# Modeling Skewness and Kurtosis Using Normal Mixture

Rafael Minghini (BW Gestão de Investimentos)

Paulo Beltrão Fraletti (FGV-EASP e Insper)

Pedro Nielsen Rotta (University of South Carolina)

**January 31st, 2017**

***Abstract***

*The aim of this paper is to provide a practical approach of a statistical tool proposed by Jin Wang (2001), defined as a mixture of "n" normal distributions, the Normal Mixture, capable of dealing with the non-normality issues of modeling financial series as fat-tails**. In order to illustrate the potential of this technique, an application to real data is proposed for the daily rates of returns of four financial series (IBOVESPA, VALE5, USDBRL, PRE-DI) from January 2006 to December 2016. To fit the parameters of the normal mixture it is employed the Maximum Likelihood Estimation (MLE) through the Expectation Maximization (EM) algorithm according to Hastie et al (2001) and Söderlind (2010). A step-by-step description of the commands in R software used to perform the estimation, goodness-of-fit test and simulation are presented in the last section of the paper.*

***Keywords:*** *financial series modeling, normal mixture*, *Maximum Likelihood (ML),* *Expectation Maximization (EM) algorithm*

# 1. Introduction

The assumption of normality for the distribution of financial market returns is a key part of the financial modeling research. The most common premise for decades has been that financial asset data follows a gaussian stationary process. An example of the wide use of this assumption is the fact that the Parametric VaR developed in JP Morgan (1995), one of the most widespread techniques in financial risk management, is based on the hypothesis that the probability distribution of the returns of individual assets follows a normal distribution.

In many situations, this approximation is a justified approach from a practical business standpoint, as it is easy to understand, and produces satisfactory results. However, several studies such as Duffie na Pan (1997), Venkataraman (1997) and Hull and White (1998) has shown that financial data exhibits strong skewness and kurtosis (fat tails). That is, they are remarkably distinct from a normal bell shape curve.

Considering the need that market players currently have to implement a easy to use framework capable of addressing these issues, the aim of this paper is to provide evidence on the benefits of employing a flexible statistical tool, proposed by Jin Wang (2001), defined as a mixture of *n* normal distributions, the Normal Mixture, able to deal with the violation of the normality assumption in empirical finance.

In order to illustrate the potential of this technique, an application to real data is proposed for the daily rates of returns of four financial series (IBOVESPA, VALE5, USDBRL, PRE-DI) from January 2006 to December 2016. To fit the parameters of the normal mixture it is employed the Maximum Likelihood Estimation (MLE) through the Expectation Maximization (EM) algorithm according to Hastie et al (2001) and Söderlind (2010).

In summary, section 2 of this paper will describe the statistical properties of the normal mixture, followed by a brief discussion (section 3) on the stylized facts of financial series that can be accommodated by the methodology discussed. In section 4, it is presented an application of the method using real-world financial data to evaluate the goodness-of-and simulation power. A description of the computational procedure is presented in section 5 as a "step-by-step" guide in software R.

# 2. Normal Mixture

## 2.1 Simple Normal Distribution

The simple normal distribution, also known as the Gaussian or Gaussian distribution, is one of the most important distributions in statistics. Its history is linked to the discovery of probabilities in mathematics, in the 17th century, initially to solve questions related to games of chance. The person most responsible for its development was Abraham de Moivre, a French mathematician exiled to England, who defined it in 1730, continuing the work of Jacob Bernoulli (Theorem of Large Numbers) and his nephew Nicolaus Bernoulli. In 1809, when Gauss rigorously demonstrated its properties in studies related to astronomy, his name was directly associated with distribution.

This distribution is entirely described by the mean and standard deviation parameters, that is, by knowing these, it is possible to determine any probability in a Normal.

The main characteristics of the simple normal distribution are:

* distribution mean, named µ
* distribution standard deviation, named σ
* distribution mode, which ocorre em x=µ
* symmetric around their mean

A continuous random variable  have Normal distribution if its probability density function is given by:

In a symmetrical distribution like normal, the way the values are distributed to the left of the mean is the same to the right. The positive and negative deviations have the same preponderance and the tails of the distribution have the same shape. In this case, the third centered moment, the symmetry, is null, .

Kurtosis is related to the degree of flattening of the distribution, often established in relation to the normal distribution, and allows measuring how “fat” the tails of a distribution are. In the case of normal we have,

It is possible to characterize a distribution, in terms of kurtosis, as:

**Mesocurtic** – kurtosis equal to the normal distribution one.

**Platykurtic** – kurtosis lower than the normal distribution one.

**Leptokurtic** - kurtosis greater than the normal distribution one.

In general, the normalized symmetry and kurtosis of a random variable X are defined by Casella and Berger (1990):

where and .

## 2.2 Non-Normality of Market Assets

In 2008, 10 years after the financial crisis in Russia-Asia that led the Long-Term Capital Management (LTCM) to bankruptcy, criticism of the use of risk assessment models based on the bell-shaped probability distribution increased considerably, stimulating the research for more sophisticated methods, and showing that these events continued to surprise financial institutions negatively.

As previously mentioned, the return distributions of several market assets are often asymmetric and are substantially leptokurtic (“fat tails”). That is, the way the values are distributed to the left of the mean is not the same to the right and extreme outcomes occur much more frequently than would be expected by a modeling based on the normal distribution. We see then that assuming normality, in many cases, is far from satisfactory and appropriate.



*Figura 1.1: Função densidade de probabilidade do índice IRF-M[[1]](#footnote-1) e da normal de mesma média e variância. Foram utilizados retornos mensais no período de Set/2003 até Out/2015.*

Next, we will discuss the properties of the Normal Mixture to model asymmetric distributions and distributions with a kurtosis measure different from the normal distribution one (platykurtic and leptokurtic).

**2.3 Definition and Properties**

We can describe the cumulative distribution function of a mixture of *n* individual normal distributions by the random variable X as follows (Jin Wang [2001]):

, (2.1)

Where φ is the cumulative distribution function of *N* (0,1) and is the weight assigned to each individual normal. In this way its probability density function is represented by:

,

for j= 1,...,n,

## Properties: If *X* is a mixture of *k* individual normal distributions with probability density function (2.1), its mean, variance, asymmetry and kurtosis are:

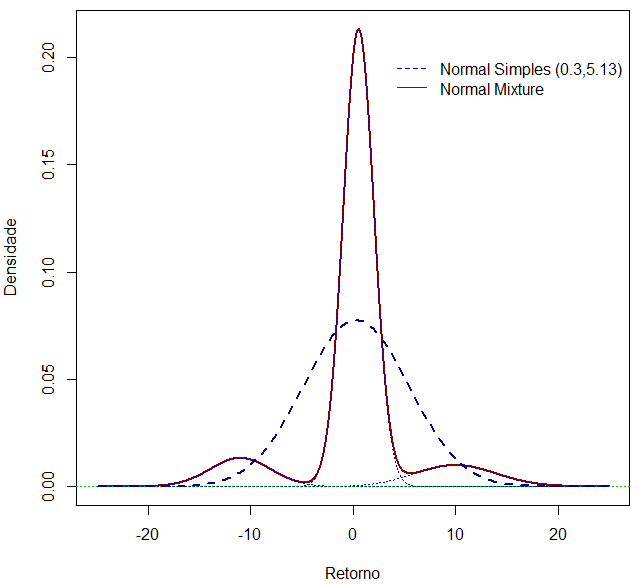
# 3. Accomodation of Stylized Facts

## a. Example “*Fat Tail*”

Consider the following three distributions:

* Normal 1 (0,5%, 1,5%) with probability of 80%;
* Normal 2 (10% , 4%) with probability of 10%;
* Normal 3 (-11%, 3%) with probability of 10%.

Applying formulas from item 2.a, the moments of the final distribution are:



The green line represents the equivalent normal distribution, with the same mean and standard deviation. By the numbers generated above with , we were able to obtain a distribution in which the way the values are distributed to the left of the mean is not the same to the right, that is, it is asymmetric. In this case, we have negative asymmetry, where the mode is less than the median, which, in turn, is less than the average.

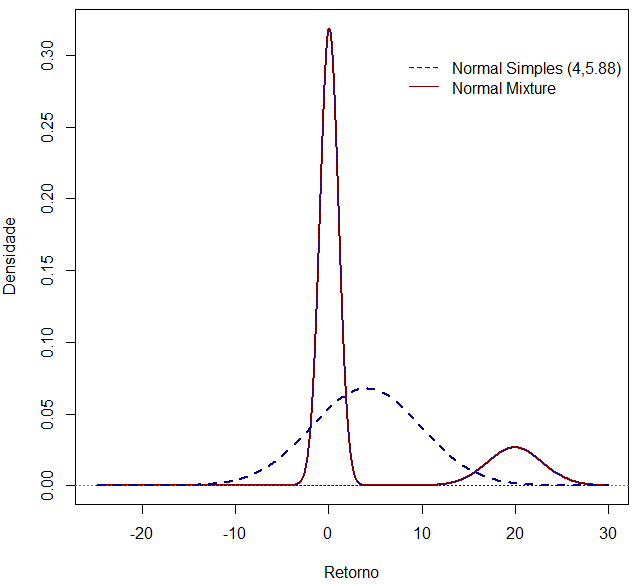
We can also verify that, with , the simulated distribution has tails that are "fatter" than the normal. That is, extreme outcomes occur with a higher frequency than would be expected by modeling based only on the normal distribution, where .

## Example of Discontinuous Marekts (*JUMP*)

Consider the following two distributions:

* Normal 1 (0%, 1%) with probability of 80%;
* Normal 2 (20% , 3%) with probability of 20%;

Applying formulas from item 3.a, the moments of the final distribution are:



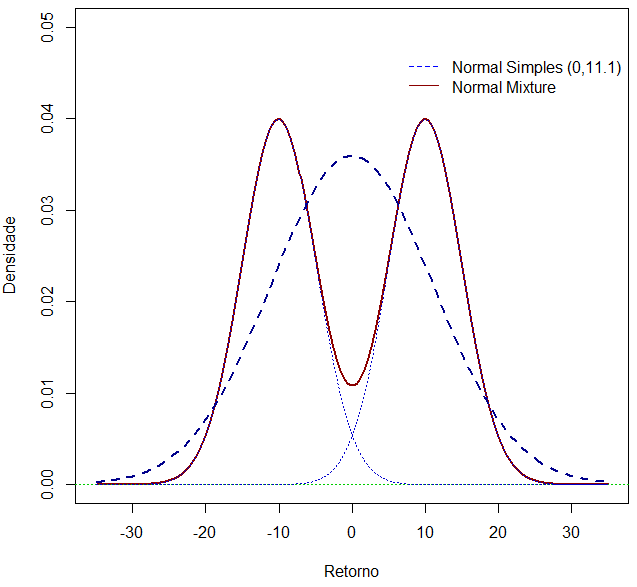
The green line represents the equivalent normal distribution, with same mean and standard deviation. By mixing these two normal distributions we can simulate non-continuous markets (jump), where asymmetry and kurtosis are evident features. In this case, the modeling is based, for example, on an expectation of asymmetric discontinuity of the quotation pattern of a given currency, originated by an eventual change in the exchange rate regime where a maximum devaluation is expected with some probability.

## Bimodal Market Example

Consider the following two distributions:

* Normal 1 (10%, 5%) with probability of 50%;
* Normal 2 (-10% , 5%) with probability of 50%;

Applying formulas from item 2.a, the moments of the final distribution are:



A linha verde representa a distribuição normal equivalente, de mesma média e desvio-padrão. A distribuição final acima representa a resultante de dois cenários completamente opostos com probabilidades equivalentes. Tal formato pode ser representado, por exemplo, pela expectativa do mercado em relação a uma notícia onde os impactos serão absorvidos de forma extremamente antagônica. Nesse sentido, adicionando esse parâmetro de incerteza em sua distribuição, se torna possível simular o impacto financeiro em sua carteira baseado neste cenário “binário” esperado.

The green line represents the equivalent normal distribution, with same mean and standard deviation. The final distribution above represents the combination of two completely opposite scenarios with equivalent probabilities. Such shape can be represented, for example, by the market's expectation in relation to a news story where the impacts will be absorbed in an extremely antagonistic way. By adding this parameter of uncertainty in its distribution, it becomes possible to simulate the financial impact on its portfolio based on this expected “binary” scenario.

# 4. Application with Real Data

In order to illustrate the potential that the technique of combining individual normal distributions has regarding the accommodation of some of the stylized facts and characteristics of financial series, in this section an application of this method is proposed for the series of daily returns of four assets (IBOVESPA, VALE5, USDBRL, PRE-DI) covering the period from January 2007 to December 2016. To calculate the parameters that define the normal mixture distribution, the Maximum Likelihood (MV) estimate is used through the Expectation Maximization (EM) algorithm according to Hastie et al (2001) and Söderlind (2010).

# 4.1 Estimation

The EM algorithm is a tool widely used to simplify the Maximum Likelihood estimation when the calculation is excessively complex. According to the literature, this method has an excellent performance in the face of problems involving unobserved variables, which is the case of the present study, since the probability of each observation belonging to a certain distribution is not observed. In the last decade significant advances have been introduced in relation to the estimation of models of mixture of distributions, especially through the Maximum Likelihood method by the algorithm as argued in Picard (2007). For the case of mixing two normal distributions (*k = 2*):

Where s the probability density function of a normal distribution with mean and standard-deviation . And the function of the likelihood log is given by:

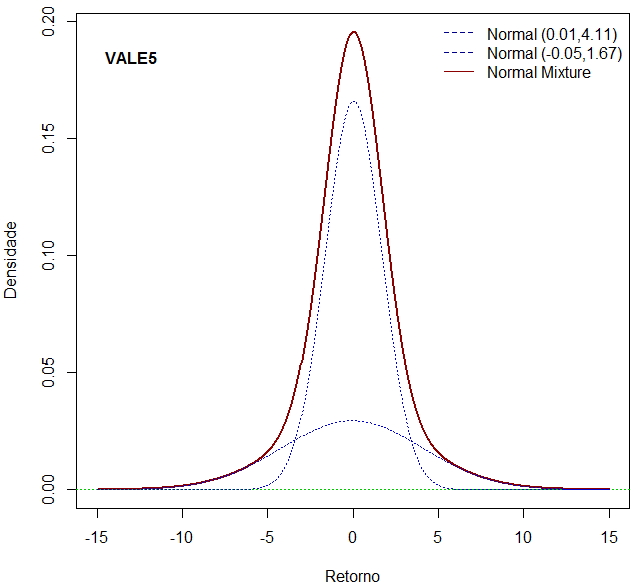
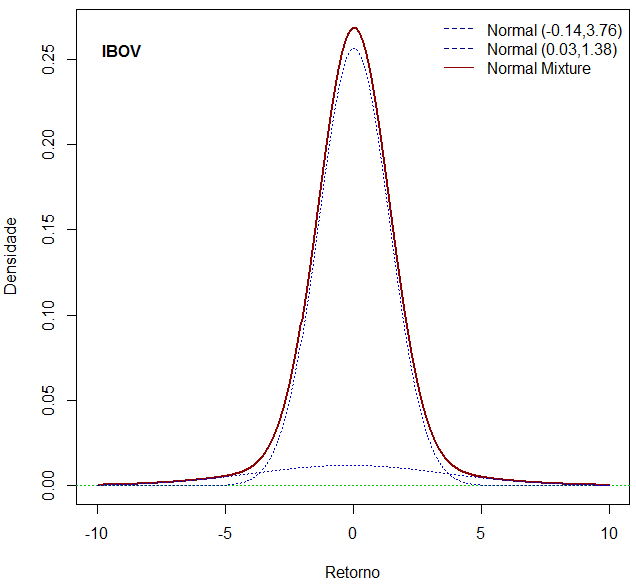
Given that the direct maximization of this likelihood log function is a complex procedure, it is within this context that the EM algorithm presents itself as an efficient alternative to solve this problem through numerical optimization. In summary, this method can be divided into two stages: in the first stage (E step) the EM algorithm determines both the expected value and the initial estimates of the parameters. In the second step (M step), the expected value is maximized. By repeating steps 1 and 2, the method converges to a local maximum of the likelihood function. The "nor1mix" package in the R software has the norMixEM function that performs the Maximum Likelihood (MV) estimate using the EM algorithm. Therefore, the procedure is easy to apply, and the results will be presented in the next subsection.

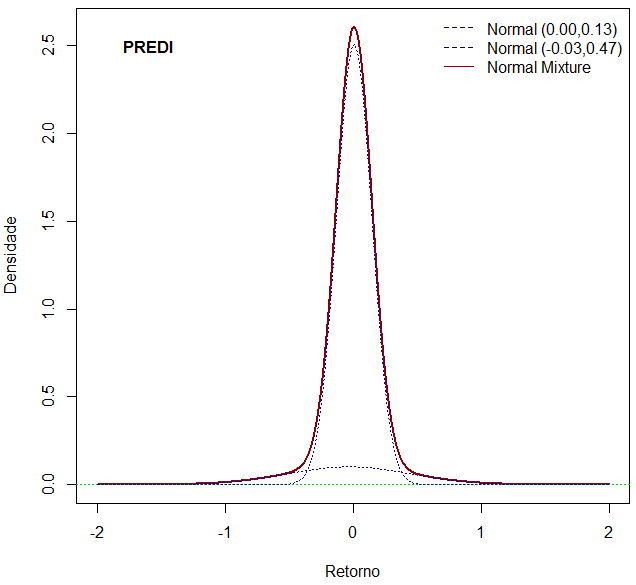
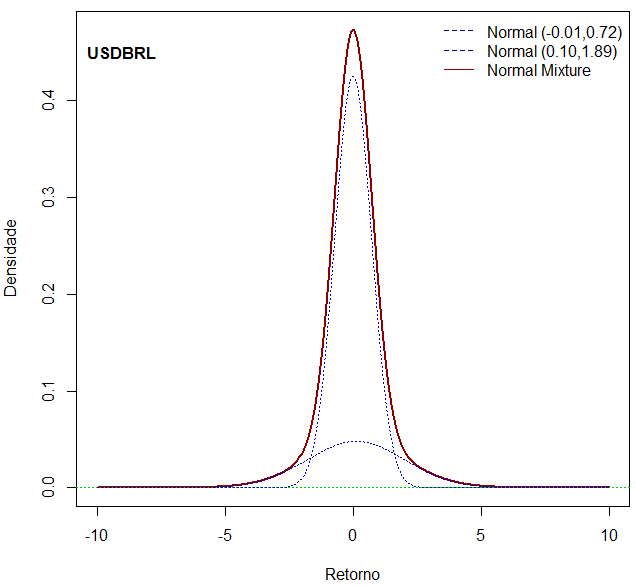
# 4.2 Results

**Tabela 1:** Estimated Parameters (%) by EM algorithm for n=2

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
| IBOV | 11.27 | -0.14 | 3.76 | 88.73 | 0.03 | 1.38 |
| VALE5 | 30.45 | 0.01 | 4.11 | 69.55 | -0.05 | 1.67 |
| USDBRL | 77.24 | -0.01 | 0.72 | 22.76 | 0.10 | 1.89 |
| PREDI | 87.79 | 0.00 | 0.13 | 12.20 | -0.03 | 0.47 |

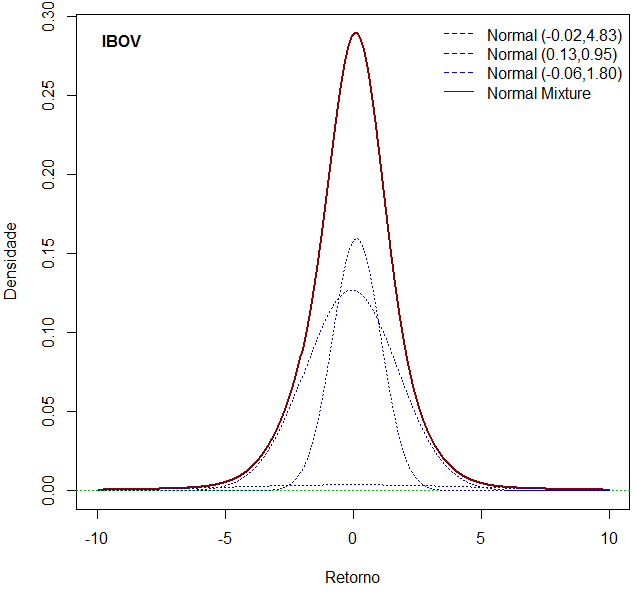
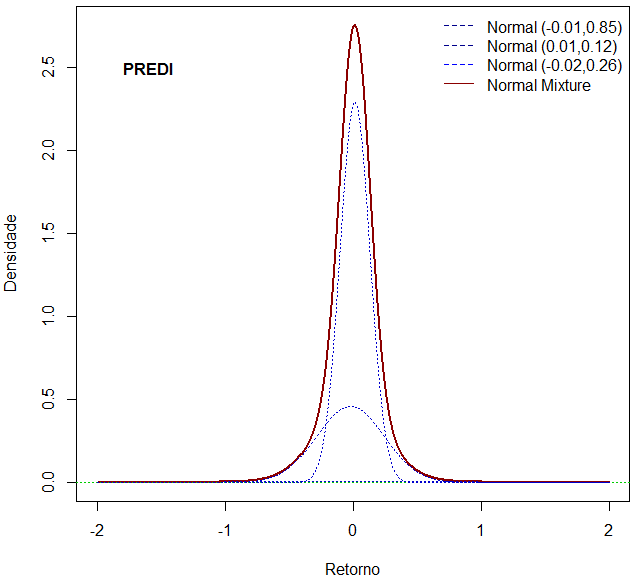
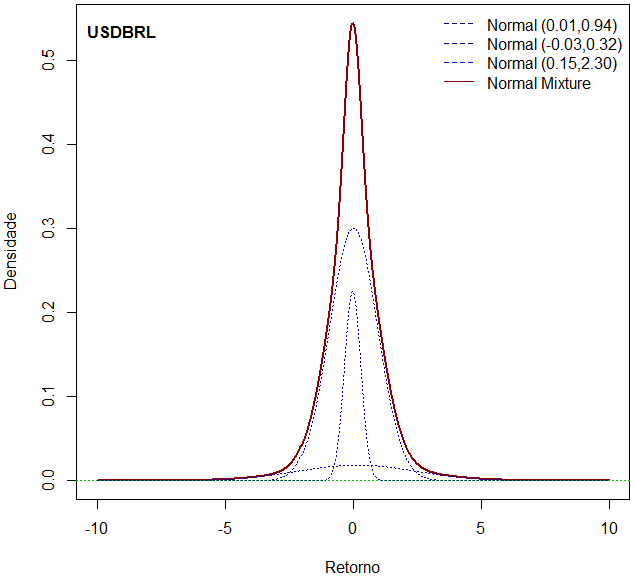
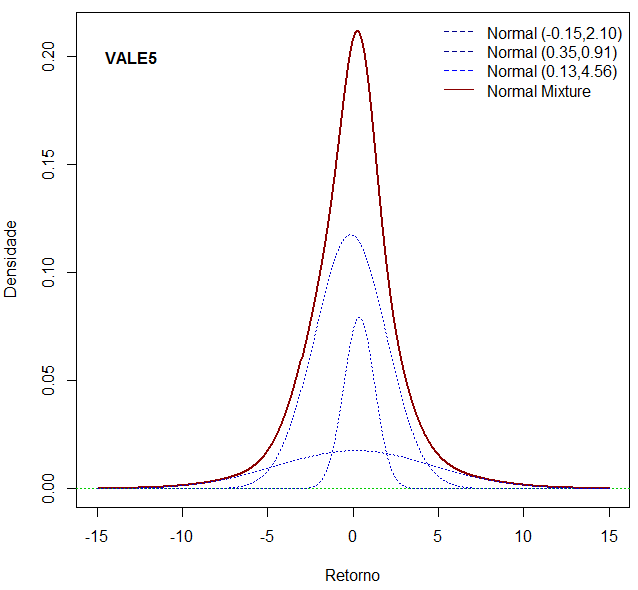
**Gráfico 1:** Normal Mixture n=2 and associated components

****

****

**Tabela 2:** Estimated Parameters (%) by EM algorithm for n=3

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |
| IBOV | 4.55 | -0.02 | 4.83 | 37.97 | 0.13 | 0.95 | 57.48 | -0.06 | 1.80 |
| VALE5 | 61.79 | -0.15 | 2.10 | 18.02 | 0.35 | 0.91 | 20.19 | 0.13 | 4.56 |
| USDBRL | 71.07 | 0.01 | 0.94 | 18.18 | -0.03 | 0.32 | 10.75 | 0.15 | 2.30 |
| PREDI | 1.95 | -0.01 | 0.85 | 67.55 | 0.01 | 0.12 | 30.50 | -0.02 | 0.26 |

**Grafico 2:** Normal Mixture n=3 and associated components ****

# 4.3 Goodness-of-fit Test

A measure commonly used to verify the quality of the fit of the model (goodness-of-fit) is the Chi-Square statistic, which has the following decision rule:

In order to verify the real gain from the use of the normal mixture distribution over the use of a simple normal distribution, a comparative table was constructed with the Chi-Square statistics for each asset and for each distribution used.

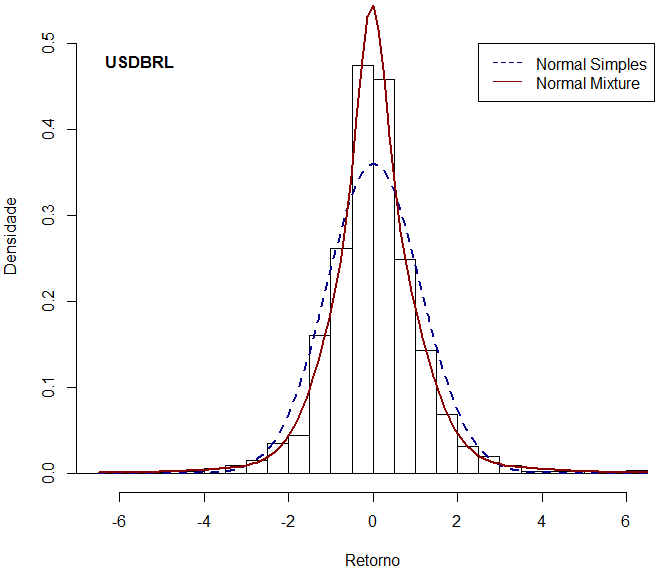
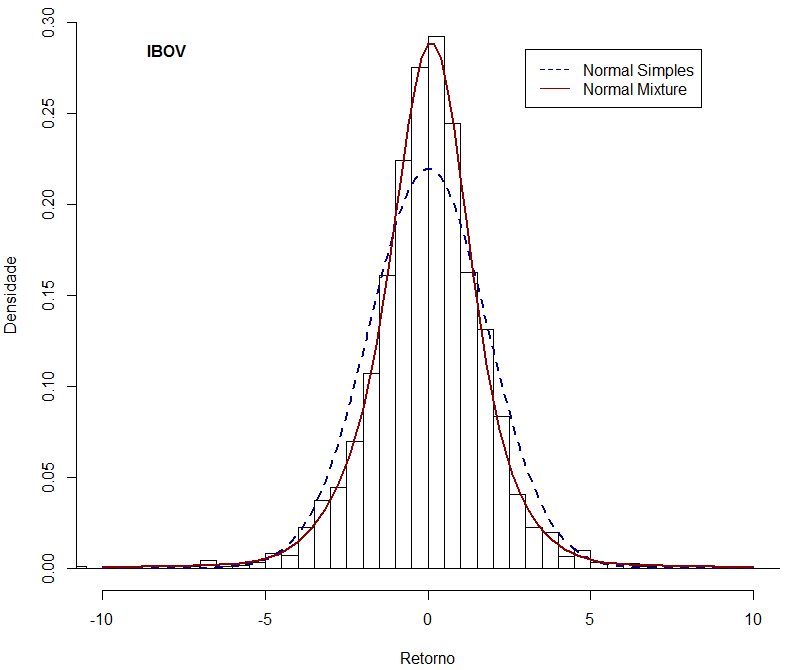
**Tabela 3:** Goodness of fit Test

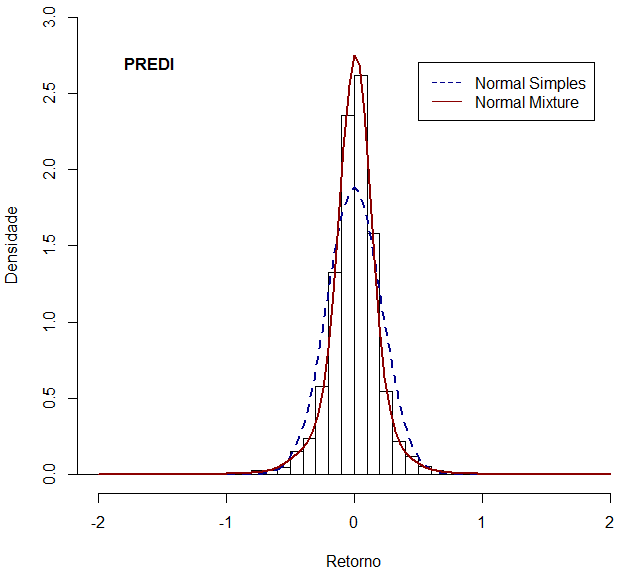
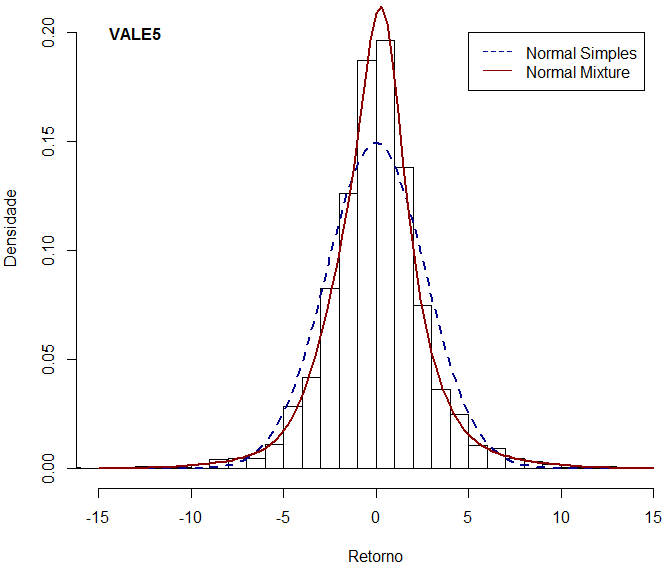
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Normal n=1 | | Normal Mixture n=2 | | Normal Mixture n=3 | | |
|  | CS | p-value | CS | p-value | | CS | p-value | |
| IBOV | 87987 | 0.0005\* | 144.69 | 0.0025\* | | 51.983 | 0.4303 | |
| VALE5 | 609390 | 0.0004\* | 128.82 | 0.0054\* | | 72.285 | 0.1539 | |
| USDBRL | 54105 | 0.0004\* | 257.96 | 0.0010\* | | 97.113 | 0.0530 | |
| PREDI | 1813120 | 0.0000\* | 45736 | 0.0004\* | | 65.253 | 0.0630 | |

\*reject at 1% significance level

The results indicate that, for all assets, the data modeling done with the mixture of normal distributions has a higher quality adjustment than using a single normal distribution. With a 95% confidence level, the null hypothesis is rejected for all assets when the simple normal distribution is used. In the case of the normal mixture adjustment using two components (n = 2), a significant improvement is already observed, but the best adjustment is found when we use normal mixture with three components (n = 3), because in all cases the hypothesis null is not rejected at the significance level of 1%, and the lowest value of the Chi-square statistic is also obtained. Overall, the graphs below directly illustrate the quality of the adjustments.

**Gráfico 3:** Normal Mixture and Simple Normal





Another way to show the gain from using the mixture of normal distributions over the simple normal can be seen in Table 4 below, which shows the values of asymmetry and kurtosis for each asset. In this way, we were able to verify the improvement in the ability to capture, with greater accuracy, the characteristics of the distribution.

**Tabela 4:** Asymmetry and Kurtosis

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Empirical | | Simple Normal | | Normal Mixture n=3 | | |
|  | Asymmetry | Kurtosis | Asymmetry | Kurtosis | | Asymmetry | Kurtosis | |
| IBOV | 0.01 | 8.63 | 0.00 | 3.00 | | -0.04 | 8.71 | |
| VALE5 | -0.06 | 6.04 | 0.00 | 3.00 | | -0.10 | 5.94 | |
| USDBRL | 0.21 | 7.75 | 0.00 | 3.00 | | 0.19 | 7.26 | |
| PREDI | -1.09 | 18.11 | 0.00 | 3.00 | | -0.14 | 17.68 | |

# 4.4 Normal Mixture Simulation

Once the goodness-of-fit test confirms the suitability of using the normal mixture for the series of returns, it may be of interest to the market participant to perform simulations of random samples originated from the distributions defined in Table 2 and Table 3. For illustration, we can consider the values obtained from Normal Mixture for Stock 1 (VALE5).

* Normal Mixture n=2 – VALE5

Normal 1 (-0.05%, 1.67%) with probrability 69.55%;

Normal 2 (0.01% , 4.11%) with probrability 30.45%;

Generate random variable U that follows Uniform distribution (0, 1).

If U < 0.6955, then generate random variable X that follows Normal distribution (-0.35%, 7.30%).

If U ≥ 0.3045, then generate random variable X that follows Normal distribution (1.26% , 16.60%).

* Normal Mixture n=3 – VALE5

Normal 1 (-0.15%, 2.10%) with probrability 61.79%;

Normal 2 (0.35% , 0.91%) with probrability 18.02%;

Normal 3 (0.13% , 4.56%) with probrability 20.19%;

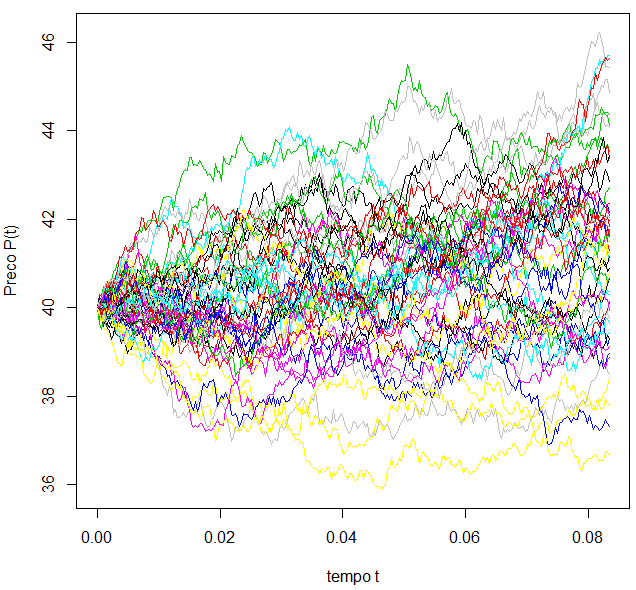
Generate random variable U with Uniform distribution (0, 1).

Se U < 0.3309, then generate random variable X that follows Normal distribution (0.84%, 16.40%).

Se 0.3309 ≤ U < 0.4881, generate random variable X that follows Normal (-3.38%, 4.93%).

Se U ≥ 0.4881, then generate random variable X that follows Normal distribution (8.38%, 3.46%).

**Grafico 3:** Multiple Price Trajectories Simulation – VALE5



Just as the *norMixEM* function of the "*nor1mix*" package of the R software was used in the process of estimating the parameters in subsection 4.1, we can use another function, called *rnorMix*, of this same package that is made available with the objective of carrying out simulations of samples extracted from a normal mixture. This calculation will be detailed in Step 5 of the “step-by-step” description in section 5.

# 5. Step-by-Step

In this section, the necessary procedures in the R software will be described, in a step-by-step format, for modeling financial data using: the normal mixture distribution with three components (“n” = 3), the respective goodness-of-fit test and the procedure of simulation. For easy of understanding, the steps below will be based on the procedures performed for Stock 1 (VALE5). The full code for all assets and different number of components is provided in the appendix.

**Step 1) Importing Data and Installating Package:** it is necessary to initially install the packages and respective libraries available for the R software that have the functions to be used in Step 2, Step 3 and Step 4. The "*nor1mix*" package has the *norMixEM* function that performs the Maximum Likelihood (MV) estimation using the EM algorithm, the "moments" package has the skewness and kurtosis functions, and the "xlsx" package features the read.xlsx function that is used to import data from an Excel file.

*> install.packages("nor1mix", lib="/data/Rpackages/")*

*> install.packages("moments", lib="/data/Rpackages/")*

*> install.packages("xlsx", lib="/data/Rpackages/")*

*> install.packages("zoo", lib="/data/Rpackages/")*

*> library(nor1mix);library(moments);library(xlsx);library(zoo)*

*> setwd("E:/Artigo\_02/Bases")*

*> data <- read.xlsx("valed1.xlsx", sheetIndex = 1)*

**Step 2) Estimation:** after importing the data in its original format which is the daily closing price series, the *retpct* series of returns is created by calculating the logarithm difference multiplied by 100. The norMixEM function allows the use of the Maximum Likelihood (MV) method through the Expectation Maximization algorithm (EM) for the estimation of the parameters of the normal mixture distribution. Finally, the norMix function stores the parameters estimated for use in the simulation (Step 5).

*> attach(data)*

*> retpct=diff(log(close))\*100*

*> parmixn3 <- norMix(mu=c(EMn3[1],EMn3[2],EMn3[3]),sigma=c(EMn3[4],EMn3[5],EMn3[6]),*

*+ w=c(EMn3[7],EMn3[8],EMn3[9]))*

**Step 3) Statistics and Graphs:** in this step statistics are generated as average, standard deviation, asymmetry and kurtosis as well as a histogram of the returns. This information will be used in Step 4 in the goodness-of-fit test.

*> mean=mean(retpct); sd=sd(retpct); skewness=skewness(retpct); kurtosis=kurtosis(retpct)*

*> hist <- hist(retpct, prob=TRUE,breaks=50,xlab="Retorno",ylab="Densidade",main=NULL, xlim=c(-15,15),ylim=c(0,0.22))*

*+ plot(parmixn3,p.comp=TRUE,xlim=c(-15,15),xlab="Retorno",ylab="Densidade",main=NULL)*

**Step 4) Goodness-of-fit Test**: the chisq.test function requires both the specification of the empirical density of the data (hist$counts), which was obtained by the hist command in Step 3, as well as the reference distribution specified in the null hypothesis, which for the case of mixing normals it is provided by the pnorMix function also available in the "nor1mix" package.

*> breaks\_cdf\_nor <- pnorm(hist$breaks,mean=mean(retpct),sd=sd(retpct))*

*> null.probs.nor <- rollapply(breaks\_cdf\_nor, 2, function(x) x[2]-x[1])*

*> chisq.nor <- chisq.test(hist$counts, p=null.probs.nor, rescale.p=TRUE, simulate.p.value=TRUE)*

*> breaks\_cdf\_mix <- pnorMix(hist$breaks,parmixn3)*

*> null.probs.mix <- rollapply(breaks\_cdf\_mix, 2, function(x) x[2]-x[1])*

*> chisq.mix.n3 <- chisq.test(hist$counts, p=null.probs.mix, rescale.p=TRUE, simulate.p.value=TRUE)*

**Step 5) Simulation:** the function called *rnorMix* of the "*nor1mix*" package is a tool made available with the objective of carrying out simulations of samples extracted from a normal mixture distribution. The function only requires as arguments the sample size to be generated and the vector containing the parameter values that, in the case of this article, was created after the estimation in Step 1.

*> x3 <- rnorMix(500,parmixn3)*

# 6. Conclusion

This article objective was to present in a practical way a statistical tool proposed by Jin Wang (2001), defined as a mixture of “n” normal distributions, Normal Mixture, in the field of empirical financial modeling. We show that, despite its practicality and wide use by market participants, the use of the simple normal curve as a way to replicate the behavior of financial market assets is far from adequate.

When applying the technique using the series of daily returns of four important assets in the Brazilian market (IBOVESPA, VALE5, USDBRL, PRE-DI) during a period from 2007 to 2016, it was possible to demonstrate the real gain in terms goodness-of-fit through the Chi-Square statistic comparison. The applications embraced modeling bimodal, asymmetrical and fat tails distributions so commonly found in financial series.

The use of the Normal mixture distribution can be of interest for a variety of goals the finance field, such as:

1. Risk Management: perform simulations to adequately verify possible impacts on the portfolio when modeling the expected tail events;
2. *Resampling***:** conduct data resampling using the available data subsets;
3. Pricing options with exotic probability distributions;
4. Represent the mixture of different future macroeconomic scenarios of members of a committee;
5. Portfolio optimization: when modeling all moments of the distribution, this technique has great benefit when incorporating asymmetry and kurtosis effects compared to the classic mean-variance model (Markowitz);

With this, we presented the great importance of the technique for different areas in the field of finance as well as a step-by-step guide to its applicability.

**References**

Casella, G., and Berger, R. L., 2002. Statistical inference. Duxbury Press.

Dempster, A.P., N.M. Laird and D.B. Rubin, 1977. Maximum likelihood from incomplete data via the EM algorithm. J. the Royal Statistical Society Series B, 39: 1-38.

Due, D., and J. Pan, 1997. “An overview of value at risk” Journal of Derivatives, 4(3), 7-49

J.P. Morgan, 1995. RiskMetricsTM. Technical Documentation, Release 1-3, J. P. Morgan, New York, NY.

Hastie, T., R. Tibshirani and J. Friedman, 2001. The Elements of Statistical Learning: Data Mining, Inference and Prediction. Springer Verlag.

Hull, J., and A. White, 1998. Value-at-risk when daily changes in market variables are not normally distributed," Journal of Derivatives, 5(3), 9-19.

Picard, F., 2007. An introduction to mixture models, Statistics for Systems Biology Group, Research Report No 7.

Söderlind, P., 2010. Lecture Notes in Empirical Finance (PhD): Return Distributions. University of St. Gallen.

Venkataraman, S., 1997. Value at risk for a mixture of normal distributions: the use of quasi Bayesian estimation techniques," Economic Perspective, Federal Reserve Bank of Chicago, March/April, 2-13.

Wang J., 2001. Generating daily changes in market variables using a multivariate mixture of normal distributions. In Proceedings of the 2001 Winter Simulation Conference, Peters BA, Smith JS, Medeiros DJ, Rohrer MW (eds). Arlington, VA; 283–289.

1. **O IRF-M expressa a variação de mercado dos títulos prefixados do governo – LTN e NTN-F – e está exposto ao risco de oscilações nas taxas de juros, em função de, principalmente, reversões de expectativas de juros reais e inflação futuros até o prazo do investimento.** [↑](#footnote-ref-1)