# Modelling Financial Returns with Finite Mixtures of GED

## *Modellazione dei Rendimenti Finanziari con Misture Finite di GED*

Pierdomenico Duttilo and Stefano Antonio Gattone

**Abstract** As widely identified by the empirical evidence, daily returns on financial assets are not Normally distributed, because they are characterized by excess kurtosis and different degrees of skewness. Finite mixtures of distributions have been proposed in literature to capture these features. In this work a finite mixture of two Generalized Error Distributions (GED) is applied to fit the distribution of the daily returns on the Dow Jones Industrial Average (DJIA) index for the period from January 4, 2016 to January 31, 2022. Moreover, in order to highlight the flexibility of the shape parameter over time, the entire analysis period was divided in three different sub-periods and for each one the mixture of GED was estimated.

**Abstract** *Come ampiamente identificato dall'evidenza empirica, i rendimenti giornalieri delle attività finanziarie non sono distribuiti normalmente, poiché sono caratterizzati da curtosi eccessiva e diversi gradi di asimmetria. Per catturare queste caratteristiche, in letteratura sono state proposte misture finite di distribuzioni. In questo lavoro viene applicata una mistura finita di due distribuzioni GED (Generalized Error Distribution) per stimare la distribuzione dei rendimenti giornalieri dell'indice Dow Jones Industrial Average (DJIA) per il periodo dal 4 gennaio 2016 al 31 gennaio 2022. Inoltre, al fine di evidenziare la flessibilità del parametro di forma nel tempo, l'intero periodo di analisi è stato suddiviso in tre diversi sottoperiodi e per ciascuno è stata stimata la mistura di GED.*

**Key words:** daily financial returns, excess kurtosis, skewness, finite mixtures of GED

---

Pierdomenico Duttilo
University "G. d'Annunzio" of Chieti-Pescara, Viale Pindaro 42, 65127 Pescara, Italy e-mail: pierdomenico.duttilo@unich.it

Stefano Antonio Gattone
DISFIPEQ, University "G. d'Annunzio" of Chieti-Pescara, Viale Pindaro 42, 65127 Pescara, Italy e-mail: gattone@unich.it

## 1 Introduction

Commonly, financial time series such as daily returns on stocks, indices, currencies, cryptocurrencies and many others financial assets do not follow a Gaussian distribution because they are characterized by excess kurtosis and different degrees of skewness. Finite mixtures of distributions have been proposed in literature to capture these features.
Finite mixture of Normal distributions known also as mixtures of Gaussians are widely used in this field, for instance Kon (1984) [5] proposed a discrete mixture of Normal distributions to approximate the excess kurtosis and positive or negative skewness of daily returns distribution of common stocks and indices. More recent studies [1, 2, 4] argued that finite mixtures of Gaussians (with two or three components) are a good tool to fit the empirical distribution of financial returns. However, these mixtures impose a priori specific constraints on the form of the returns distribution since the components are Gaussians. The finite mixture of GED, known also as the finite mixture of generalized normal distribution can overcome this critical issue thanks to the flexibility provided by the additional shape parameter $\nu_k$. In this framework, the recent contribution of Wen et al. [8] (2020) is remarkable. They studied a univariate mixture of GED and proposed an expectation conditional maximization (ECM) algorithm for parameter estimation. Additionally, using data sets of the S&P 500 and Shanghai Stock Exchange Composite Index (SSEC), it was found that the mixture of GED better describes the excess kurtosis and skewness of daily returns compared to mixtures of Gaussians.
This work aims to enrich the existing literature on the use of mixtures of GED in finance by applying a finite mixture of two generalized error distributions to fit the distribution of the daily returns on the Dow Jones Industrial Average (DJIA) index. Besides, the likelihood-ratio test (LR Test) and information criteria were applied to compare the goodness of fit performance among the mixture of two GED, the mixture of two Gaussian distributions and the mixture of a Gaussian and a Laplace distribution [3].
Moreover, the entire analysis period was divided in three different sub-periods and for each one the mixture of GED was estimated in order to highlight the flexibility of the shape parameter over time.
The rest of the paper was organized as follows. Section 2 illustrates the methodological framework. Section 3 illustrates the empirical application. Finally, Section 4 provides the results discussion and some conclusions.

# 2 Methodological Framework

## 2.1 *Generalized Error Distribution*

A random variable $X$ is said to have the generalized error distribution with parameters $\mu$ (location), $\sigma$ (scale) and $\nu$ (shape) if its *probability density function* (p.d.f.) is given by

$$f(x|\mu,\sigma,\nu) = \frac{\nu}{2\sigma\Gamma(1/\nu)} \exp\left\{-\left|\frac{x-\mu}{\sigma}\right|^{\nu}\right\}, \tag{1}$$

with $\Gamma(1/\nu) = \int_0^\infty t^{1/\nu-1} \exp^{-t} dt$, $-\infty < x < \infty$, $-\infty < \mu < \infty$, $\sigma > 0$, $\nu > 0$.
Thanks to the shape parameter, the GED distribution is a flexible tool to capture a large class of statistical distributions [7, 8], for example with $\nu = 1$ and $\nu = 2$ GED becomes a Normal and Laplace distribution, respectively.

## 2.2 *Finite Mixtures of GED*

A finite mixture of GED with K components is given by the marginal distribution of the random variable $X$

$$\begin{aligned} f(x|\mu_k,\sigma_k,\nu_k) &= \sum_{k=1}^{K} \pi_k p(x|\mu_k,\sigma_k,\nu_k), \\ & \sum_{k=1}^{K} \pi_k \frac{\nu_k}{2\sigma_k\Gamma(1/\nu_k)} \exp\left\{-\left|\frac{x-\mu_k}{\sigma_k}\right|^{\nu_k}\right\}, \end{aligned} \tag{2}$$

where $\nu_k > 0$, $\sigma_k > 0$, $\mu_k \in$ R, $0 < \pi_k < 1$ and $\sum_{k=1}^{K} \pi_k = 1$. With $K = 2$ the mixture of two GED is given by:

$$\begin{aligned} f(x|\theta) &= \sum_{k=1}^{2} \pi_k p(x|\mu_k,\sigma_k,\nu_k), \\ &= \frac{\pi_1\nu_1}{2\sigma_1\Gamma(1/\nu_1)} \exp\left\{-\left|\frac{x-\mu_1}{\sigma_1}\right|^{\nu_1}\right\} + \frac{\pi_2\nu_2}{2\sigma_2\Gamma(1/\nu_2)} \exp\left\{-\left|\frac{x-\mu_2}{\sigma_2}\right|^{\nu_2}\right\}, \end{aligned} \tag{3}$$

where $\theta = (\pi_1,\pi_2,\mu_1,\mu_2,\sigma_1,\sigma_2,\nu_1,\nu_2)$. The mixture of two GED can be estimated via the ECM algorithm [8]. As shown by Wen et al. (2020) [8], this model has important properties because it nests several distributions as its sub-models. Especially, depending on the value of the shape parameter ($\nu_k$), the mixture of two GED reduces to:

- the mixture of two Gaussians when $\nu_1 = \nu_2 = 2$;
- the mixture of two Laplace distributions when $\nu_1 = \nu_2 = 1$;
- the mixture of a Gaussian and a Laplace distribution when $\nu_1 = 2$ and $\nu_2 = 1$;
- the mixture of a Gaussian and a GED distribution when $\nu_1 = 2$ and $\nu_2 > 0$;
- the mixture of a Laplace and a GED distribution when $\nu_1 = 1$ and $\nu_2 > 0$.

As a result the mixture of GED does not impose a priori specific constraint on the shape of each component of the mixture [6].

## 3 Empirical Application

The daily closing prices of the DJIA index between January 4, 2016 to January 31, 2022 were collected for the analysis from https://finance.yahoo.com. Next, the daily return $r_t$ in period $t$ is defined as $r_t = (\ln P_t - \ln P_{t-1})100$ where $P_t$ and $P_{t-1}$ are the closing prices at time $t$ and $t-1$, respectively. Panel (a) of Figure 1 shows the daily returns on the DJIA index.
Estimation results of the mixtures of distributions with two components for the daily returns on the DJIA index are summarized in Table 1. According to the LR Test and information criteria the mixture of two GED (Figure 1, panel b) is preferred over the mixture of two Gaussians and the mixture of a Gaussian and a Laplace distribution. Furthermore, the estimated shape parameters show that the first component has heavier tails than a Laplace distribution ($0.53 < 1$), while tail weights of the second component intermediate between the Gaussian and the Laplace distribution ($1 < 1.12 < 2$). These results are in line with those found by Wen et al. (2020) [8] who estimated the two-component mixture of GED on the daily returns of the S&P500 identifying a mixture component that allows a more extreme tail behaviour compared to the other.

In order to highlight the flexibility of the shape parameter over time, the entire analysis period was divided in three different sub-periods and for each one the two-component mixture of GED was estimated. As showed in panel (a) of Figure 1 each sub-period reflects different volatility levels: low (2016-2017), intermediate (2018-2019) and high (2020-2022). In addition, the latter two sub-periods are characterized by a higher number of large negative returns, i.e. negative skewness. Finally, Table 2 suggests that the third sub-period has a two-component mixture of GED with heavy tails compared to the other sub-periods.

It is important to note that the current version of the mixture model proposed in this work can be cast in the framework of unconditional (with respect to time) estimation which suggests a poor predictive ability. Indeed, previous studies [1, 2, 4, 5, 8] do not apply mixtures of distributions to make predictions. Nonetheless, this work provides an out-of-sample application of the Value at Risk (VaR) estimation. The first two sub-periods are taken as "in-sample" observations and the third sub-period as the "out-of-sample" observations. The estimated $\widehat{\text{VaR}}_{\alpha=0.01}$ in the "in-sample" observations is -2.736, -2.482 and -2.364 for the mixture of GED, mixture of Gaussians, and mixture of a Gaussian-Laplace, respectively. The empirical $\text{VaR}_{\alpha=0.01}$ of the third sub-period (out-of-sample) is -5.702. The evidence suggests that the third sub-period is characterised by an extreme $\mathrm{VaR}_{\alpha=0.01}$ value (due to the COVID-19 crisis) and the estimated $\widehat{\mathrm{VaR}}_{\alpha=0.01}$ of the two-component mixture of GED is the closest compared to the other two-component mixtures.

## 4 Conclusion

It has been shown that the mixture of GED is a powerful and flexible tool to fit the empirical distribution of financial returns. Considering the results of the empirical application in Section 3 at least two interesting considerations arise. Firstly, the mixture of GED with two components can model the behaviour of daily returns more appropriately and steadily compared to benchmark models, i.e. the mixture of two Gaussians and the mixture of a Gaussian and a Laplace distribution. Secondly, the estimated shape parameters change over time, they are not constant. Thus, the shape parameter changes according to the behaviour of daily returns i.e. market conditions. Consequently, the overall volatility estimated by the mixture in each sub-period (0.63, 0.97, 1.73) reflects the corresponding volatility level showed by Figure 1 (low, intermediate and high).

## Appendix

Table 1: Estimation results of the two-components mixtures of distributions.

| Parameter | Mixture of distributions | | |
|---|---|---|---|
| | Gaussian | Gaussian-Laplace | GED |
| $\pi_1$ | 0.1416 | 0.5415 | 0.3221 |
| $\pi_2$ | 0.8583 | 0.4584 | 0.6778 |
| $\mu_1$ | -0.4242 | 0.1353 | -0.0721 |
| $\mu_2$ | 0.1245 | -0.0095 | 0.1553 |
| $\sigma_1$ | 3.8912 | 0.6652 | 0.2115 |
| $\sigma_2$ | 0.8876 | 1.0912 | 0.6317 |
| $v_1$ | 2 | 2 | 0.5304 |
| $v_2$ | 2 | 1 | 1.1245 |
| $Stdev_1$ | 2.7515 | 0.4703 | 1.7759 |
| $Stdev_2$ | 0.6276 | 1.5433 | 0.7468 |
| $Stdev$ | 1.2031 | 1.1031 | 1.1855 |
| LL | -2047.13 | -2017.19 | -1997.13 |
| LR Test | 100.00* | 40.12* | |
| AIC | 4099.26 | 4039.37 | 4001.25 |
| BIC | 4130.93 | 4071.04 | 4045.59 |
| HQIC | 4114.18 | 4054.30 | 4022.15 |
| EDC | 4133.39 | 4073.20 | 4049.03 |

*$p$-value = 0.

Table 2: Estimation results of the two-components mixtures of GED in the three sub-periods.

| Parameter | Sub-periods | | |
|---|---|---|---|
| | 2016-2017 | 2018-2019 | 2020-2022 |
| $\pi_1$ | 0.4890 | 0.7942 | 0.1180 |
| $\pi_2$ | 0.5110 | 0.2058 | 0.8820 |
| $\mu_1$ | 0.0481 | 0.1507 | -0.5622 |
| $\mu_2$ | 0.1220 | -0.4872 | 0.1198 |
| $\sigma_1$ | 0.3940 | 0.7469 | 3.4994 |
| $\sigma_2$ | 0.8773 | 2.2960 | 0.9137 |
| $\nu_1$ | 2.1787 | 1.5205 | 1.1075 |
| $\nu_2$ | 1.3384 | 1.9479 | 1.2465 |
| $Stdev_1$ | 0.2674 | 0.6344 | 4.2266 |
| $Stdev_2$ | 0.8402 | 1.6464 | 0.9469 |
| $Stdev$ | 0.6302 | 0.9716 | 1.7267 |
| LL | -421.34 | -647.61 | -857.05 |

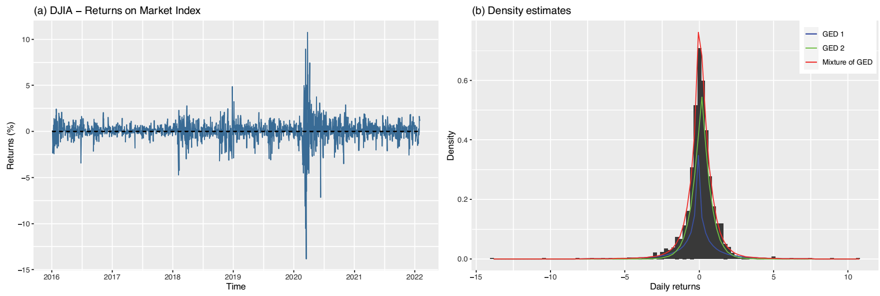


Fig. 1: Daily returns on the DJIA index and estimation of the density of the two-component mixture of GED.

## References


1. Behr, A., Pötter, U.: Alternatives to the normal model of stock returns: Gaussian mixture, generalised logF and generalised hyperbolic models. Ann. Finance **5**, 4-9–68 (2009)
2. Bellalah, M., Lavielle, M.: A Decomposition of Empirical Distributions with Applications to the Valuation of Derivative Assets. Multinational Finance Journal **6**, 99–130 (2002)
3. Haas, M., Mittnik, S., Paolella, M.: Modelling and predicting market risk with Laplace-Gaussian mixture distributions. Applied Financial Economics. **16**, 1145–1162 (2006) doi:10.1080/09603100500438817.
4. Han, L., Yan, H., Zheng, C.: Normal mixture method for stock daily returns over different sub-periods. Communications in Statistics - Simulation and Computation **48**, 447–457 (2019) doi:10.1080/03610918.2017.1383423
5. Kon, S.J.: Models of Stock Returns–A Comparison. The Journal of Finance **39**, 147–165 (1984)
6. Mohamed, O.M., Jaïdane, M.: Generalized Gaussian mixture model. 17th European Signal Processing Conference, 2273–2277 (2009)
7. Nadarajah S.: A generalized normal distribution. Journal of Applied Statistics **32**, 685–694 (2005) doi:10.1080/02664760500079464
8. Wen, L., Qiu, Y., Wang, M., Yin, J., Chen, P.: Numerical characteristics and parameter estimation of finite mixed generalized normal distribution. Communications in Statistics - Simulation and Computation (2020) doi:10.1080/03610918.2020.1720733