# Fitting Heavy Tail Distributions With Mixture Models

Jorge Basilio[1;3] and Amilcar Oliveira[1;2]

[1] Universidade Aberta, Lisboa, Portugal
(E-mail: 1600751@estudante.uab.pt)
[2] CEAUL - Centro de Estatística e Aplicações da Universidade de Lisboa, Portugal
[3] National College of Ireland, Dublin, Ireland

**Abstract.** The normal probability distribution as assumption for financial returns have been recognized as inappropriate, and a source of inaccurate estimation of Value at Risk. Empirical evidence also have been shown that financial returns shows a more accentuated leptokurtic distribution when compared with a Normal distribution and also skewed. This is usually a cause of underestimated values of *VaR*, specially when the quantiles are very low. Therefore it is necessary to focus on the tail of the distribution and identify models to fit that behavior. We will highlight the differences between the quality of fitting in the tails of the distribution and the fitting for all the distribution.

This work compares and interprets the results obtained by applying mixture models as a method to estimate the behavior on the extremes for heavy tail data distributions. This results will be then used to describe an analytical solution of *VaR* under mixture models.

**Keywords:** Mixture Models, Extreme Values, *VaR*, Risk Analyses .

## 1 Introduction

Extreme value theory is used to model unusually low or high value data that is observed in the tail of the distribution.

These unusual events represented by extreme data points are often complex to model and requiring advanced techniques to fit a distribution that includes the heavy tail with satisfactory results.

In statistics the concept of mixture distribution concerns the combination of two or more distributions.

Mixture distributions are of importance in order to model complex processes allowing for a more flexible approach than using a single distribution.

Gaussian mixtures are a possible choice to model this type of complex processes and are formed by linear combinations of two or more Gaussian distributions as a weighted sum of that Gaussian distributions in order to form a new distribution.

One promising use case to apply mixture models is to model heavy tailed distributions. Even though financial returns are usually modeled as normally distributed this assumption proved to be inconsistent with empirical evidence.

$6^{th}$ *SMTDA Conference Proceedings, 2-5 June 2020, Barcelona, Spain*

Asset returns are considered heavy tailed which means that extreme values are more likely to happen in practise than suggested by a normal distribution. This discrepancies could cause estimation errors of major impact if for example we are using those assumption to build risk measures such as *VaR*. Then normality assumption can lead to inappropriate risk management measures [10].

Some research work involving mixture models assumes that the several distributions are drawn from the same probability density functions. On other hand heterogeneous mixture models can be a valid option for modelling more complex phenomenons and improving modelling capabilities.

An option is to model asset returns distribution with a mixture of distributions, a mixture of normal, a mixture of Gaussian and Gumbel and also Gaussian and GEV as extreme value distributions.

The modelling approach presented explores the capabilities of a mixture model in order to fit the financial returns [14].

In this work we will fit the financial returns with different models and analyse the impact of each model to estimate Value at Risk (*VaR*).

## 2 Value at Risk

*VaR* is a central tool in risk, asset, and portfolio risk management and it also plays a key role in systemic risk. *VaR* is defined as the maximum loss an asset/portfolio/institution can incur in a given time period at a defined significance level $\alpha$.

This probability represents a quantile for risk. With a random variable $X$ and a distribution function $F$ that model losses, verified for an asset in a time period. $VaR_\alpha$ is defined as:

$$VaR_\alpha = F^{-1}(1-\alpha) \tag{1}$$

A *VaR* of $d$ days at $\alpha\%$ significance level means that on % of $d$ days, we won't see a loss higher than the *VaR*, but for the $(1-\alpha)\%$ of times the loss will be higher.

For example, a 1 day *VaR* at 99% confidence level of 5% means that only 1 of every 100 days we will see a loss higher that 5% of the initial capital [9].

The concept of *VaR* becomes central to the study of systemic risk, becoming also a standard concept for the definition of several of the most significant systemic risk measures mentioned in the literature. Yet this method faces challenges dealing with the risk associated with events involving a volatility component with dependencies between extremes values in distinct data sets and modelling extreme values with volatility.

## 3 Financial Institutions Returns

### 3.1 The Data Set

Institution financial details frequently are not available on the public domain and are informed only on a periodic basis, so this methodology is an option to obtain *VaR* based on market public data. As an assumption, the market value, the market capitalization of each institution reflects the book value of the assets.

The process could be summarised as follows:

- Obtain the data:
  - collect stock prices, we will use weekly based stock prices (Friday's price)
- market value of equity (MVE)
  - stock price $\times$ shares outstanding
- Assume market value of assets (MVA)
- getting returns as: $X_t^i = \frac{(MVA_t^i - MVA_{t-}^i)}{MVA_{t-}^i}$, for each financial institutions

Based on the list of banks that are part of the *STOXX Europe 600* Banks index, information on daily quotations for each title, public available for consultation at *https://finance.yahoo.com/*.

### 3.2 Normal distribution assumption

There are an open ongoing discussion about the application of normal distribution to model financial related data.

In fact, the use of of normal distribution in order to model financial returns is considered a traditional assumption in finance since Markowitz developed his portfolio theory in 1952, as it is the backbone of traditional (mean-variance) premise [4].

Despite that, some recent research work have been rejecting this assumption based on the study of skewness, kurtosis and in special in heavy tail in the distribution of financial returns [8].

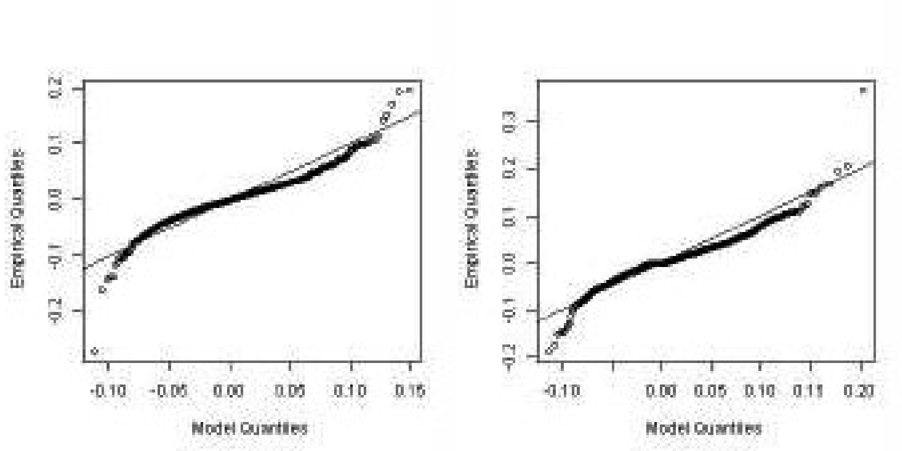
Even though is noticeable that financial returns distributions are at least close to a bell shaped curve, even if this does not translate directly for a normal distribution.

By using the graphs bellow, we can notice a bell shape pattern on the returns of financial institutions. We also can notice an heavy tail behavior, best described in this case, in terms of extreme tails by the $t$-Student curve than by the normal curve.

An additional complementary graphical analysis in order to compare the behavior of the returns across the distinct quantiles and compare that behavior with the expect behavior for a Normal population is to use a Q-Q plot.

The Q-Q plot, or quantile-quantile plot, is a type of graph that allow us assess if it is plausible to assume that a data set came from some theoretical distribution such as a Normal. Even if it is just a visual check, and somewhat subjective, it allows us to see at-a-glance if our assumption is plausible, and how the assumption is eventually violated and which data points cause that violation.

By applying a simple normality test to financial institution returns series, it seems to show up very clear that we should strongly consider other options to model the financial institution returns.




(a) HSBA returns Q-Q plot vs normal distribution (b) BBVA returns Q-Q plot vs normal distribution

Lets start by applying some conventional statistical tests for normality such as Shapiro-Wilk's test and Kolmogorov-Smirnov (K-S) normality test.

| Financial Institution | *W* statistic | *K-S* statistic | *W* p-value | *K-S* p-value |
|---|---|---|---|---|
| HSBA | 0.94344 | 0.06652 | 2.2e-16 | 0.00017 |
| BBVA | 0.93228 | 0.07789 | 2.2e-16 | 5.249e-06 |
| BARC | 0.77149 | 0.11839 | 2.2e-16 | 2.565e-13 |

Table 1: Normality testing

Based on the above results, we should consider the financial institutions returns (and financial system returns too) as not normally distributed.

However it is also know that normality tests are in fact very sensitive to what happens on in the extreme tails. This fact can then restrain all the conclusions based on those type of tests.

As we have a relatively large sample of data on our data set it will also worthwhile try a visual approach to investigate normality. Lets then compare the histogram of returns for some financial institutions and the system.

It become also clear we have a different behavior on extreme tails of returns distribution, and in the tail it is not following a normal behavior. This pattern must be included in the modelling.

### 3.3 Heavy Tails Distributions

In statistics the term heavy tail is associated to distributions with a relatively height probability of extreme outcomes.

Even thought there are not a definitive and formal definition of heavy tail usually is assumed that a distribution has a heavy tail when the probability in the tail is thicker when compared with a normal distribution.

Taken as example Cauchy distribution we can notice a ticker tail in Cauchy when compared with a normal.

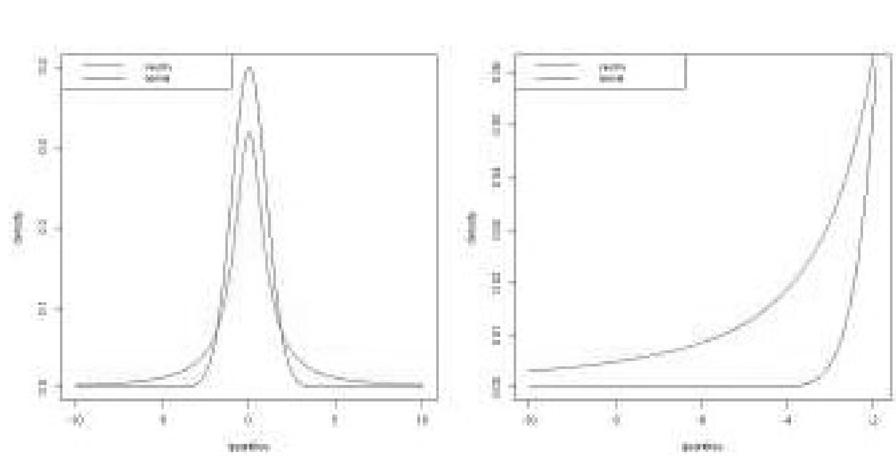


Fig. 2: Heavy tail distribution versus normal

Cauchy among others, like $t$-Student for example are know as heavy tailed distributions.

In order to close the gap in terms of modelling the extreme tail of financial institution returns as first approach we will study an approach of univariate fitting were we will include the modelling of the tails of the distribution.

To evaluate how heavy tails impacts our financial data set, we will compare results from fitting three theoretical distributions: Normal, Cauchy and $t$-Student.

### 3.4 Cauchy distribution

In probability theory, Cauchy distribution is the probability distribution whose probability density function is defined by:

$$f(x) = \frac{1}{\pi(1 + x^2)}$$

with x $\in \mathbb{R}$ in Cauchy standard form, when it mean is 0, and the and first and third quantiles are 1 and 1 respectively.

Cauchy distribution then is defined as any distribution that belongs to this family and, if $X$ is random variable with standard Cauchy distribution then, let $\mu \in \mathbb{R}$ be an arbitrary value and $\sigma > 0$. The random variable $Y$, defined as:

$$Y = \mu + \sigma X$$

also follows a Cauchy distribution with median $\mu$ and whose first quantile is $\mu - \sigma$ and the third quantile is $\mu + \sigma$. The probability density function is therefore defined as:

$$f(y) = \frac{1}{\pi\sigma(1 + \frac{(x-\mu)^2}{\sigma^2})}$$

Standard Cauchy distribution can also be defined as a ratio of two normal distributions. Let $X$ and $Y$ be two independent random variables. If $X \sim N(0,1)$ an $Y \sim N(0,1)$ then:

$$\frac{X}{Y} \sim \text{Cauchy}(0,1)$$

In addition to its application in physics, Cauchy distribution is commonly used in models in finance to represent deviations in returns from the predictive model [7]. The reason for this is that practitioners in finance are wary of using models that have light-tailed distributions as Normal, on their returns, and they generally prefer to go the other way and use a distribution with very heavy tails as Cauchy. The history of finance has a vast record of catastrophic predictions based on models that did not have heavy enough tails in their distributions. The Cauchy distribution has sufficiently heavy tails as its moments does not exist, and so it is an ideal candidate to give an error term with extremely heavy tails [5].

### 3.5 *t*-Student

Also $t$-Student is an option to deal with heavy tails and have been an option for researchers too [16].

The $t$-Student distribution can be defined as a variable $Z \sim N(0,1)$ and a variable $W \sim X_v^2$, then the standardized quotient of the two follows a $t$-Student distribution with $v$ degrees of freedom:

$$T = \sqrt{v}.\frac{Z}{W} \sim t$$

$t$-Student distribution could be very useful for financial analysis as we can adapt to the tail behavior of the data. In its conventional form, $t$-Student could not be a very flexible model because of the absence of a location and a scale parameter. An alternative definition can then be described as:

$$\text{If}\, T \sim t_v \Longrightarrow S = \mu + \lambda T \sim t_v(\mu, \lambda^2)$$

with $\mu$ as location parameter and $\lambda^2(\frac{v}{v-2})$ as scale parameter. The tail decay is therefore polynomial, that is, the density function goes to zero proportional to $x^{-(v+1)}$ for $x \to \infty$. For low values of $v$ this is a much slower rate than for the Gaussian [17].

Since the typical assumption involving the Gaussian distribution had failed and is now hardly accepted due to the probabilities at the extremes are much larger than those supported by Gaussian distributions. This invalid assumption is specially dangerous for risk management related application.

# 4 Best fit for returns

If it has been established that financial returns shows heavier tail and peaked than Gaussian it was not yet established which distribution better fit those financial return, remaining as research field. In the process of fitting the data to a theoretical distribution, one could found that usually more than one distribution would be of interest.

The importance of correctly model financial returns and identify distributions that are able to adjust and fit financial return series have been putted at the spotlight every time a new financial crises is faced. Recent crises were not exception and it become even more clear that the typical assumption with Gaussian distributions was anymore acceptable.

The financial return series, like the example below, has phases where it exhibits different volatility. While small returns occur more frequently, there are phases were positive and negative returns are persistently larger as well. This is mentioned as volatility clusters and is understood as characteristic of financial data [18].

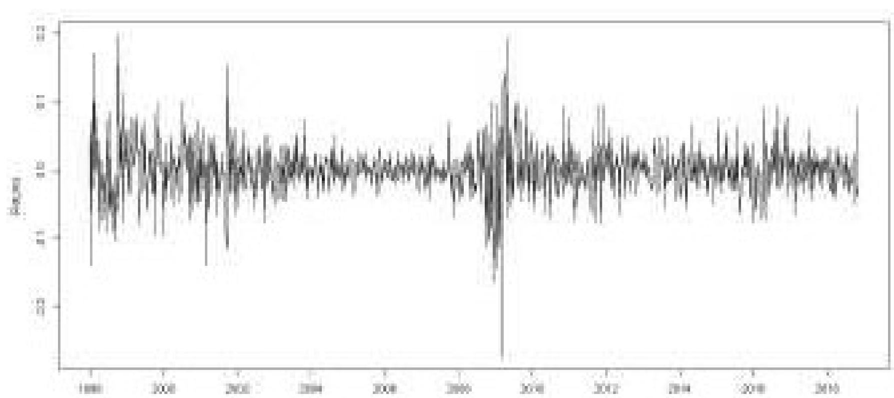


Fig. 3: HSBA weekly returns

However, when working with longer return series, we neglect the exact times of occurrence of each return, and focus on the distribution of returns only.

Starting with HSBA returns series as example, and fitting this series, after normalization to a the three above mentioned distributions: Normal, Cauchy and $t$-Student, the resulted fit are as showed in the figure:

By visually analyse this results one can notice some differences in the quality of the adjustment, depending on the quantile we take. Due to this fact it becomes harder to identify a single distribution that performs well across all quantiles.

In this particular case, if we take the global results of the fit, we can see that even though, Normal and Cauchy have clear distinct shapes, in terms of

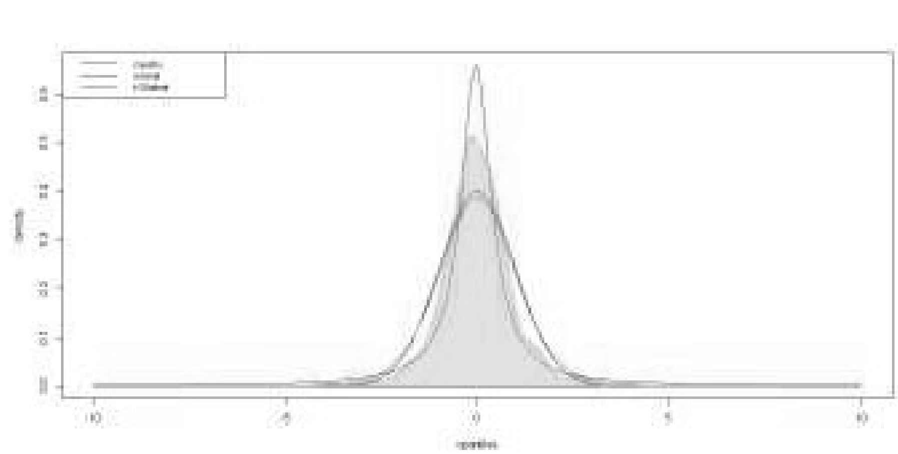


Fig. 4: Fit of HSBA weekly returns

| Goodness-of-fit statistics | Cauchy | Normal | $t$-Student |
|---|---|---|---|
| Kolmogorov-Smirnov | 0.0552 | 0.0652 | 0.0739 |
| Cramer-von Mises | 0.7421 | 1.6626 | 2.2699 |
| Anderson-Darling | 8.8937 | 10.2788 | 13.9965 |
| Goodness-of-fit criteria | | | |
| BIC | 3043.091 | 3013.615 | 2931.064 |

Table 2: Goodness-of-fit

Goodness-of-fit it turns that the results are quite similar only with a short advantage over Cauchy. However our concerns are more with the behavior on the tail of the distribution, and the following image can proportionate a closer look at the more extreme quantiles in detail:

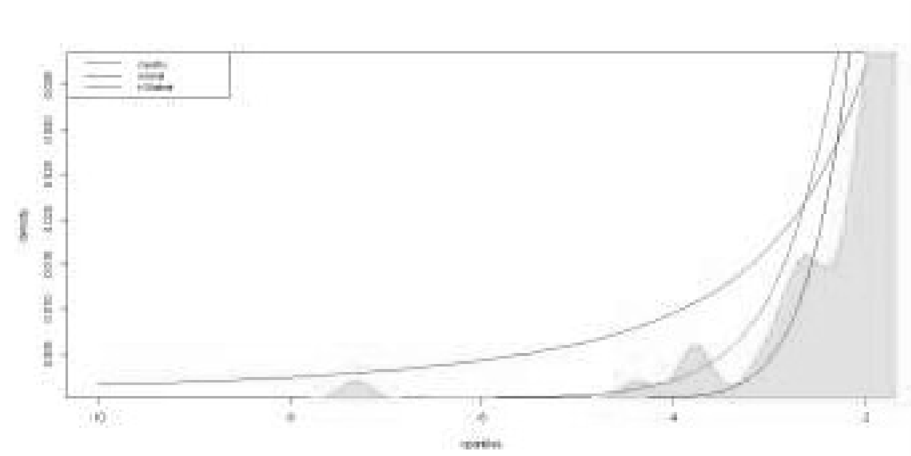


Fig. 5: Fit of HSBA weekly returns

Using the image above, is now easier to identify some limitation on this approach to model extreme values in the tails. There a relevant density on returns series that is not captured by Normal (neither by $t$-Student), and Cauchy estimate it far for excess.

### 4.1 Modelling financial returns with a mixture of normal distributions

As empirical evidence and results suggests, the normality assumption of financial institution returns is not verified as it is heavy tailed, and modelling this behavior using only one distribution have shown limitation, one option is consider an approach that uses more than one distribution.

Heavy tailed distributions can be modeled instead by a mixture of distribution. In this case we will first approach the problem by applying a mixture of normal distributions.

Assuming the returns are following a stochastic process for a financial institution $i$ as:

$$R_{it} = \lambda_{it} R_{it}^{\alpha} + (1 - \lambda_{it}) R_{it}^{\beta} \tag{2}$$

where $R_{it}^{\alpha} \sim N(\mu_\alpha, \sigma_\alpha)$, $R_{it}^{\beta} \sim N(0, \sigma_\beta)$, and $\lambda_{it}$ is 1 with probability $p$ and 0 otherwise.

These three random variables $R_{it}^{\alpha}$, $R_{it}^{\beta}$ and $\lambda_{it}$ are independent each other.

Depending on $\lambda$ and with probability $p$ the distribution to apply will be $N(\mu_\alpha, \sigma_\alpha)$, for example for the most normal situations. With probability (1 - $p$), $\lambda$ will be equal to 0 and the distribution to apply is $N(0, \sigma_\beta)$ and it could be interpreted as an exceptional case.

The challenge now is with the estimation of the parameters involve; $p, \mu_\alpha, \sigma_\alpha, \sigma_\beta$.

Despite several alternative methods that are possible to use to estimate the parameters of a mixture of normal distribution, if we consider the traditional maximum likelihood method, it could then be formulated as:

$$l\Big((p, \mu_\alpha, \sigma_\alpha, \sigma_\beta)|R_{it}\Big) = \sum_t log\Big[\frac{p}{\sigma_\alpha} exp\Big(-\frac{\big(R_t - \mu_\alpha\big)^2}{2\mu_\alpha^2}\Big) + \frac{1-p}{\sigma_\beta} exp\Big(-\frac{(R_t^2)}{-\mu_\beta}\Big)\Big]$$

Due the existence of both poles and saddle points, the maximization of the mixture of normals likelihood could be challenging and the global maximum for that function could not exist [6].

This problem however could be described as an incomplete data problem since the data we observe in our sample can be viewed as a subset of the "complete" data.

### 4.2 Expectation-Maximization Algorithm

The Expectation-Maximization (EM) Algorithm is an appropriate tool for that type of problems. EM Algorithm is an approach for maximum likelihood estimation in the presence of latent variables and can be used to predict the latent variables values with the condition that the general form of the probability distribution governing those latent variables is know.

The algorithm is implemented as an iterative procedure given a set of incomplete data and considering a set of starting parameters will iterate on two steps

- Expectation step (E – step). Using the available observed data and the current model parameters the missing or latent variables are estimated by.
- Maximization step (M – step). After estimate missing values, this step will be used to update the parameters by compute the parameters that maximize the expected log-likelihood of the model based on the values estimated on E-step.

EM Algorithm includes statistical considerations to compute the maximum-likelihood (ML), source distribution that would have created the observed data, including the effects of counting statistics. Specifically, it assigns greater weight to high-count elements of a profile and less weight to low-count regions [3].

### 4.3 EM Algorithm and mixture of Gaussians

Taken the case of a mixture of Gaussians let $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ be a sample of n i.i.d. observations of a mixture of two Gaussian and $\mathbf{z} = (z_1, z_2, \ldots, z_n)$ the latent variables that determine the component where the observation originates[15].

$$X_i|(Z_i = 1) \sim \mathcal{N}_d(\boldsymbol{\mu}_1, \Sigma_1) \quad and \quad X_i|(Z_i = 2) \sim \mathcal{N}_d(\boldsymbol{\mu}_2, \Sigma_2),$$

where

$$\mathrm{P}(Z_i = 1) = \tau_1 \quad \mathit{and} \quad \mathrm{P}(Z_i = 2) = \tau_2 = 1 - \tau_1$$

The goal of this process is to estimate the parameters for the mixture of Gaussians:

$$\theta = \big(\boldsymbol{\tau}, \boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \Sigma_1, \Sigma_2\big)$$

The likelihood function therefore is:

$$L(\theta; \mathbf{x}, \mathbf{z}) = \exp\left\{ \sum_{i=1}^{n} \sum_{j=1}^{2} \mathbb{I}(z_i = j) \big[ \log \tau_j - \tfrac{1}{2} \log |\Sigma_j| - \tfrac{1}{2} (\mathbf{x}_i - \boldsymbol{\mu}_j)^\top \Sigma_j^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_j) - \tfrac{d}{2} \log(2\pi) \big] \right\}.$$

The result of the application of EM Algorithm to HSBA financial returns come as a mixture of two normal distributions:

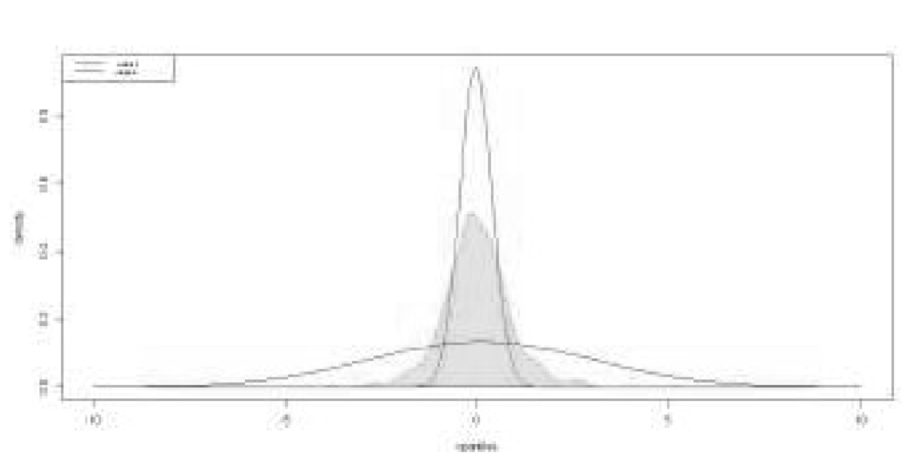


Fig. 6: Fit of HSBA weekly returns: the two components of the mixture

In order to represent correctly the join density of the mixture distribution we have to build the mixture distribution according to $\tau$ parameter as estimated by EM Algorithm. In this case the estimated $\tau$ value was 0.761.

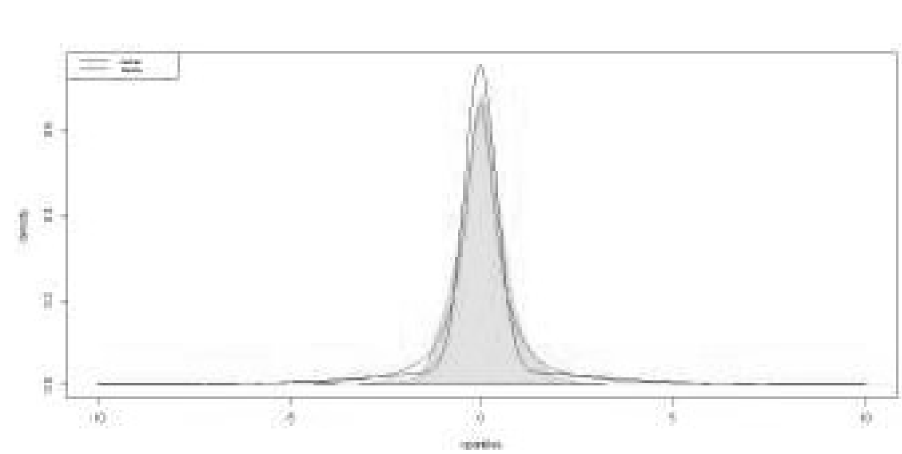


Fig. 7: Fit of HSBA weekly returns with a mixture of Normal

With the two normal mixture model we obtained a BIC criterion of 2871.93, what makes this fit slightly better than the fit provided by Cauchy, Normal and $t$-Student, according to BIC criterion. The results on the tail still not being the as good as desirable,

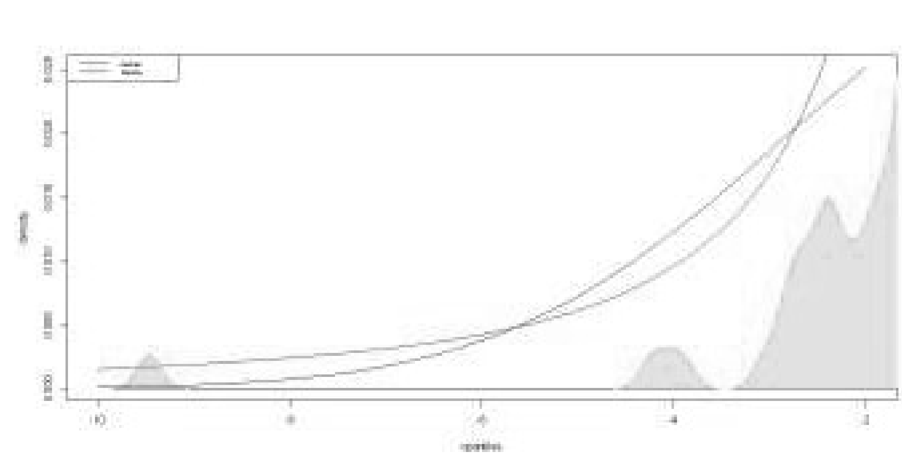


Fig. 8: Fit of HSBA weekly returns with a mixture of Normal

### 4.4 Extreme Value Mixture Models

The idea behind the use of Mixture of Extreme Value distribution is to combine the flexibility of using a distribution to capture the main component (the bulk distribution), that could be for example a Normal, and also the tails, as extreme values. With this mixture model one will get an entire distribution function by splitting the distribution in a bulk component and a tail components.

There are several approaches that consider only one tail and also approaches that consider both upper and lower tail. In this case the mixture function will be compounded potentially by a mixture of distribution from distinct families.

In our case we are specially interested in explore a mixture of a Normal distribution as bulk distribution with two Gamma tail distribution in both upper and lower tail. MacDonald et al. (2011) [12] proposed a two tailed mixture model where the standard kernel density estimator is spliced with two extreme value tail models.

This model uses a kernel density estimators to estimate the non-extreme value distribution and Generalized Pareto distribution (GPD) to estimate the tail distribution. A boundary-corrected kernel density estimator is also used in the case of a population with bounded support. This kernel density estimator assumes a particular kernel, in this case the normal density, which is centered at each data point, and uses only one parameter to define bandwidth. The model uses also the standard cross-validation likelihood to define bandwidth, combined with the likelihood for the peaks over threshold tail model, to give a full likelihood for all of the observations. The term tail fraction refers to the proportion of the distribution above the threshold. This parameter will be identified by $\Phi_u$and u represents the threshold.

The distribution function comes as:

$$F(x|\Theta) = \begin{cases} \phi_{ul} 1 - G(-x| - u_l, \sigma_{ul}, \epsilon_l) & x < u_l, \\ H(x|\mu, \sigma) & u_l \leq x \leq u_r \\ (1 - \phi_{ur}) + \phi_{ul} G(x|u_r, \sigma_{ur}, \epsilon_r) & x > u_r \end{cases}$$

where $\phi_{ul} = H(u_l|\mu, \sigma)$ and $\phi_{ur} = 1 - H(u_r|\mu, \sigma)$ and $H(.|\mu, \sigma)$ is the normal distribution with mean $\mu$ and standard deviation $\sigma$. $G(.| - u_l, \sigma_{ul}, \epsilon_l)$ and $G(.| - u_r, \sigma_{ur}, \epsilon_r)$ are GPD distributions for lower and upper tails respectively.

By applying the methodology to HSBA returns series we obtained the following estimates for the parameters:

The graph bellow shows the results for a mixture of a normal $N$(-0.00084,0.023) bounded at left by parameter $u_l$ =-0.0354 and on right by parameter $u_r = 0.025$.

The Gamma parameters obtained are respectively:

| **left tail** | **right tail** |
|---|---|
| $\phi_{ul} = 0.120$ | $\phi_{ur} = 0.119$ |
| $\mu_l = 0.133$ | $\mu_r = 0.0743$ |
| $\sigma_l = 0.593$ | $\sigma_r = 0.728$ |

The goodness of fit for this model is also slight better than previous ones with an BIC criterion value estimated as 2860.661. The advantage and flexibility of this mixture model is essentially in the tails of the distribution as it is able to take advantage of the capabilities of Gamma distribution to adapt to the tail. The graphs obtained for the extreme value mixture are as follows:

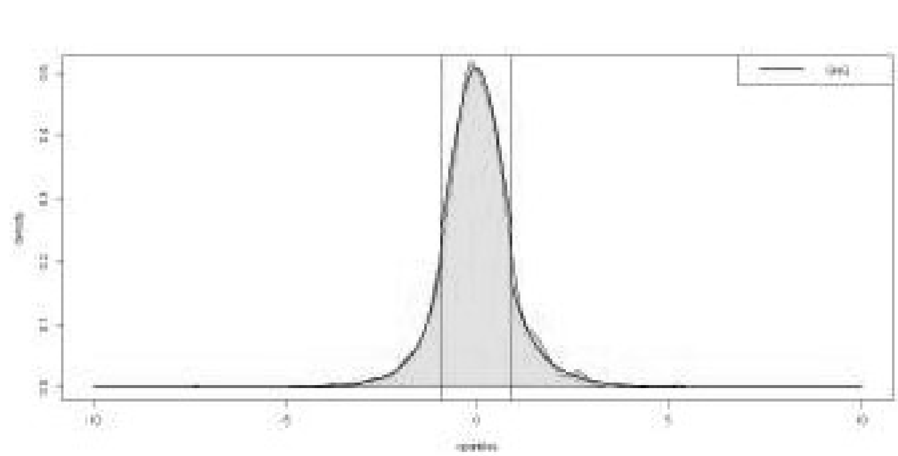


Fig. 9: Fit of HSBA weekly returns with a mixture of Gamma-Normal-Gamma

By visual analyse of the graph obtained with the Gamma-Normal-Gamma (GNG) fitting it is possible to identify a very close adjustment.

Also in the tail of the distribution we can notice a good approximation including a decay of the distribution function when it goes further on the left.

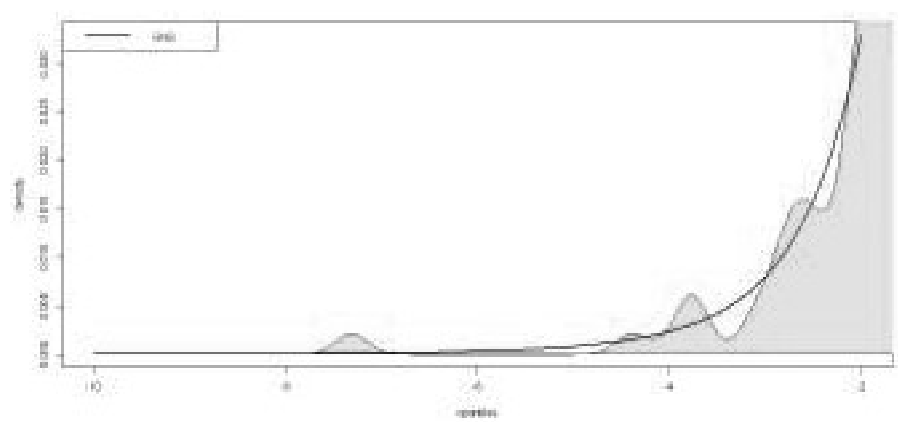


Fig. 10: Fit of HSBA weekly returns with a mixture of GNG on the left tail

## 5 Value at Risk Estimation

Based on the results we will compare the impact of the different models tested in terms of the estimates obtained for *VaR*.

Based on the quantiles for each model, Cauchy, Normal, *t*-Student and extreme value mixture GNG we estimate the *VaR* for several levels using several $\alpha$ values in a range between 0.001 and 0.1 as showed in the table bellow.

| | 0.1% | 1% | 2.5% | 5% | 10% |
|---|---|---|---|---|---|
| Cauchy | -5.7016 | 0.5702 | -0.2278 | -0.1133 | -0.0554 |
| Normal | -0.1139 | -0.0856 | -0.0720 | -0.0603 | -0.0468 |
| *t*- Student | -0.1541 | -0.1023 | -0.0823 | -0.0667 | -0.0503 |
| GNG | -0.1820 | -0.0992 | -0.0728 | -0.0547 | -0.0383 |
| Historical *VaR* | -0.1610 | -0.0975 | -0.0717 | -0.0548 | -0.0384 |

Table 3: *VaR* estimates

The values showed in the table above represent the percentage of the value of each institution at risk for each one of the risk levels (0.1% to 10%).

In this example is also visible the issue by using Cauchy as model in extreme values as the *VaR* estimate is to high and value is not reasonable. In general Cauchy based estimates are to excessive when compared with historical *VaR* value. In this case we can conclude Cauchy consistently overestimate the risk and is not also a good option for risk management application.

## 6 Conclusion

The results obtained by comparing the goodness of fit obtained by applying distinct statistics modelling techniques, highlighted a concern regarding the quality of the global adjustment versus the quality of the adjustment on the tails of the distribution. In certain applications the analyses of the tail of the distributions is of major importance, as for example in risk analyses. The results obtained for *VaR* estimates for each model implemented also showed that more complex model could be advantageous as they are more flexible in adapt to the tail of the distribution providing better adjustments.

Complex phenomenons requires also more complex models and the complexity of certain phenomenons like behavior of financial returns requires more complex and versatile models.

## 7 Acknowledgements

This research is based on the partial results of the Funded by FCT - Fundacão para a Ciência e a Tecnologia, Portugal, through the project UID/MAT/00006/2019.

## References


1. Adrian, Tobias and Brunnermeier, Markus K, *CoVaR*, National Bureau of Economic Research, 2011
2. Bingham, Nicholas H and Kiesel, R diger and Schmidt, Rafael and others, *A semi-parametric approach to risk management* Quantitative Finance, 3, 6 , 426-441, Taylor & Francis, 2003
3. Dempster, Arthur P and Laird, Nan M and Rubin, Donald B, *Maximum likelihood from incomplete data via the EM algorithm*, Journal of the Royal Statistical Society: Series B (Methodological) 39, 1, 1-22, Wiley Online Library 1977,
4. Fabozzi, Frank J and Gupta, Francis and Markowitz, Harry M, *The legacy of modern portfolio theory.* The Journal of Investing, 11,3, 7-22, 2002
5. Guerrero-Cusumano, José-Luis, *An asymptotic test of independence for multivariate t and Cauchy random variables with applications.* Information sciences, 92,1-4, 33-45,Elsevier, 1996
6. Hamilton, James D, *A quasi-Bayesian approach to estimating parameters for mixtures of normal distributions.* Journal of Business & Economic Statistics, 9, 1, 27–39, Taylor & Francis, 1991
7. Harris, David E, *The distribution of returns.* Available at SSRN 2828744, 2017
8. Jondeau, Eric and Poon, Ser-Huang and Rockinger, Michael, *Financial modelling under non-Gaussian distributions.* Springer Science & Business Media, 2007
9. Kaplanski, Guy and Kroll, Yoram , *VaR risk measures versus traditional risk measures: an analysis and survey.* (Journal of Risk, 2002) volume 4, number 3, pp. 1-27
10. Kimball, Ralph C., et al. *Failures in risk management.* New England Economic Review, 2000, 3-12.
11. Lakshmi, R. Vani, and V. S. Vaidyanathan. *Parameter estimation in gamma mixture model using normal-based approximation.* Journal of Statistical Theory and Applications 15, 1, 25-35, Atlantis Press, 2016
12. MacDonald, A and Scarrott, Carl John and Lee, D and Darlow, B and Reale, Marco and Russell, G . *A flexible extreme value mixture model.* Computational Statistics & Data Analysis, 55, 6, 2137-2157, Elsevier, 2011
13. Padoan, Simone A *Multivariate extreme models based on underlying skew-t and skew-normal distributions.* Journal of Multivariate Analysis, 102,5,977-991, Elsevier, 2011
14. Razzaghi, Mehdi and Kodell, Ralph L *Risk assessment for quantitative responses using a mixture model.* Biometrics 56.2 Wiley Online Library 2000 519-527
15. Reynolds, Douglas A *Gaussian Mixture Models* Encyclopedia of biometrics 741, Berlin, Springer, 2009
16. Terzić, Ivica and others *Empirical estimation and comparison of normal and student t linear var on the belgrade stock exchange*, Sinteza 2014-Impact of the Internet on Business Activities in Serbia and Worldwide 298-302, Singidunum University ,2014
17. Theodossiou, Panayiotis *Financial data and the skewed generalized t distribution*, Management Science, 44, 12-part-1, 1650-1661, INFORMS, 1998
18. Tseng, Jie-Jun and Li, Sai-Ping *Asset returns and volatility clustering in financial time series*, Physica A: Statistical Mechanics and its Applications, 390, 7, 1300-1314 Elsevier, 2011