boldsymbol{x}) = -\boldsymbol{w}'\boldsymbol{x}$, *then* $l(\boldsymbol{X}) \sim GM(\pi,\{\mu_{lj}\}_{j=1,\ldots,g}, \{\sigma_{lj}^2\}_{j=1,\ldots,g})$, *with* $\mu_{lj} = -\boldsymbol{w}'\mu_j$ *and* $\sigma_{lj}^2 = \boldsymbol{w}'\Sigma_j\boldsymbol{w}$, *for each* $j=1, \ldots, g$.

Regarding estimation, we can obtain parameter estimators through the usual methods of moments or maximum likelihood. Lopez de Prado and Foreman (2013) introduce a method that exactly fits the first three sample moments. On the other hand, the likelihood equation (written for a univariate g-component Gaussian Mixture)

$$\frac{\partial}{\partial\boldsymbol{\theta}}\ln L(\boldsymbol{\theta}) = \sum_{j=1}^{n}\ln\left[\sum_{i=1}^{g}\pi_i f_i(\boldsymbol{y_j};\mu_i,\sigma_i)\right] = \boldsymbol{0}$$

does not admit a closed-form solution. So it is necessary to use a numerical algorithm to solve the equation for the parameters $\theta = (\pi,\{\mu_j\}_{j=1,\ldots,g}, \{\Sigma_j\}_{j=1,\ldots,g})$ and obtain Maximum Likelihood Estimators. For this purpose, we favor the EM algorithm published by Dempster, Laird and Rubin (1977). For details on the EM algorithm, see McLachlan and Krishnan (1997) and about its application to Gaussian mixtures, see McLachlan and Peel (2000).

3 Loss Distribution and Risk Measures

In this section we derive the -aggregated- Portfolio Loss Distribution following the lines of McNeil, et al (2005). We then linearize the loss function through a loss operator that is approximately equal to it for small changes in the underlying risk factors. Finally we formally define both market risk measures to be calculated on the loss distribution and introduce their estimators under three different models.

3.1 Portfolio Loss Distribution

Given a portfolio of assets subject to market risk, consider the aggregated -profit and- loss random variable for the time interval $[t\Delta,(t+1)\Delta]$:

$$L_{[t\Delta,(t+1)\Delta]} = L_{t+1} = -(V_{t+1} - V_t) = -[f(t+1, \boldsymbol{Z}_t + \boldsymbol{X}_{t+1}) - f(t, \boldsymbol{Z}_t)]$$

Where

1) $L_{[t\Delta,(t+1)\Delta]}$ is the loss over the time interval $[t\Delta,(t+1)\Delta]$,
2) Δ is the time horizon (we will assume that t is measured in days and that $\Delta=1$), therefore
3) L_{t+1} is the portfolio loss from day t to day $t+1$,
4) $V_t = f(t, \boldsymbol{Z}_t)$ is the portfolio market value at time t,
5) $f: \mathbb{R}^+ \times \mathbb{R}^d \rightarrow \mathbb{R}$ is a measurable function,
6) $\boldsymbol{Z}_t \in \mathbb{R}^d$ is the d-dimensional vector of risk factors at time t,
7) $\boldsymbol{X}_t = \boldsymbol{Z}_t - \boldsymbol{Z}_{t-1}$ contains the risk factor returns from $t-1$ to t.

According to the above definitions, losses are positive and profits are negative. As the value of risk factors $\boldsymbol{Z}_t$ is known at time t, the loss distribution is completely determined by the distribution of risk factor returns in the following period. It is then possible to define the loss operator $l_t: \mathbb{R}^d \rightarrow \mathbb{R}$ that maps risk factor returns into portfolio loss:

$$l_t(x) = -[f(t+1, Z_t + x) - f(t, Z_t)], x \in \mathbb{R}^d \tag{1}$$

Observe that $L_{t+1}=l_t(\boldsymbol{X}_{t+1})$. If the function f is differentiable, it is possible to write the linear approximation (*delta* in derivatives nemotecnia) of the loss operator l_t in equation (1) as

$$l_t^{\Delta}(\boldsymbol{x}) = -\left[f_t(t, \boldsymbol{Z}_t) + \sum_{j=1}^{d} f_{Z_j}(t, \boldsymbol{Z}_t) x_j\right] = -\left(\boldsymbol{c}_t + \boldsymbol{\omega}_t' \boldsymbol{x}\right) \tag{2}$$

where
$c_t = f_t(t, \boldsymbol{Z}_t) \approx 0$ for small time increments, such as one day,
$\omega_t' = (f_{Zj}(t, \boldsymbol{Z}_t))_{j=1,\ldots,d}$ is the vector of risk factor sensitivities, and
$f_u(t,\cdot)=\partial f(t,)/\partial u$.

If the function f has non-vanishing second-order derivatives, the approximation (2) can include them, producing a *Delta-Gamma* model. The loss operator (and random variable) moments are, from equation (2):

$$E(L_{t+1}) \approx E l_t^{\Delta}(\boldsymbol{X}) = -\sum_{j=1}^{d} f_{Z_j}(t, \boldsymbol{Z_t}) E X_{t+1,j} = \boldsymbol{\omega}_t{}' \boldsymbol{\mu} = \mu_L$$

$$Var(L_{t+1}) \approx Var\left(l_t^{\Delta}(\boldsymbol{X})\right) = Var\left(\sum_{j=1}^{d} f_{Z_j}(t, \boldsymbol{Z_t}) E X_{t+1,j}\right) = \boldsymbol{\omega}_t{}' \boldsymbol{\Sigma} \boldsymbol{\omega}_t = \sigma_L^2 \tag{3}$$

with $\boldsymbol{\mu}' = (EX_{t+1})_{j=1}^{d}$ and $\Sigma_{ij} = cov(X_{t+1,i}, X_{t+1,j})$.

In what follows we will assume that returns $\boldsymbol{X}_t$ come from a stationary process to ease notation, that is, they are independent and identically distributed (iid) random vectors and so we can omit the t subscript.

3.2 Market Risk Measures

Market Risk measures to be estimated are Value at Risk and Expected Shortfall, as defined below, according to McNeil et. al. (2005).

Definition 3.1 (VaR□). *Let L be a -positive- loss random variable and F_L:ℝ→[0,1] its distribution function. We define Value at Risk at confidence level $\alpha \in (0,1)$ as*

$$VaR_\alpha := \inf_{u \in \mathbb{R}} \{F_L(u) \geq \alpha\}.$$

Definition 3.2 (ES□). *Let L and F_L be as above. Suppose also that $E|L|<+\infty$. Expected Shortfall at confidence level $\alpha \in (0,1)$ is defined as*

$$ES_\alpha := \frac{1}{1-\alpha}\int_\alpha^1 VaR_u du$$

It is clear that VaR_α is just the α-quantile of the loss distribution: $VaR_\alpha = q_\alpha(F_L) = F_L^{-1}(\alpha)$ and that ES_α is the average of all quantiles above confidence level α, as long as the loss distribution is continuous.

Acerbi and Tasche (2002) provide a generalized *ES* definition that includes the case of non-continuous loss distributions (such as the empirical distribution), introducing a term to correct the bias of VaR_α as an estimator of the α -quantile:

Definition 3.3 (GES□). *Let L and F_L be as in Definition 3.2. Generalized Expected Shortfall at confidence level $\alpha \in (0,1)$ is*

$$GES_\alpha := \frac{1}{1-\alpha}\left[\int_\alpha^1 VaR_u du + q_\alpha\big(1-\alpha-P(L \geq VaR_\alpha)\big)\right]$$

For continuous distributions Definitions 3.2 and 3.3 coincide and the following proposition provides a useful tool for calculations. The proof is in the Appendix A.1.

Proposition 3.4. *If L is a loss random variable with continuous distribution function F_L and $E|L|<\infty$, then*

$$ES_\alpha = E[L|L > VaR_\alpha] \tag{4}$$

If the distribution of *L* is a location and scale distribution, VaR calculation depends only on the moments described in equations (3):

$$VaR_\alpha = \boldsymbol{\omega}'\boldsymbol{\mu} + q_\alpha(\boldsymbol{\omega}'\boldsymbol{\Sigma\omega})^{1/2} = \mu_L + q_\alpha\sigma_L \tag{5}$$

where q_α is the quantile in Definition 3.1 for a distribution function F_L with location parameter zero and scale parameter one.

Property 2.3 guaranties that under the Delta-Normal model, the random variable *L* follows a univariate Normal distribution and in this case equation (5) provides our VaR estimator. For the non-parametric model (HS) the distribution of *L* is the empirical distribution and it suffices to take the appropriate order statistic to obtain

$$\widehat{VaR}_\alpha = L_{(\lfloor n\alpha \rfloor)} \tag{6}$$

where $L_{(j)}$ is the *jth* order statistic, *n* is the sample size, and $\lfloor x \rfloor$ is the biggest integer that is less than or equal to *x*. Finally, for the Delta-GM model, Property 2.4 ensures that the distribution of *L* is a univariate finite Gaussian Mixture and it is necessary to solve for $q_{\square}$the following equality:

$$F_L(q_\alpha; \boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\sigma}) - \alpha = 0 \tag{7}$$

With respect to ES, the Appendix contains derivations of closed expressions for the estimator for both parametric models under consideration, whereas for the HS model it is built using the empirical distribution and Definition 3.2 or 3.3 together with expression (6). Final formulas for each model are

$$\begin{array}{lll} \text{HS} & \widehat{ES}_\alpha = & \frac{1}{n-\lfloor n\alpha \rfloor}\sum_{j=\lfloor n\alpha \rfloor+1}^{n} L_{(\lfloor j \rfloor)} \\ \text{Delta\_Normal} & ES_\alpha = & \mu + \frac{\sigma}{1-\alpha}\phi\left(\Phi^{-1}(\alpha)\right) \\ \text{Delta-MG} & ES_\alpha = & \frac{1}{1-\alpha}\sum_{j=1}^{g}\left[\pi_j \Phi\left(-z_{j,\alpha}\right)\left(\mu_j + \sigma_j \frac{\phi(z_{j,\alpha})}{\Phi(-z_{j,\alpha})}\right)\right] \end{array} \tag{8}$$

where $z_{j,\alpha}=(q_\alpha - \mu_j)/\sigma_j$ and $F_L(q_\alpha)=\alpha$.

3.3 Backtesting and Model Comparison

Once risk figures are systematically estimated over time, the performance of the estimation model can be monitored. This process of monitoring is known as *backtesting* and can also be used to compare performance among different models, as suggested by Campbell (2005).

Let us assume that for each time *t* we have one-period α-level estimations for VaR and ES, denoted $\widehat{VaR}_\alpha$ and $\widehat{ES}_\alpha$, respectively. We can then define excess indicator and excess loss random variables

$$\begin{aligned} &\hat{1}_\alpha(L_{t+1}) := 1_{(\widehat{VaR}_{\alpha,t},+\infty)}(L_{t+1}) \text{ and} \\ &\widehat{M}_{\alpha,t+1}(L_{t+1}) := \left(L_{t+1} - \widehat{ES}_{\alpha,t}\right)1_\alpha(L_{t+1}) \end{aligned} \tag{9}$$

where $1_A(u)$ is the indicator function of the set *A*. The process $\{1_\alpha(L_t)\}_{t\in\mathbb{N}}$ is a process of iid Bernoulli random variables with success probability $1-\alpha$. After estimating VaR figures for times $t=1, \ldots, n$, we can expect that

$$\sum_{t=1}^{n} \hat{1}_\alpha(L_t) \sim Bin(n, 1-\alpha)$$

Kupiec (1995) constructs a test for $H_0{:}p=p_0$ *vs* $H_A{:}p\neq p_0$ based on the likelihood ratio as test statistic. Asymptotically, this statistic follows a chi-square distribution with one degree of freedom, but for any given sample size exact rejection regions can be calculated for the binomial distribution, as shown in Casella and Berger (2002) based on work by Casella (1986). We have written a Matlab function that implements the exact test at a confidence level equal to that of the corresponding VaR estimation and returns a non-rejection interval.

Turning now to ES, we should expect that excess losses behave like realizations of iid variables from a distribution with mean zero and an atom of probability mass of size α at zero. Testing the latter property is equivalent to backtesting VaR, so we will concentrate on a procedure to test the former: zero-mean behaviour.

Recall first the one sample test under normal population for $H_0{:}\mu=\mu_0$. This can be conducted using the test statistic $z=\sqrt{n}(\mu-\mu_0)/\sigma$, which follows a normal distribution if σ is known or a Student t distribution if it is estimated.

Efron and Tibshirani (1994) propose a non-parametric bootstrap hypothesis test based on the empirical distribution of the above test statistic under the null hypothesis.

We will use the non-parametric bootstrap test for HS and a parametric bootstrap version for the Delta-Normal and Delta-GM models.

The bootstrap test is conducted as follows: draw N samples of size n with replacement from $\{m_{\square,}(l_t)\}_{t=1,\ldots,n}$, as defined in (9), or from the fitted parametric distribution and for each bootstrap sample, say $m_1, \ldots, m_n$, compute the statistic

$$t(m) = \frac{\bar{m}}{s/\sqrt{n}} \tag{10}$$

where s is the standard deviation of the bootstrap sample. The Achieved Significance Level (ASL) for the alternative hypothesis $H_A{:}\mu>0$ is estimated as

$$\widehat{ASL}_{boot} = \#\{t(m) > t_{obs}\}/N \tag{11}$$

where $t_{obs} = t(\boldsymbol{m}_{\square})$ is the value of the statistic (10) observed in the original sample. We test against a one-sided alternative based on the evidence of lack of symmetry of $\boldsymbol{m}_{\square}$ (see Figure 2).

As noted by Efron and Tibshirani (1994), the estimate ASL_{boot} has no interpretation as an exact probability, but like all bootstrap estimates is only guaranteed to be accurate as the sample size goes to infinity.

4 VaR and ES estimation in Πractice

In this section we propose a portfolio of assets with exposure to the three usual risk factor classes (interest rates, equities and foreign exchange). We then fit multivariate Normal and Gaussian Mixtures distributions to the historical daily risk factor returns (using the EM algorithm to maximize the likelihood of the latter). From daily sensitivities to each risk factor and assumed distributions, we estimate market risk measures (VaR and ES) for both parametric models (Delta-Normal and Delta-GM) as well as for the empirical distribution (HS model) at three different confidence levels (95, 97.5 and 99%) for each asset and the portfolio, for 1700 consecutive days (from July 2007 until March 2014). Finally we compare models through *backtesting* for each risk figure and asset.

Table 1: Portfolio Description

Asset	Instrument	Face value (MXN mln) or mln shares
FX	USDMXN	50
Equity	Naftrac02	10
Bond	Cetes185d	15 000

The proposed portfolio contains three assets: a short position of 50 million in US dollars (USD, this can be thought of as a debt), 15000 million face value (in MXN) of a Mexican sovereign zero coupon bond maturing in 6 months (Cetes), and 10 million shares of Naftrac02. This is an Exchange Traded Fund (ETF) that replicates the performance of Mexican Stock Exchange Index (IPC). For the sake of simplicity, it will be treated as an individual common share and not as a fund. Table 1 shows the main features of selected assets, while Table 2 displays market values and risk factor sensitivities as of April 30, 2013, under the assumption that losses are positive. For VaR and ES estimation, sensitivities are updated for each historical scenario.

Table 2: Portfolio Market Value and Sensitivities as of Apr 30, 2013

Instrument	Precio (MXN)	MtM (MXN mln)	Sensitivity	Value (MXN mln)
USDMXN	12.1401	-607.005	Delta FX	-607.005
Naftrac02	42.26	422.600	Delta IPC	422.600
Cetes185d	9.8096	14 714.463	DV01 (MXN rate)	-0.742

In order to estimate risk measures under HS, as well as parameters of parametric distributions (Normal and Gaussian Mixture) for each historical scenario, we took samples of 1000 daily returns (approximately 4 years) from USDMXN foreign exchange (FX), Mexican 6-month sovereign rate and Naftrac02. Estimators for the normal distribution are the usual unbiased estimators based on maximum likelihood. In the case of GM, we have implemented the EM algorithm in VBA for Microsoft Excel and fitted a tri-variate Gaussian Mixture with two components. Table 3 and Table 4 show an example of estimators for Normal and Gaussian Mixture distributions, correspondingly, with standard errors in parenthesis.

Table 3: ($x10^{-4}$) with Standard Errors for Normal

$\boldsymbol{\mu}$			$\boldsymbol{\Sigma}$		
			0.7	-0.6	47.1
			(0.03)	(0.03)	(8.3)
0.8	2.9	-2 937.4		2.3	-74.3
(2.3)	(4.1)	(972.4)		(0.09)	(14.8)
					126 604.1
					(4 894.8)

Note that with obtained Gaussian Mixture distribution estimators, the stylized two-component case interpretation holds: the first component describes the behaviour of the risk factor returns under the business as usual regime, while the second component describes it under stressed times, so that its mean is well-separated and its variance is significantly higher than that of the first component. If we take the USDMXN risk factor, for instance, over the sample time span MXN experienced an average daily depreciation of 0.008%, which can be decomposed into two regimes: a slight daily appreciation of 0.013% under business as usual (81% of the time), with and annual volatility of 8.97%($=\{3.2*10^{-5}*250\}^{1/2}$), and a daily depreciation of 0.098% for the remaining 19% of the time, with and annual volatility of 24.13%($=\{2.33*10^{-5}*250\}^{1/2}$), 2.7 times the volatility under the business as usual regime.

Table 4: Example of Estimators ($x10^{-4}$) for Gaussian Mixture

j	π_j	μ_j			Σ_j		
					0.32	-0.21	3.71
					(0.000)	(0.000)	(0.004)
	8111	-1.30	7.82	-487		0.85	-0.31
1	(1.14)	(0.004)	(0.004)	(0.190)		(0.000)	(0.005)
							26 770
							(8.1)
					2.33	-2.43	245
					(0.001)	(0.001)	(0.114)
	1889	9.83	-18.84	-13 460		8.22	-419
2	(1.14)	(0.012)	(0.002)	(7.050)		(0.004)	(0.228)
							541 220
							(269.2)

Another important feature is that under any of the parametric assumptions the mean of the daily portfolio -profit and- loss distribution is the same (MXN 290 906), while standard deviations for both distributions are quite similar: MXN 10.708 million under normality and MXN 10.699 million under GM. This means that the Gaussian Mixture model does not modify neither the mass center nor the dispersion of the returns joint distribution, but only decomposes them into components, while showing a higher kurtosis.
We now turn to risk estimation under the three models (HS, Delta-Normal and Delta-GM) at three confidence levels (95, 97.5, and 99%) for each asset and the portfolio. To obtain portfolio risk measures, in each historical scenario we take the weighting vector ω to be sensitivities calculated as shown in the last column of Table 2.
According to Definition 3.1, VaR has been estimated as the corresponding quantile of the loss distribution. Calculations are straightforward for both the empirical distribution and Normal assumption (equations (5) and (6)), but not for the Gaussian Mixture. For this, we have developed a Matlab code to estimate any given quantile for a univariate Gaussian Mixture with and arbitrary number of components using equation (7). Table 5 displays average VaR figures over 1700 scenarios for each instrument and the portfolio under the three considered models.

Table 5: Average VaR(99%) (figures in MXN mln)

Method	USDMXN	Naftrac02	Cetes	Portfolio
HS	13.461	14.034	7.148	27.695
Delta-Normal	11.560	10.961	6.462	21.803
Delta-GM	15.226	13.961	8.538	29.936

As for ES, equation (8) provides closed-form expressions for its calculation under the three models. Table 6 averages ES(97.5%) for each instrument and the portfolio over the 1700 historical scenarios for each one of the models.

Table 6: Average ES(97.5%) (MXN mln)

Method	USDMXN	Naftrac02	Cetes	Portfolio
HS	15.805	13.917	8.102	29.205
Delta-Normal	11.616	11.016	6.494	21.911
Delta-GM	15.283	14.011	8.555	30.004

To assess the performance of the different models and discriminate among them, we have conducted backtesting for VaR and ES following the procedures described in Section 3.3. Table 7 shows the number of times loss in any given day exceeded estimated VaR the day before ($L_{t+1} > VaR_{a,t}$) over 1700 scenarios for each confidence level, asset and model. The first column also shows non-rejection intervals at corresponding confidence level. We have written in *italics* the violations, whether figures were too conservative (less violations than the lower bound: risk over-estimation) or too aggressive (more violations that the upper bound: risk under-estimation). Fixed income risk is over-estimated by all models at 95 and 97.5% levels. Those are the only violations of Delta-GM model, making it the strongest one. On the other hand, Delta-Normal is the only model that under-estimates FX, equity and portfolio risks at the 99% level, making it the weakest of the three. Historical Simulation stands in the middle, due to under-estimation of FX risk at 95 and 99% levels.

Table 7: VaR Backtesting

Confidence	Model	USDMXN	Naftrac02	Cetes	Portfolio
	HS	*115*	88	*48*	79
95%	D-N	80	87	*32*	77
[68,103]	D-GM	96	93	*42*	75
	HS	54	43	*25*	40
97.5%	D-N	*59*	56	*20*	51
[29, 58]	D-GM	44	44	*18*	34
	HS	*31*	24	14	19
99%	D-N	*40*	*37*	14	*31*
[7, 17]	D-GM	19	28	7	17

Figure 1 shows historical VaR(99%) development for the three models as well as daily losses. Even though most excesses are concentrated in the months after the bankruptcy of Lehman-Brothers with 13 out of 17 for Delta-GM from September 2008 to May 2009, it is worth mentioning the speed of adjustment for this model after sudden changes in volatility. Over that period, HS shows 16 excesses while Delta-Normal experienced 22.

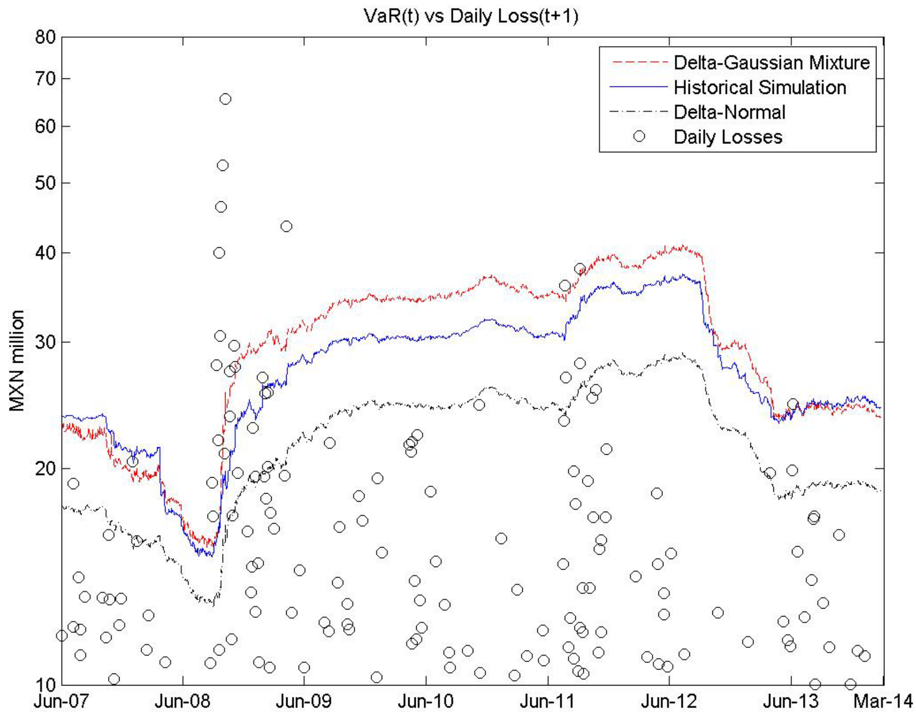

Figure 1: Backtesting VaR(99%)

Table 8 shows estimated bootstrap ASLs for each asset, model and confidence level, according to equation (11). We compare each figure against one minus the corresponding confidence level. For the Delta-Normal model, the null hypothesis that ES properly estimates average excess loss is to be rejected for every asset class and the portfolio, besides Fixed Income. HS and Delta-GM models, on the other hand, are equivalent in the sense that every time the null hypothesis is rejected for one of them, it is also rejected for the other. Moreover, the null hypothesis is rejected only in the case of Equities at 97.5 and 99% confidence levels. This is consistent with findings of McNeil and Frey (2000) for Normal and Generalized Pareto Distributions. At any other instance, ES is a reasonable estimator of average excess losses for both models.

Figure 2 displays excess losses over VaR(97.5%) for each model. Not only does the Normal model shows more excesses, but they are bigger than those of the other models. Negative excess losses are close to zero in the Normal case due to the small difference between VaR and ES (see Tables 5 and 6). HS displays more and higher excesses than GM (*t*-statistics are 0.52 and 0.19), so the latter is slightly -but not significantly- superior than the former.

Table 8: ES no-parametrical Significance Levels

Confidence level	Method	ES p-values USDMXN	Naftrac02	Cetes	Portfolio
95%	HS	0.5212	0.6518	0.4068	0.4884
	D-N	0.0054	0.0008	0.0968	0.0092
	D-GM	0.5504	0.3904	0.6884	0.7782
97.5%	HS	0.0576	0.0094	0.1438	0.0436
	D-N	0.0084	0.0008	0.0998	0.0144
	D-GM	0.0382	0.0062	0.1376	0.0500
99%	HS	0.0146	0.0014	0.0770	0.0170
	D-N	0.0064	0.0002	0.0840	0.0056
	D-GM	0.0230	0.0018	0.1216	0.0300

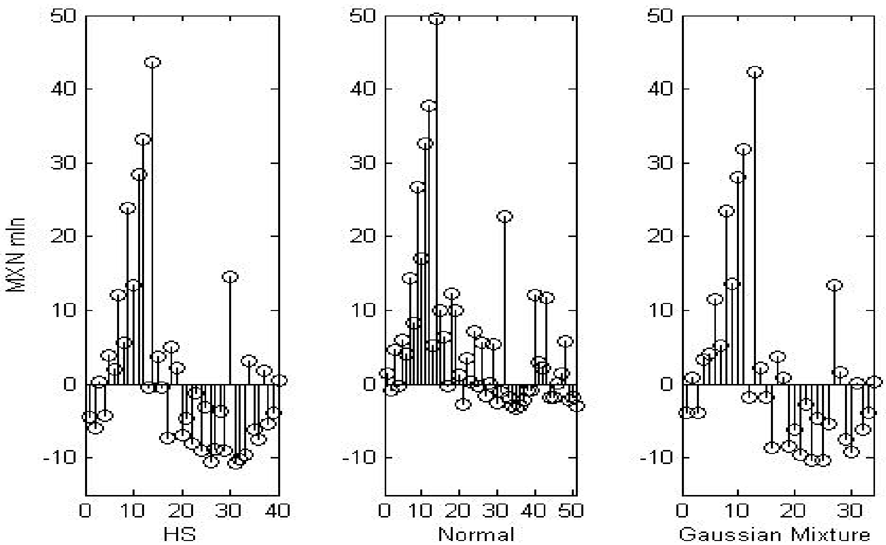

Figure 2: Backtesting ES(97.5%)

5 Conclusions

Among the three models under study, Delta-Normal is the most aggressive, in the sense that it consistently produces the smallest figures for VaR and ES. We believe, however, that its major drawback is that, having so little mass at the tail of the distribution, switching from VaR(99%) to ES(97.5%) (as proposed by Basel Committee) means a uniform increase of 0.5% in risk figures. This has the benefit of saving capital, but it can expose financial institutions to significant losses when high volatility episodes happen.

When using Historical Simulation, there is a significant adjustment of 47, 37 and 30% between VaR and ES for α=95, 97.5 and 99%. This is a confirmation of its strong dependence on the sample window, given the fact that the sample window includes the whole credit crisis time span. Over this period there existed returns much higher than the mean of the empirical distribution as well as recurrent changes in monetary policy rate that influenced short-term interest rates.

With respect to the finite Gaussian Mixture model, since it explicitly includes a component to model high volatility periods, it usually (but not always) produces the most conservative VaR figures: 10% higher than HS and 29% higher than Delta-Normal on average. Going from VaR(99%) to ES(97.5%) represents an increase of only 0.2% on average, but it fluctuates across assets and along time, as volatility changes. This is a distribution that displays excess kurtosis and can fit historical volatility to each risk factor simultaneously.

A technical but relevant detail, noted in Section 4, is that the GM model does not modify the mass center or the dispersion of the returns distribution, but only segments them into components. This implies that Lopez de Prado and Foreman's (2013) critique does not hold and therefore it is not necessary to explicitly fit sample moments with ad-hoc estimators. We then have maximum likelihood estimators, with their advantage over moment estimators, that in turn perfectly fit the first sample moment and quite well the second one; making it unnecessary to compute higher moments.

We believe that we have shown strong evidence that the finite Gaussian Mixture model is appropriate to estimate tail risk measures in the context of changing volatility. We have, however, based our model on a stationary assumption for the returns distribution (or equivalently, the one-period loss distribution). We should now relax this assumption and fit a regime-switching model to test whether adding new parameters produces more precise estimators and assess its impact on parsimony.

References

[1] Acerbi, C. and D. Tasche, On the coherence of Expected Shortfall, Dept of Math TU-München *Working Paper*, (2002).

[2] Alexander, C. and E. Lazar, Normal Mixture GARCH(1,1): Applications to Exchange Rate Modeling, *Journal of Applied Econometrics*, **21**,(2002), 307-336.

[3] Behr, A. and U. Poetter, Modeling Marginal Distributions of Ten European Stock Market Index Returns, *International Research Journal of Finance and Economics*, (28), (2009), 104-119.

[4] Basel Comitte on Banking Supervision, *Basel III: A Global Regulatory Framework for more Resilient Banks and Banking Systems*, available at www.bis.org/publ/bcbs189.pdf, 2011.

[5] Basel Comitte on Banking Supervision, *Fundamental Review of the Trading Book: A Revised Market Risk Framework*, available at www.bis.org/publ/bcbs265.pdf, 2013.

[6] Campbell, M., S. Mittnik and M.S. Paolella, Mixed Normal Conditional Heteroskedasticity, *Journal of Financial Econometrics*, **2**, (2005), 211-250.

[7] Casella, G., Refining Binomial Confidence Intervals, *Canadian Journal of Statistics*, **14,** (1986), 113-129.

[8] Casella, G. and R. L. Berger, *Statistical Inference*, Duxbury, California, 2002.

[9] CNBV, *General Dispositions Applicable to Credit Institutions* (in Spanish), available at www.cnbv.gob.mx, Mexico, 2005.

[10] Dempster, A.P., N.M. Laird and D.B. Rubin, Maximum Likelihood from Incomplete Data via the EM Algorithm, *Journal of the Royal Statistical Society B*, **39**, (1977), 1-38.

[11] Efron, B. F. and R. J. Tibshirani, *An Introduction to the Bootstrap*, Chapman & Hall, New York, 1994.

[12] Haas, M., S. Mittnik and M.S. Paolella, Mixed Normal Conditional Heteroskedasticity, *Journal of Financial Econometrics*, **2**, (2004), 211-250.

[13] Hardy, M., A Regime-Switching Model of Long-Term Stock Returns, *North American Actuarial Journal*, **5**(2), (2001), 41-53.

[14] Kamaruzzaman, Z.A. and Z. Isa, Estimation of Value at Risk and Conditional Value at Risk Using Normal Mixture Distribution Model, *Proceedings of the 20th National Symposium on Mathematical Sciences*, **1522** (2013), 1123-1131.

[15] Kamaruzzaman, Z.A., Z. Isa and M.T. Ismail, Mixtures of Normal Distributions: Application to Bursa Malaysia Stock Market Indices, *World Applied Sciences Journal*, **16**(6), (2012), 781-790.

[16] Kupiec, P., Techniques for verifying the accuracy of Risk Management Models, *Journal of Derivatives*, **3**, (1995), 73-84.

[17] Lopez de Prado, M. and M. Foreman, A Mixture of Gaussians Approach to Mathematical Portfolio Oversight: The EF3M Algorithm, *Quantitative Finance,* forthcoming (2013).

[18] McLachlan, G. and T. Krishnan, *The EM Algorithm and Extensions*, Wiley, New York, 1997.

[19] McLachlan, G. and D. Peel, *Finite Mixture Models*, Wiley, New York, 2000.

[20] McNeil, A.J., R. Frey and P. Embrechts, *Quantitative Risk Management: Concepts,Techniques and Tools*, Princeton UP, New Jersey, 2005.

[21] McNeil, A.J. and R. Frey, Estimation of tail related risk measures for heteroscedastic financial time series: an extreme value approach, *Journal of Empirical Finance*, **7**, (2000), 271-300.

[22] Tan, K. and M. Chu, Estimation of Portfolio Return and Value at Risk using a class of Gaussian Mixture Distributions, *The International Journal of Business and Finance Research*, **6**(1), (2012), 97-107.

[23] Theodossiou, P., Skewed Generalized Error Distribution of Financial Assets and Option Pricing, School of Business, Rutgers U *Working Paper*, (2000).

[24] Yoon, S.M. and S.H. Kang, A Skewed Student-t Value-at-Risk Approach for Long Memory Volatility Processes in Japanese Financial Markets, *Journal of International Economic Studies*, **11**(1), (2007), 211-242.

[25] Zhang, M.H. and Q.S. Cheng, An Approach to VaR for Capital Markets with Gaussian Mixture, *Applied Mathematics and Computation*, (168), (2005) 1079-1085.

Appendix

Proof and Derivations

In this appendix we prove Proposition 3.4 and derive closed-form expressions for Expected Shortfall under the two studied parametric models. In Section A.2 we derive ES under Normal assumption, filling in the details of the proof that can be found in McNeil, et al (2005). The reason to include this proof is that from there we can adapt the result to obtain the corresponding expression for finite Gaussian Mixture distribution in Section A.3.

A.1 Proof of Proposition 3.4

If L is continuous with distribution function F_L and $E|X|<+\infty$, then $ES_\alpha = E[L|L>VaR_\alpha]$.

Proof. It is well-known that the random variable $U=F_L(L)$ follows a uniform distribution in $[0,1]$ with density function $f_U(u)=1$, $u \in [0,1]$; therefore

$$\int_\alpha^1 VaR_u du = \int_0^1 VaR_u 1_{(\alpha,1)}(u) f_U(u) du = E\left[VaR_U 1_{(\alpha,1)}(U)\right],$$

where $1_A(\cdot)$ is the indicator function of the set A. Using that $VaR_U=F_L^{-1}(U)$ and that continuity of F_L implies F_L^{-1} is strictly increasing:

$$\int_\alpha^1 VaR_u du = E\left[F_L^{-1}(U) 1_{\left(F_L^{-1}(\alpha), F_L^{-1}(1)\right)}\left(F_L^{-1}(U)\right)\right] = E\left[L 1_{(VaR_\alpha,+\infty)}(L)\right].$$

Dividing by 1-α, we obtain:

$$ES_\alpha = \frac{1}{1-\alpha}\int_\alpha^1 VaR_u du = \frac{E\left[L 1_{\{L \geq VaR_\alpha\}}\right]}{P(L \geq VaR_\alpha)} = E[L|L \geq VaR_\alpha].$$

A.2 ES for Normal

Let $L \sim N(\mu, \sigma^2)$ and let $q_\alpha = \mathrm{VaR}_\alpha$ be the α-quantile of F_L, i.e., $F_L(q_\alpha)=\alpha$. Let $f_L(\cdot)=\varphi(\cdot;\mu,\sigma^2)$ be the density function of L and let $\varphi(\cdot)=\varphi(\cdot;0,1)$ be the standard normal density function with α -quantile equal to z_α. Using equation (4) and the distribution of L, we have:

$$ES(\alpha) = E[L|L > VaR(\alpha)] = \frac{1}{1-\alpha}\int_{q_\alpha}^{+\infty} u f_L(u) du = \frac{1}{1-\alpha}\int_{q_\alpha}^{+\infty} u\phi(u;\mu,\sigma^2) du \quad (12)$$

with the change of variable $u = \sigma z + \mu \; (z_\alpha = (q_\alpha - \mu)/\sigma \, , du = \sigma dz)$

$$\begin{aligned} ES(\alpha) &= \frac{1}{1-\alpha}\int_{z_\alpha}^{+\infty} (\sigma z + \mu)\phi(z) dz = \frac{1}{1-\alpha}\left[\sigma \int_{z_\alpha}^{+\infty} z\phi(z) dz + \mu \int_{z_\alpha}^{+\infty} \phi(z) dz\right] \\ &= \frac{1}{1-\alpha}\left[-\sigma\phi(z) + \mu\Phi(z)\right]\Big|_{z_\alpha}^{+\infty} = \frac{1}{1-\alpha}\left[\sigma\phi(z_\alpha) + \mu\left(1 - \Phi(z_\alpha)\right)\right] \\ &= \frac{1}{1-\alpha}\left[\sigma\phi\left(\Phi^{-1}(\alpha)\right) + \mu(1-\alpha)\right] = \mu + \frac{\sigma}{1-\alpha}\phi\left(\Phi^{-1}(\alpha)\right) \quad (13) \end{aligned}$$

A.3 ES for Gaussian Mixtures

Let $L \sim GM(\pi,\mu,\sigma)$, recall that $q_\alpha = VaR(\alpha)$ is the solution of equation (7). From Definition 3.1, equation (3.4) and the distribution of L, we have

$$ES(\alpha) = \frac{1}{1-\alpha}\int_{q_\alpha}^{+\infty} u f_L(u)du = \frac{1}{1-\alpha}\int_{q_\alpha}^{+\infty} u \sum_{j=1}^{k} \pi_j \, \phi\left(u;\mu_j,\sigma_j^2\right) du$$

$$= \frac{1}{1-\alpha}\sum_{j=1}^{k} \pi_j \int_{q_\alpha}^{+\infty} u\phi\left(u;\mu_j,\sigma_j^2\right)du \qquad (14)$$

The integral within the sum is the same as (12), with the only difference that the lower limit of the integral depends on the specific component. Making the change of variable $u=\sigma_j z + \mu_j$ and defining $z_{j,\alpha}=(q_\alpha-\mu_j)/\sigma_j$ we obtain an analogous result to (13):

$$ES(\alpha) = \frac{1}{1-\alpha}\sum_{j=1}^{k} \pi_j \left[\sigma_j \, \phi\left(z_{j,\alpha}\right) + \mu_j \left(1 - \Phi\left(z_{j,\alpha}\right)\right)\right]$$

$$= \frac{1}{1-\alpha}\sum_{j=1}^{k} \pi_j \, \Phi\left(-z_{j,\alpha}\right)\left[\mu_j + \sigma_j \, \frac{\phi\left(z_{j,\alpha}\right)}{\Phi\left(-z_{j,\alpha}\right)}\right] \qquad (15)$$

Note that $z_{j,\alpha}$ depends on α through q_α and on the component through parameters μ_j and σ_j, but it is not the α -quantile of the *j-th* component distribution, that is to say, it is not the case that $\Phi(z_{j,\alpha})=\alpha$. In other words, $\mu_j + \sigma_j \, \varphi(z_{j,\alpha}) / \Phi(-z_{j,\alpha})$ is not the ES_α corresponding to the *j-th* component.

It is possible, however, to write the finite Gaussian Mixture Expected Shortfall as the weighted summation of the component-specific Expected Shortfalls. To see this, let $L_j \sim N(\mu_j, \sigma_j^2)$, then, according to Section A.2:

$$ES_\alpha\left(L_j\right) = \mu_j + \sigma_j \, \frac{\phi\left(\Phi^{-1}(\alpha)\right)}{1-\alpha}$$

$$ES_\alpha(L) = \sum_{j=1}^{k} \lambda_j \, ES_\alpha\left(L_j\right)$$

.

where, $\lambda_j = \pi_j \, \frac{\Phi\left(-z_{j,\alpha}\right)}{1-\alpha} \frac{\mu_j + \sigma_j \frac{\phi\left(z_{j,\alpha}\right)}{\Phi\left(-z_{j,\alpha}\right)}}{\mu_j + \sigma_j \frac{\phi\left(\Phi^{-1}(\alpha)\right)}{1-\alpha}}$