INSTITUTO
UNIVERSITÁRIO
DE LISBOA

UNIVERSIDADE
DE LISBOA

Patterns of skewness risk

Daniela Sofia dos Santos Alberto

MSc in Mathematical Finance

Dissertation supervised by:
Professor Joaquim Paulo Viegas Ferreira de Carvalho, Invited Assistant Professor,
ISCTE Business School

November 2021

Department of Finance
Department of Mathematics

Patterns of skewness risk

Daniela Sofia dos Santos Alberto

MSc in Mathematical Finance

Dissertation supervised by:
Professor Joaquim Paulo Viegas Ferreira de Carvalho, Invited Assistant Professor,
ISCTE Business School

November 2021

Acknowledgements

Nothing can be made by itself. And so the help and support of many people have been crucial. I start to acknowledge the ones with my gratitude. First, I would like to thank professor Paulo Viegas de Carvalho for allowing me to participate in this theme and research, and getting the knowledge and work this area deserves. I would particularly like to thank the master professors, especially João Pedro Nunes for arousing my interest in the applications of mathematics in finance and for all the support given during the course. To all the professors with an open door and, in this atypical year, an "open e-mail", my biggest thanks for answering my questions and for helping me. It would not be possible to list every helpful colleague as in working or emotional area. As this last year went by, the closest to us are the ones who have been at the top of the support. By saying, I would like to thank my family for being next to me and believing in me and in my work, mother, father, sister and Daniel. My closest friends, the ones who stayed with me for this long journey, also deserve my gratitude. I could not be less thankful to those who own their work into risk, researching and improving all the knowns and unknowns of mathematical finance. To the researchers who released their works so mine could have a positive impact on the area.

Resumo

É convenientemente assumida a distribuição Normal como representante do comportamento de retornos de ativos e de índices. No entanto, a Skewed Generalised t (SGT), tendo também uma forma de sino, consegue refletir melhor o mercado e as quedas e aumentos extremos. A SGT tem a capacidade de acompanhar a distribuição empírica quando mostra altos valores de enviasamento e de curtose, os quais são ignorados quando se utiliza uma distribuição Normal.

Esta tese contribuiu para a investigação sobre os padrões de enviesamento da distribuição real das rendibilidades de ativos financeiro, utilizando séries de índices de ações como objeto de análise. O trabalho realizado analisa os padrões de enviesamento das distribuições relativas a diferentes horizontes temporais, períodos e frequências. Para o efeito, consideramos a distribuição SGT, uma distribuição de cinco parâmetros inicialmente introduzida por Theodossiou (1998), para fazer o enquadramento da distribuição empírica dos retornos financeiros a uma distribuição parametrizada. Para tal, foram considerados índices dos principais mercados de ações no mundo: FTSE 100, Dax 30, S&P 500 e o Nikkei 225; com base no número de empresas cotadas.

Os resultados obtidos validam a hipótese de que os retornos financeiros apresentam efetivamente enviesamento para o lado esquerdo, confirmando que a SGT supera a Normal em termos da capacidade em seguir o comportamento da distribuição empírica. Estes resultados são encorajadores para encarar a SGT em termos paramétricos na estimação do Value at Risk (VaR) e futuras metodologias financeiras.

Abstract

The Skewed Generalised t (SGT), which has a bell shape, is more suitable to reflect market returns behaviour and its extreme drops and rises than the Normal distribution. This distribution is efficient in following the empirical distributions of financial returns, as it may incorporate high values of skewness and kurtosis, typically ignored by the normal distribution.

This dissertation contributes to the investigation on skewness present in financial returns, by using stock indices as the object of analysis. The study focuses on identifying patterns of skewness in distributions with different time windows, time periods and frequencies. For the analysis, we use the SGT distribution, a five-parameter distribution initially introduced by Theodossiou (1998), to reflect the empirical distribution of returns into a parameterized distribution. To this end, the leading indices of the most traded currencies in the world were considered: FTSE 100, DAX 30, S&P 500 and Nikkei 225.

The achieved results validate the hypothesis that the financial returns show a negative asymmetry, confirming that the SGT fits the empirical distribution better than the Normal distribution does. These outcomes are uplifting to use the SGT in parametric terms to estimate the Value at Risk (VaR) and future financial methodologies.

Keywords: Skewness; Value at Risk; Market risk; Skewed Generalized t distribution; Generalised Pareto

JEL Codes: C12; C13.

Contents

List of Figures x

List of Tables xiii

1. Introduction 1

2. Literature Review 3

2.1. Deviations from the assumptions of the normal distribution 3

2.2. Alternative distributions 5

2.2.1. The Skewed Generalised T Distribution 5

2.2.2. Extreme Value Theory 7

2.2.3. Generalised Pareto distribution 11

3. Concepts and Methodologies 15

3.1. Maximum Likelihood Estimation 15

3.1.1. SGT Likelihood function 16

3.1.2. GP Likelihood function 16

3.1.3. Normal Likelihood function 17

3.2. Goodness of fit 17

3.2.1. Kolmogorov-Smirnov test 18

3.2.2. Anderson-Darling test 19

3.2.3. Kuiper Test 19

4. Empirical Study 21

4.1. Data Description 21

4.2. Estimation results 26

4.3. Goodness of fit results 27

5. Conclusions 41

A. Moments 47

B. Demonstration: equality of BenSaïda's and Theodossiou's formula 51

C. Kolmogorov-Smirnov table 53

List of Figures

2.1. Nested distributions of the SGT distribution modified from Sikora et al. (2019) . 7
2.2. EVT distributions comparison from Embrechts et al. (2013) 10
2.3. Nested distributions of the GPD 13

4.1. Nikkei225 fitted distributions for (a) weekly returns (b) daily returns (c) p1 returns 33
4.2. SP500 fitted distributions for (a) weekly returns (b) daily returns (c) p1 returns 34
4.3. FTSE100 fitted distributions for (a) daily returns (b) weekly returns (c) p4 returns 35
4.4. FTSE100 GPD fitted distributions for (a) daily returns 20% (b) daily returns 5% (c) weekly returns 20% 36
4.5. Fitted distributions for (a) Portfolio P3 returns (b) Portfolio daily returns (c) DAX daily returns 36

C.1. Kolmogorov-Smirnov table 53

List of Tables

4.1. Portfolio preliminary statistics . . . 23
4.2. FTSE100 preliminary statistics . . . 23
4.3. SP500 preliminary statistics . . . 24
4.4. DAX30 preliminary statistics . . . 25
4.5. NIKKEI225 preliminary statistics . . . 25
4.6. SP500 estimated parameters . . . 28
4.7. DAX30 estimated parameters . . . 29
4.8. FTSE100 estimated parameters . . . 30
4.9. Portfolio estimated parameters . . . 31
4.10. NIKKEI225 estimated parameters . . . 32
4.11. NIKKEI225 goodness-of-fit . . . 37
4.12. SP500 goodness-of-fit . . . 38
4.14. DAX30 goodness-of-fit . . . 39
4.15. Portfolio goodness-of-fit . . . 40

1. Introduction

In this thesis, we aim to test differences in skewness in terms of sample types, between different time periods, in normal market conditions, as well as during periods of Market Stress. For that purpose, we are going to analyse a five-parameter distribution, firstly introduced by Theodossiou (1998). The Skewed Generalized T distribution (henceforth named SGT) is a skewed version of the Generalized T (GT) Distribution presented by McDonald and Newey (1988). Afterwards, we are going to compare the SGT fitting into the tails with an Extreme Value Theory distribution which also accounts for skewness.

A convenient assumption usually made by financial institutions is that financial returns are normally distributed. The particular reason for this choice is that as the time scale increases, the empirical distribution of financial returns tends to look identical to a Normal distribution. Moreover, the normality assumption, although not required, is a commonly choosed feature in Markowitz's portfolio theory, in the CAPM theory, and in implementing the Black-Scholes option pricing model. It is likely that the financial returns follow a geometric Brownian motion process, in which the increments adjusts to a Normal distribution.

The main problem of the Normal distribution is the imposition of symmetry and mesokurtosis (kurtosis of 3). Empirical evidence shows that the observed distribution of daily returns data has a heavier lower tail and is leptokurtic and skewed to the left. Thus, the Normal distribution may not reveal the most proper fit.

Leptokurtosis can result from jumps in prices, correlations between shocks, changes in moment dependencies and the leverage effect. It can be reduced when standardizing the returns and transforming the data as shown by Ayán and Díaz (2008). Once the leptokurtosis is observed in the distribution of the given returns, the probability of large losses is greater than the one implied by the Normal distribution, as mentioned by Harris and Küçüközmen (2001). And so, using the Normal distribution, in this case, would mislead probabilities and result in deceitful Value at Risk (VaR) estimates.

SGT is a highly flexible distribution which provides a good fit to the empirical distribution of data, as studied by Theodossiou (1998). BenSaïda and Slim (2016) also came to the same conclusion when analysing the SGT and the Generalized Hyperbolic, confirming the superiority of these two distributions in approximating the empirical distribution of the chosen data at a notable exactness. This way, for the fitting process, we use the maximum likelihood estimates method, also utilized and programmed by BenSaïda for the Skewed Generalised t, and the *Distribution Fitter* app for the Generalised Pareto. The results show that the SGT surpass the skewness and kurtosis problem of the Normal distribution. Overall, these findings are encouraging for using the SGT distribution to estimate market risk measures, such as VaR. An empirical study (Liu et al., 2009) inclusively shows that the SGT distribution benefits the VaR estimations.

In addition to this Section, the study is organized into five sections. Section 2 reviews the state-of-the-art literature regarding the assumptions of the Normal and alternative distributions. Section 3 focuses on the concepts and methodologies used in the fitting process and on the validation of the studied distributions. Section 4 reveals the empirical study, describing the data and the estimated parameters as well as the results of the goodness-of-fit process. Finally, Section 5 concludes the dissertation by restating the critical outcomes and conclusions and showing insights for future research on this subject by analysing alternative assets, distinct time periods and distributions, and their use to estimate VaR.

2. Literature Review

2.1. Deviations from the assumptions of the normal distribution

The most popular distribution in statistics is the normal distribution. It is common to choose this distribution to describe a sample or the results of a procedure. The reason for this preference to other distributions is due to the simplicity of the parameter's evaluations and partly to the central limit theorem and some direct implications.

To fit a normal distribution in a specific data sample, we need to know how wrong the best approximation may be. Therefore, we need to verify some assumptions. The Assumption of Normality says, in a straightforward way, that if we take the mean from independent samples and write it down, these values will have a perfect bell shape. Therefore, the distribution of the means will converge to a normal distribution as the number of samples increases. Another reason to assume the sample follows a normal distribution is that the sample mean and the sample variance are independent when observations are independent of each other. The only distribution that held this assumption is the Normal distribution, which explains its preference of choice versus alternative distributions.

Following Mordkoff (2016), the first property of the Normal distribution claims that the distribution of the sample means is centred on the population's mean and there is perfect symmetry around it. The second property says that it is a unimodal distribution. Third, the normal distribution is asymptotic, meaning that the further we get from the mean, the closer we get to the x-axis. Lastly, the mean, the mode and the median are equal. This shows the perfect bell shape of the distribution, which has just one peak exactly on the mean, mode and median, and half the values are in the right while the other half on the left side of the mean.

The Normal distribution is known for having a skewness value of zero due to the symmetric property and kurtosis of 3. The coefficients of skewness and kurtosis are both the third and the fourth standardized moment of a distribution (a detailed description of moments is presented in the Appendix).

Sometimes, a distribution may not be symmetric with respect to its mean. One of the tails may be heavier than the other, in which case, we say that the distribution is skewed. The skewness can take positive or negative values depending on which tail is heavier. If the heavier tail is on the right, we have a positive asymmetry and if it is on the left side, we have negative asymmetry. While a Normal distribution is symmetric to its mean, the majority of the time the empirical distribution shows a relative skewness. Johansson (2005, 11) analyses the Swedish stock market in which research he claimed that "an asset with positive (negative) co-skewness reduces (increases) the risk of the portfolio to large absolute market returns and should yield a lower (higher) expected return in equilibrium". The co-skewness term represents the correlation of residuals from the regression of the market returns with the squared market returns.

Hair Jr et al. (2021) offer a general guideline to interpret the excess skewness, saying that a substantially skewed distribution occurs when the skewness is lower than -1 or greater than +1. Additionally, when the absolute value for skewness is lower than 0.5, the distribution is fairly symmetrical, and it is moderately skewed when it is in a range of [-1,-0.5] or [0.5,1]. Finally, for values greater, in absolute terms, than 1, the distribution is highly skewed.

In terms of kurtosis, Peiro (1999) studies skewness and kurtosis in financial returns and highlights the high kurtosis appearance in the empirical distribution of returns. All his studied distributions showed a more peaked and heavier tail than the normal distribution. He also connects the high skewness with the leptokurtic data.

Meanwhile, some implicated problems are detected in financial modelling theory. Since the Capital Asset Pricing Model (CAPM) only assumes the mean and the variance of returns, higher moments such as skewness and kurtosis do not take place on its evaluation. Thus, assuming a normal distribution is favourable. Despite this, a skewness implementation would benefit the model since a positive skewness is tempting, since it means a few large gains and recurrent small losses for the investor. The Black-Scholes option-pricing model also presents some accuracy problems due to the lack of a skewness parameter. As explained by Corrado and Su (1996), the model misprices deep-in-the-money and deep-out-of-the-money options.

Lastly, we focus the calculation of Value at Risk (VaR). The VaR quantifies the monetary value of the expected losses within a given probability of occurrence. So, it represents the maximum potential loss of a portfolio over a particular time horizon at a certain confidence level. There is not a single way of computing VaR. Analysts may use a historical simulation method, a Monte Carlo simulation or a variance-covariance method. The last approach assumes that the financial returns follow a Normal distribution. Therefore, only the expected return and the standard deviation is required to estimate VaR. But, the skewness and kurtosis risk also have implications in the evaluation of the VaR. If either is ignored, the VaR calculations will be flawed. As Liu et al. (2009) referred, the assumptions of the SGT distribution benefit the VaR estimation, which shows a need to consider fat-tails, leptokurtosis and skewness behaviours in VaR models. Bali and Theodossiou (2007) also analyses the conditional VaR incorporating the SGT and compared the relative performance of the SGT and the normal distribution in the construction of the average expected value of losses above VaR, also known as expected shortfall.

2.2. Alternative distributions

2.2.1. The Skewed Generalised T Distribution

The skewed generalised t-distribution (SGT), firstly introduced by Theodossiou (1998), is a skewed version of the Generalized t (GT) distribution proposed ten years before by McDonald and Newey (1988). The SGT is a highly flexible five-parameter distribution that accommodates the skewness and excess kurtosis, unlike the broadly used Normal distribution. It has enormous applications, such as on the VaR models to estimate the risk, or the CAPM application for evaluation of regression estimation methods and intercept bias.

Our study focus on the fit of the SGT distribution into the empirical unconditional distribution of financial data. However, Kucukozmen et al. (2004) applied this same distribution into the modelation of the conditional distribution of daily equity returns. The Generalised Autoregressive Conditional Heteroscedasticity (GARCH) class of models is widely used in the estimation of the variance of conditional distribution, since the conditional distribution is needed in order to present a time-varying volatility. The use of SGT distribution is a great improvement in the fit of these GARCH and EGARCH models.

Countless researchers have come to different ways to present the SGT probability density function. In a first approach, Theodossiou (1998) presented it as a non-location-parameter distribution. This way the new skewness parameter would be the rate of descent of the density around $x = 0$. Right after then, it is introduced the first moment, which represents the mean, so that it incorporates the location parameter when using a transformed random variable $z = x - \mu$, where μ is the mean of the random variable x ($\mu = E(x)$). Hence, we represent the probability density function of the SGT distribution of the random variable z as:

$$f(z|\mu,k,n,\lambda,\sigma^2) = \begin{cases} f_1 = C(1+(\frac{k}{n-2})\theta^{-k} \times (1-\lambda)^{-k}|\frac{z+\mu}{\sigma}|^k)^{-(n+1)/k} & z \leq -\mu \\ f_2 = C(1+(\frac{k}{n-2})\theta^{-k} \times (1+\lambda)^{-k}|\frac{z+\mu}{\sigma}|^k)^{-(n+1)/k} & z \leq -\mu \end{cases} \quad (2.1)$$

where k, n, λ, μ, and σ^2 are scaling parameters. The following restrictions are required $k > 0$, $n > 2$, $-1 < \lambda < 1$. k and n are shape parameters controlling the height and tails of the density while λ is the skewness parameter. σ^2 is the variance or second centralized moment, with μ being the mean, which represents a location parameter. C and θ are normalizing constants in order to verify all the properties of a probability density function for f and are defined as follows:

$$C = \frac{k}{2\sigma} B\left(\frac{1}{k}, \frac{n}{k}\right)^{-3/2} B\left(\frac{3}{k}, \frac{n-2}{k}\right)^{1/2} S(\lambda) \quad (2.2)$$

$$\theta = \left(\frac{k}{n-2}\right)^{1/k} B\left(\frac{1}{k}, \frac{n}{k}\right)^{1/2} B\left(\frac{3}{k}, \frac{n-2}{k}\right)^{-1/2} \frac{1}{S(\lambda)} \quad (2.3)$$

$$S(\lambda)=\left(1+3\lambda^2-4\lambda^2 B\left(\frac{2}{k},\frac{n-1}{k}\right)^2 B\left(\frac{1}{k},\frac{n}{k}\right)^{-1} B\left(\frac{3}{k},\frac{n-2}{k}\right)^{-1}\right)^{1/2} \tag{2.4}$$

where $B(\cdot)$ is the beta function.

Most researchers use a direct form for the SGT distribution density function, instead of a branch function, as introduced by the sign function. For this study, we are going to use the one presented by BenSaïda and Slim (2016). This one, apart from the way it is written, is also programmed in a flexible distributions toolbox containing essential tools related to the SGT distribution provided by Ahmed BenSaïda, which we must grasp and use to estimate the parameters and fit this distribution to real data. That way for the fitting process we are going to use his MatLab toolbox. Accordingly, for any random variable $x \in \mathbb{R}$ following an SGT distribution, the probability density function is given as:

$$f_{SGT}(x;k,n,\lambda,\mu,\sigma)=\frac{k}{2\theta^*\sigma B\left(\frac{1}{k},\frac{n}{k}\right)\left[1+\frac{|x-\mu^*|^k}{(1+sign(x-\mu^*)\lambda)^k\sigma^k\theta^{*k}}\right]^{(n+1)/k}} \tag{2.5}$$

We change the nomenclature to preserve the writing and mathematical guideline developed by Theodossiou (1998). Although we use of the computational program for the fitting process, we keep the original nomenclature and present values accordingly. It should also be noted that the new constant θ^* is just a slight change to the one used in Theodossiou's work.

$$\theta^*=\frac{B\left(\frac{1}{k},\frac{n}{k}\right)}{\sqrt{(1+3\lambda^2)B\left(\frac{1}{k},\frac{n}{k}\right)B\left(\frac{3}{k},\frac{n-2}{k}\right)-4\lambda^2 B\left(\frac{2}{k},\frac{n-1}{k}\right)^2}}=\left(\frac{k}{n-2}\right)^{-1/k}\times\theta \tag{2.6}$$

Moreover, the $x-\mu^*$ is going to be equivalent to $z+\mu$ from the first presented formula. Let $E(x)=m$, the value of μ^* is chosen according to the restriction:

$$\mu^*=m-2\sigma\theta^*\lambda\frac{B\left(\frac{2}{k},\frac{n-1}{k}\right)}{B\left(\frac{1}{k},\frac{n}{k}\right)}$$

It is easy to verify that the second term of the previous equality corresponds to the first moment which represents the expected value m (please see the Appendix for details about moments). This way it's introduced the location parameter regarding the mean, which informs where the distribution is located with relation to a null mean. To validate the veracity of BenSaïda and Slim's probability density function, we present in the appendix the demonstration starting at the new function and aiming at Theodossiou's function.

It is beneficial to remark that conditioning parameters can restrict qualities for the SGT distribution and generate other known distributions. Therefore, for $\lambda=0$, the SGT distribution, formerly presented, is symmetric and has all odd moments equal to zero, so it turns out to be directly McDonald and Newey (1988) generalised-t distribution. For $k=2$, the SGT distribution can spawn the skewed student t distribution. When $\lambda=0$ and $k=2$, n is interpreted

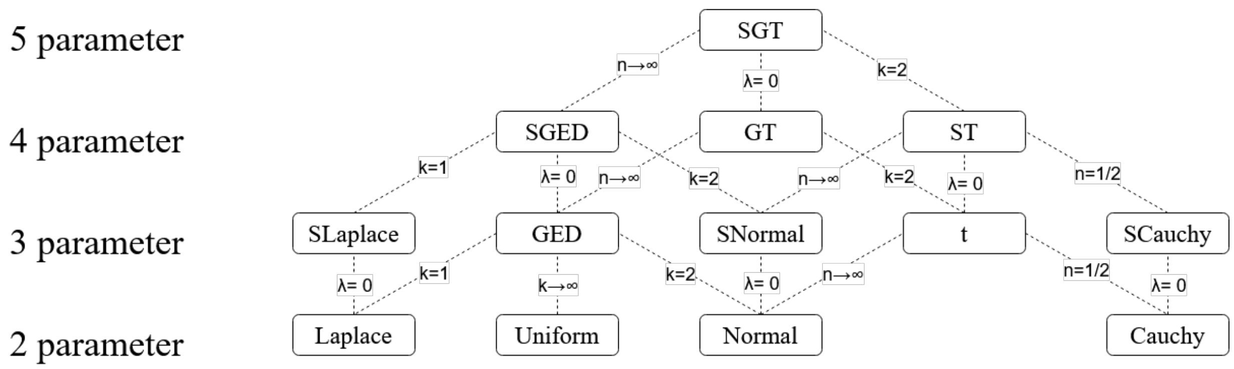


Figure 2.1.: Nested distributions of the SGT distribution modified from Sikora et al. (2019)

as the degrees of freedom and the student t distribution is obtained. For $\lambda = 0$ and $n \to \infty$, we get the Subbotin's power exponential or Box-Tiao distribution. The widely known normal distribution is obtained when $k = 2,\ \lambda = 0$ and $n \to \infty$. Cauchy's distribution is taken after imposing $\lambda = 0,\ k = 2$ and $n = 1$, while Laplace or double exponential distribution takes after $\lambda = 0,\ k = 1$ and $n \to \infty$. Lastly, the uniform distribution is obtained by restricting $\lambda = 0,\ k = \infty$ and $n = \infty$.

Figure 2.1 presents the aforementioned nested distributions of the SGT distribution including some jointed ones.

Next, the third and fourth standardised moments, respectively the skewness and kurtosis, for the SGT distribution, are given as follows:

$$S_k = \frac{m_3}{\sigma^3} = 4\theta^{*3}\lambda(1+\lambda^2)\frac{B(\frac{4}{k},\frac{n-3}{k})}{B(\frac{1}{k},\frac{n}{k})} - 2\theta^*\lambda\frac{B(\frac{2}{k},\frac{n-1}{k})}{B(\frac{1}{k},\frac{n}{k})}\left[3+4\theta^{*2}\lambda^2\frac{B(\frac{2}{k},\frac{n-1}{k})^2}{B(\frac{1}{k},\frac{n}{k})^2}\right] \tag{2.7}$$

$$\begin{aligned} K_u = \frac{m_4}{\sigma^4} = &\frac{\theta^{*4}[1+5\lambda^2(2+\lambda^2)]B(\frac{5}{k},\frac{n-4}{k})}{B(\frac{1}{k},\frac{n}{k})} + 24\theta^{*2}\lambda^2\frac{B(\frac{2}{k},\frac{n-1}{k})^2}{B(\frac{1}{k},\frac{n}{k})^2} \\ &\times\left[6+12\theta^{*2}\lambda^2\frac{B(\frac{2}{k},\frac{n-1}{k})^2}{B(\frac{1}{k},\frac{n}{k})} - 8\theta^{*2}(1+\lambda^2)\frac{B(\frac{4}{k},\frac{n-3}{k})}{B(\frac{2}{k},\frac{n-1}{k})}\right] \end{aligned} \tag{2.8}$$

where m_3 and m_4 are, respectively, the third and the fourth central moments. The proof and in-depth analysis of moments are presented in the Appendix. For the upcoming fitting process of the SGT distribution towards the empirical financial data, we apply the classical maximum likelihood estimation method also used by BenSaïda and Slim (2016).

2.2.2. Extreme Value Theory

The Extreme Value Theory (EVT) is a powerful framework capable of predicting the occurrence of rare events. Under this approach, the focal point is the tail behaviour. Probabilistically, it deals with the stochastic behaviour of problems of maxima and minima of i.i.d. (independent and identically distributed) random variables. In a wide selection of researches, the main field of study of this theory and its formed distributions is its applications in climatology, epidemiology, environmental studies, although there have also been some studies in the financial literature Bali (2003).

There is not a single methodology for modelling extreme events through EVT. It is possible to model a problem using a distribution of minimum or maximum procedures or just model the exceedances of a particular threshold. The first method leads to the Generalised Extreme Value distribution, proposed by Jenkinson (1955), while the other entails the creation of the Generalised Pareto distribution of Pickands III (1975). According to Gencay and Selcuk (2004, 290), "the GPD and the extreme value theory are an indispensable part of risk management in general and the VaR calculations in particular, in emerging markets". The afore paper shows the importance and the dominance of the GPD in modelling extremal events at a 0.999 percentile with 95% confidence intervals. In the next section, the Generalised Pareto distribution is going to be the main distribution in the study.

As presented in Embrechts et al. (2013), we introduce the main formalization of the Extreme Value Theory. The EVT can be compared to the central limit theorem, thereby all the methodology hereafter explained is comparable with the conclusions drawn on the central limit problem formally featured in Embrechts et al. (2013). The focal point of the EVT is the convergence of maxima. The question that Embrechts et al. (2013) tries to answer is which distributions satisfy for all $n \geq 2$ the identity in law $\max(X_1, ..., X_1) \stackrel{m}{=} c_n(X + d_n)$ for appropriate constants $c_n > 0$ and $d_n \in \mathbb{R}$. Suppose that $(X_t)_{t=1,2,...,n}$, is a sequence of independent and identical distributed random variables, all with common distribution function $F(x) = P[X_t \leq x]$ with mean μ and variance σ^2. Noting the definition of a max-stable distribution:

Definition 2.2.1 (Max-stabel distribution)**.** *If* X_t *is a non-degenerate random variable that satisfies the identity in law* $\max(X_1, ..., X_n) \stackrel{m}{=} c_n * X + d_n$ *for appropriate constants* $c_n > 0$ *and* $d_n \in \mathbb{R}$ *and every* $n \geq 2$*, it is said max-stable.*

We denote the sample of maxima of X_t by $M_1 = X_1,\ M_n = \max(X_1, ..., X_n)$ for $n \geq 2$. Then, assuming that (X_n) is a sequence of i.i.d. max-stable random variables, the law $\max(X_1, ..., X_1) \stackrel{m}{=} c_n * X + d_n$ can also be written as

$$c_n^{-1}(M_n - d_n) \stackrel{m}{=} X \tag{2.9}$$

This results in every max-stable distribution being a limit distribution for maxima of i.i.d random variables. Therefore, the *Limit property of max-stable laws* is given as follows:

Theorem 2.2.1 (Limit property of max-stable laws)**.** *The class of max-stable distributions coincides with the class of all possible (non-degenerate) limit laws for (properly normalised) maxima of i.i.d random variables.*

The Proof is explicitly formulated in Embrechts et al. (2013). Considering that the theory deals with the convergence of maxima, we present the *Fisher-Tippett theorem* as follows:

Theorem 2.2.2 (Fisher-Tippet theorem)**.** *Suppose that* (X_n) *is a sequence of i.i.d. random variables. For a sequence of* $c_n > 0$ *and* $d_n \in \mathbb{R}$ *and a non-degenerate distribution function* H *such that* $c_n^{-1}(M_n - d_n) \xrightarrow{d} H$*, then* H *belongs to one of the following three distribution functions:*

1. *Gumbel (type I):* $\Delta(x) := e^{-e^{-x}}$ *for* $x \in \mathbb{R}$

2. *Fréchet (type II):* $\Phi_\alpha(x) := \begin{cases} 0 & x \leq 0 \\ e^{-x^{-\alpha}} & x > 0; \alpha > 0 \end{cases}$

3. *Weibull (type III):* $\Psi_\alpha(x) := \begin{cases} e^{-(-x)^\alpha} & x \leq 0; \alpha > 0 \\ 1 & x > 0 \end{cases}$

The formal proof of the *Fisher-Tippet theorem* was first given in Gnedenko (1943). This last theorem is the most important EVT result since it offers three distributions where we can fit the asymptotic distribution of maxima disregarding the original empirical distribution. Most of all, these three families of distributions can be generalised with the incorporation of location (μ) and scale (σ) parameters. *Fréchet* and *Weibull* families gain the *Gumbel* family form when its α parameter extends through $+\infty$ and $-\infty$, respectively. That way, the three distributions may be nested in a single parameter family, the named Generalised Extreme value distribution. This family of distributions are present in Figure 2.2. Reparametrizing $\xi = 1/\alpha$ and bearing in mind $1 + \xi x > 0$, it is attainable the representation introduced by Jenkinson (1955) and Von Mises (1936).

$$GEV_\xi(x) = \begin{cases} e^{-(1+\xi x)^{-\frac{1}{\xi}}} & \xi \neq 0 \\ e^{-e^{-x}} & \xi = 0 \end{cases} \tag{2.10}$$

Tolikas (2008) also shows a written form for the probability density function of a standardised variable. By standardising x we get the probability density function given by:

$$gev_{\mu,\sigma,\xi}(X) = \alpha^{-1} e^{-(1+\xi)y} e^{-e^{-y}} \tag{2.11}$$

where

$$y = \begin{cases} \xi^{-1} ln\left(1 + \xi \frac{X-\mu}{\sigma}\right) & \xi \neq 0 \\ \frac{X-\mu}{\sigma} & \xi = 0 \end{cases} \tag{2.12}$$

and, σ, μ, ξ are called the scale, location, and shape parameters, respectively.

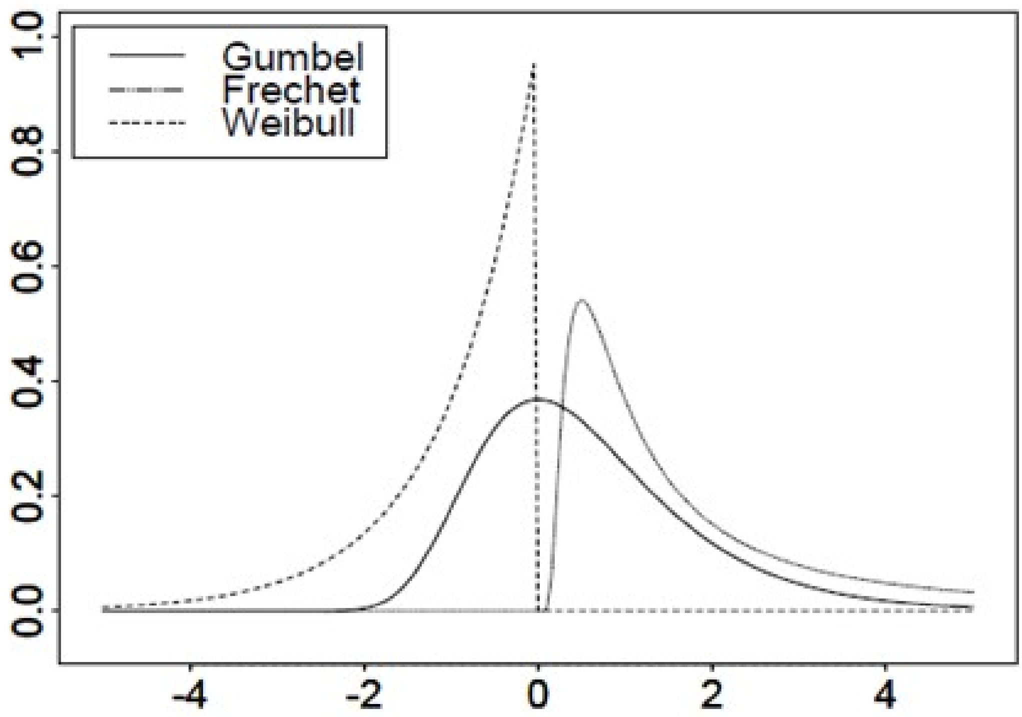


Figure 2.2.: EVT distributions comparison from Embrechts et al. (2013)

High σ values suggest that the distribution of extremes is widely spread out, while a high μ implies large extremes. Recapping the meaning of different values of ξ and the associated distributions we obtain:

- $\xi > 0$ corresponds to the *Fréchet distribution*, which has an infinite right endpoint.
- $\xi = 0$ corresponds to the Gumbel distribution, which has an infinite right endpoint and the tail decays much faster than the tail of *Fréchet distribution*.
- $\xi < 0$ corresponds to the *Weibull distribution* which is a short-tailed distribution with finite right endpoint (after a certain point there are no extremes).

Bali (2003) introduced an even more generalisation of the Generalised extreme value distribution. Using the Box and Cox (1964) transformation, Bali proposed a new distribution nesting the GEV distribution of Jenkinson (1955) and the hereafter spoken Generalised Pareto distribution of Pickands III (1975). The new Box-Cox-GEV distribution is given as follows:

$$BCGEV_{\max,\xi}(X;\mu,\sigma,\lambda) = \left(\frac{\left[exp\left\{ -\left[\xi\left(\frac{M-\mu}{\sigma}\right)\right]^{-1/\xi} \right\} \right]^{\lambda} - 1}{\lambda} \right) + 1 \tag{2.13}$$

When $\lambda = 1$ the distribution reduces to the GEV distribution and when $\lambda = 0$ the distribution reduces to the GPD by *L'Hopital's rule*, as indicated in Bali (2003). Additionally Bali rejects GEV distribution and GPD over the newly presented *Box-Cox-GEV distribution*.

2.2.3. Generalised Pareto distribution

Despite the existence of an inclusive distribution, the one formerly introduced, the most popular approach to model the financial extremes is the Generalised Pareto distribution (GPD), firstly presented by Pickands III (1975). This one is a generalisation of the Pareto distribution, studied in detail by Arnold (2014). There is abundant literature that concludes the importance of the GPD (e.g., (Embrechts et al., 2013; Das and Halder, 2016)). Embrechts et al. (2013) and Packham et al. (2017) show that the GPD is a useful tool in modelling extremal events and distributions over a high threshold and in avoiding tail risks. Embrechts et al. (2013) test and evaluates the suitability of the GPD model through the VaR spread. The VaR spread stands for the spread between the VaR from a GARCH model following a GPD and a VaR from a GARCH model following a Normal distribution.

Although the GPD is used for making interpretations about the upper tail, in this work, we analyse only the lower tail. In fact, it is more important to prevent great drops and analyse the behaviour of financial losses than to anticipate great profits and analyse their forward behaviour. Aiming to understand the GPD, we use the definition and explanation given in Embrechts et al. (2013). Since excesses over thresholds are fundamental in different fields,a *mean excess function* (MEF) is defined:

$$e(u) = E(X - u | X > u) \tag{2.14}$$

where X is a random variable with density function F and right endpoint x_F, where, for a fixed $u < x_F$, the excess density function of the random variable over the threshold u is given as

$$F_u(x) = P(X - u \leq x | X > u) \qquad \text{for } x \geq 0 \tag{2.15}$$

$F_u(x)$ donates the distribution function of exceedances above the threshold u. By the conditional probabilities, F_u can also be defined as

$$F_u(x) = \begin{cases} \frac{F(u+x)-F(u)}{1-F(u)} & x \geq 0 \\ 0 & \text{else} \end{cases} \tag{2.16}$$

Therefore, for any u and x, $[1 - F(u+x)]/[1 - F(u)]$ is the conditional probability that an observation is greater than $x+u$, given that it is greater than u, i.e.:

$$P(X > x + u | X > u) = 1 - F_u(x) \qquad \text{for } x > 0 \tag{2.17}$$

Solving F, the distribution of the threshold exceedances may also be solved. Let y be a random variable given as $y = x + u$ for a restricted $X > u$. Then, we can represent $F_u(x)$ as follows:

$$F(y) = [1 - F(u)]F_u(x) + F(u) \tag{2.18}$$

Pickands III (1975), presented a theorem, alongside Balkema and De Haan (1974), that shows that for a sufficiently high threshold u, the distribution function of excess may be approximated by the GPD. As mentioned in Das and Halder (2016, 46) "high values of the threshold are preferred".

Therefore, the generalised Pareto distribution is a three parameters distribution with a location parameter μ, a scale parameter σ, and a shape parameter $\xi = 1/\alpha$, where α is known as the tail index. The shape parameter ξ is the most important, since the GPD can be specified using only this parameter. As Das and Halder (2016) mentioned, the 2-parameter GPD respecting $\mu = 0$ is more practical in some cases. As a matter of fact, it eases the fitting process. The standardized cumulative distribution function of the GPD of a random variable x can be written as:

$$GPD_{\mu,\sigma,\xi}(x) = \begin{cases} 1 - \left(1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-\frac{1}{\xi}} & \xi \neq 0, \sigma < 0 \\ 1 - exp\left\{-\frac{x-\mu}{\sigma}\right\} & \xi = 0, \sigma < 0 \end{cases} \tag{2.19}$$

with restrictions

$$\begin{array}{lll} x \geq \mu & \text{if} & \xi \geq 0, \quad \sigma < 0 \\ \mu \leq x \leq \mu - \frac{1}{\xi} & \text{if} & \xi = 0, \quad \sigma < 0 \end{array} \tag{2.20}$$

Since the MEF (*mean excess* function) presented before can also be seen as the sum of the excesses over a threshold u divided by the number of data point that exceeds such threshold, if the empirical MEF is positively sloped and is a straight line, the respective GPD must have a positive shape parameter ξ.

The corresponding probability density function (pdf) is simply the first derivative in order to the random variable x of the cumulative distribution function (cdf). For the GPD, it is:

$$f_{GPD}(x) = \frac{dGPD_{\mu,\sigma,\xi}}{dx} = \begin{cases} \frac{1}{\sigma}\left(1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-\frac{\xi+1}{\xi}} & \xi \neq 0, \sigma < 0 \\ \frac{1}{\sigma}exp\left\{-\frac{x-\mu}{\sigma}\right\} & \xi = 0, \sigma < 0 \end{cases} \tag{2.21}$$

When $\mu = 0$ and $\sigma = 1$ the GPD is known as the standard GPD. As observed in the previous SGT distribution, different values of the parameters contribute to a number of nested distributions within the GPD. We can start by relating the GPD with the GEV distribution mentioned above, its relationship can be given for $ln(GEV_{\mu,\sigma,\xi}(x)) > -1$ as:

$$GPD_{\mu,\sigma,\xi}(x) = 1 + ln(GEV_{\mu,\sigma,\xi}(x)) \tag{2.22}$$

The other embedded distributions are given restricting the parameters as follows: For $\xi = 0$ ($\alpha \to \infty$) the GPD corresponds to the exponential distribution. When $\xi < 0$ the distribution has a heavy upper tail making it equivalent to a Pareto II type distribution. When $\xi > 0$, the GPD corresponds to the original Pareto distribution; restricting the parameters as $\xi = 0.5$ the distribution takes the form of a Triangular distribution; when $\xi = 1$, it takes the form of a uniform distribution. It should be noted that for $\xi > 0$, the nth central moment, $E[X^n]$, is infinite for every $n > 1/\xi$. For instance, with the triangular distribution, $\xi = 0.5$, the distribution has infinite variance since the variance is also the 2^{nd} central moment. When $\xi = 1/3$ the distribution has infinite skewness, and when $\xi = 0.25$, it has infinite historical kurtosis.

The nested distributions restricting the 3-parameter GPD to a 2-parameter distribution can be seen in the diagram presented in Figure 2.3. Papastathopoulos and Tawn (2013) proposed an extension of the GPD that incorporates an additional shape parameter while keeping the tail behaviour unaffected.

For the upcoming fitting process of the GPD towards the empirical financial data, GPD parameters may be estimated through the *classical maximum likelihood method*. This process was also used in several works of literature.

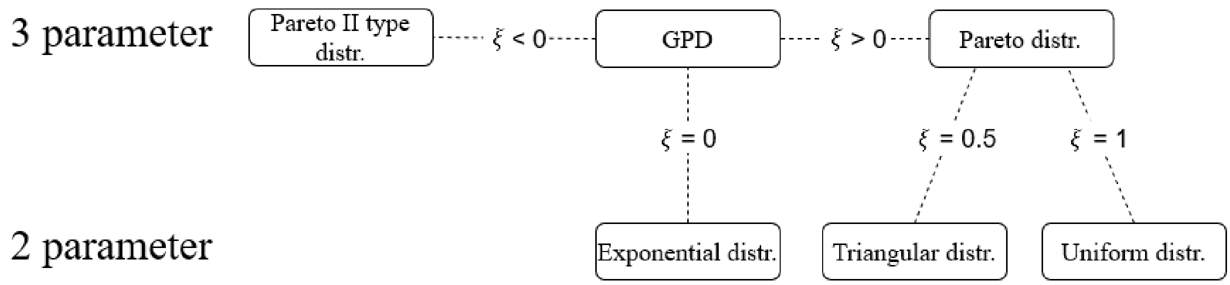


Figure 2.3.: Nested distributions of the GPD

3. Concepts and Methodologies

The parameter estimation of each distribution can be performed through different methods. Sikora et al. (2019) performed a *Chan-Karolyi-Longstaff-Sanders* (CKLS) *model* parameters estimation. This one is defined by the Itô stochastic differential equation $dX = (\alpha + \beta X)dt + \sigma X^d dB$, where $B(t)$ is a Brownian Motion, which is an extension of the *Vasicek model*. In this and some other literature (e.g. Das and Halder (2016)), the method of moments and its generalisation (GMM) is also discussed, as well as the *Probability Weighted Moments* (PWM) introduced by Greenwood et al. (1979), and the *Maximum Penalised Likelihood* (MPE) introduced by Coles and Dixon (1999). In some empirical evidence, the *maximum likelihood method* (MLE) is more sophisticated and harder to implement than the method of moments (MM). Then again, the ahead fitting process is going to be performed upon the *Maximum Likelihood Method*, since, as cited in Das and Halder (2016, 57), "MLE provides the estimates with lowest standard error".

3.1. Maximum Likelihood Estimation

For the SGT distribution, we need to estimate five parameters, whereas, for the GPD, only two parameters need to be estimated, since the location parameter will be approximated according to the empirical mean.

To start with the interpretation of the method at hand, we introduce the likelihood function, which for independent observations $X_1, X_2, ..., X_n$ from a probability distribution, is a k-parameters function

$$L(x_1, x_2, ..., x_n; p_1, p_2, ..., p_k) = \prod_{i=1}^{n} f(x_i; p_1, p_2, ..., p_k)$$

where $f(\cdot)$ represents the probability density function (p.d.f.) of the observation sample when it is absolutely continuous and represents the probability mass function (p.m.f.) when it is discrete. The joint p.d.f of $X_1, X_2, ..., X_n$, can be seen as a function of the k parameters $p_1, p_2, ..., p_k$, or simply a function of θ, being $\theta = (p_1, p_2, ..., p_k)$ a vector, as shown in Hoel et al. (1954).

The likelihood function examines the probability or the probability density behaviours through different functions in F, and thus explains the observed sample results based on the evaluation of the likelihood for the different probability distributions of the F family.

The maximum likelihood estimation method (MLE) is a parameter estimation method for a certain probability function. The likelihood function can act as a function completely composed of unknown parameters, $\theta = (p_1, p_2, ..., p_k)$. From DeGroot (2012), the frugal point is to find the vector in the restricted space Ω, letting Ω be the set $-\infty < \theta < \infty$, that maximises the likelihood function. The parameters of this vector are the estimate parameters. Subsequently, supposing that a function of $x_1, x_2, ..., x_n$ is found, $u(x_1, x_2, ..., x_n)$, such that, replacing θ with, the likelihood function L is maximum. In other words, for every $\theta^* \in \Omega$, $L(x_1, x_2, ..., x_n; \theta) \geq L(x_1, x_2, ..., x_n; \theta^*)$ and θ is called the maximum estimator, represented as $\hat{\theta}$.

When θ is a vector of parameters, $\hat{\theta}$ is obtained by solving a system of equations $\frac{\partial log(L)}{\partial p_i} = 0$ for $i = 1, k$. Suppose the system has a unique solution, such that its corresponding matrix for the second-order derivative is positive semi-definite, from optimisation theory. In this case, the solution shall be considered as a maximum and, therefore, is the estimate. If there are multiple solutions, the existence of the estimate shall not be considered.

The likelihood function will change in respect to the distribution in use. We shall present the likelihood function for the SGT, GP and Normal distribution:

3.1.1. SGT Likelihood function

In the SGT distribution, the estimators will be obtained using the log-likelihood function. On that account, the log-likelihood function is:

$$L(\theta) = \sum_{i=1}^{n} L_i(\theta) = \sum_{i=1}^{n} \ln f_{SGT}(x_i; k, n, \lambda, \mu, \sigma) \tag{3.1}$$

remembering that the f_{SGT} function corresponds to the function present in equation 2.1 and with respect to the vector $\theta = (k, n, \lambda, \mu, \sigma)$ with the parameters delimited by $|\lambda| < 1$, $n \leq 4$ and $k \leq 0$.

The maximisation of the above log-likelihood function can be revealed as complicated. Despite this, interactive algorithms can overcome these cases. Since the fitting process for the SGT distribution is going to be performed according to the BenSaïda and Slim (2016)'s given program, and this one is implemented in MatLab (2021), maximisation and equation solver algorithms are performed. As a rule, the previous parameters delimitations are critical in the program implementation since maximisation algorithms can easily "flee out" the restrictions for $\sigma^2 > 0$ and $|\lambda| < 1$. The programming process description is presented further on.

3.1.2. GP Likelihood function

For the GPD, the likelihood function for estimation of two parameters (considering the location parameter, μ, as a first approximation and not an estimation) is

$$L(\theta) = \prod_{i=1}^{n} f_{GPD}(x_i; \xi, \sigma) \tag{3.2}$$

where the f_{GPD} function corresponds to the function present in equation 2.21 and with respect to the vector $\theta = (\xi, \sigma)$. As suggested in Das and Halder (2016), to get better fitting parameters for the GPD, it is essential to have at least 500 observations and the ξ values must be within the bounds $-1/2$ and $1/2$.

Grimshaw (1993) presents an algorithm for computing the GPD maximum likelihood estimates ξ and σ. Later on, more researchers started studying the topic with different maximisation algorithms and even proposed an equivalent MLE method taking into account the efficiency and robustness as revealed in Pham et al. (2019). These reseachers presented an algorithm structure using the Newton-Raphson algorithm and wrote it in R.

3.1.3. Normal Likelihood function

For the Normal distribution, the probability density function is:

$$f_{normal} = \frac{1}{\sqrt{2\pi\sigma^2}} \exp - \frac{(x-\mu)^2}{2\sigma^2} \tag{3.3}$$

And the likelihood function for this case is:

$$L(\theta) = \prod_{i=1}^{n} f_{normal}(x_i; \mu, \sigma) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp - \frac{(x-\mu)^2}{2\sigma^2} \tag{3.4}$$

3.2. Goodness of fit

The goodness-of-fit test allows the verification of the adequacy of a specific probability distribution into an observed distribution. This way, it is possible to analyse if the empirical data was derived from a population that follows that distribution function.

The goodness-of-fit is an essential step in fitting an estimated distribution to the real data distribution. It helps to determine whether the observed data resembles the expected values. It is represented by a statistical method that inferences about the experimental values to see if the actual values are sufficiently close to the distribution generated values. As presented in Dodge (2008), the process involves hypothesis testing: Being the null hypothesis H_0 is the equality of the unknown distribution function of the underlying population with the presumed distribution function versus the alternative hypothesis H_1 of the inequality.

Once estimated the parameters through the MLE and presuming a certain distribution, the comparison among estimates is done using the Kolmogorov-Smirnov (KS) test, the Anderson and Darling (1954) (AD) test as well as the Kuiper (1960) test. The choice of use of these tests were to vary the tests used in other literature. BenSaïda and Slim (2016) also tests the Bayesian Information criterion per observation (BIC0), while Theodossiou (1998) uses the Kolmogorov-Smirnov test as well as the Log-Likelihood ratio test (LR) for testing the normality against the SGT distribution. Chu et al. (2019) presented several goodness-of-fit tests for Pareto distributions, and despite the existence of many other goodness-of-fit tests, the tests mentioned above are held for both distributions.

Although the Chi-Square test is the most common goodness-of-fit test, it is typically used in discrete distributions. Sprent (2012) expressed that sometimes one or more parameters must be estimated from the data. In conjunction with this, it has a few other problems, such as the number of samples required to produce a reliable result. Thus, the choice of the goodness-of-fit tests is fundamental.

3.2.1. Kolmogorov-Smirnov test

The Kolmogorov-Smirnov (KS) test is a refinement of the Kolmogorov test. It is a nonparametric goodness-of-fit test that evaluates the adequacy of a unidimensional continuous probability function with the observed distribution. Although it can sometimes be applied for discrete distribution, the test criteria are no longer exact, following inefficient results. The main focus is to quantify the distance between the empirical cumulative distribution function of the observed data and the cumulative distribution function of the presumed distribution. Contrary to the Chi-Square test, the KS test is exact and does not depend on the test's underlying cdf. Despite these advantages, the KS has some significant limitations that should be cared (NIST/SEMATECH, 2021). It tends to be more sensitive near the centre than at the tails, which contrasts with the purpose of this dissertation, where the main goal is to analyse the skewness. Moreover, it is recommended for large samples of over 2000. Despite this, numerous pieces of literature still use it, and it is an almost standard goodness-of-fit test on its simplicity. The Kolmogorov-Smirnov hypothesis testing considers in the null the hypothesis that the empirical data follows the theoretical distribution. The respective test statistic is defined as:

$$D = \sup_x |F(x) - F_n(x)| \tag{3.5}$$

where F is the hypothesised cumulative distribution, which must be continuous, and F_n is the empirical distribution function. F_n may be defined as:

$$F_n(x) = \begin{cases} 0 & x < x_1 \\ \frac{k}{n} & x_k \leq x \leq x_{k+1} \\ 1 & x \geq x_n \end{cases} \tag{3.6}$$

The null hypothesis is rejected if the test statistic, D, is greater than the critical value. The KS test is distribution-free, meaning that it does not depend on the distribution to evaluate the critical values. The statistical table is in Appendix C. MatLab has two types of *KS tests*. Since the aim is to compare the Normal, the SGT, and the GP distributions fit the empirical data, the *kstest2* is used.

3.2.2. Anderson-Darling test

The Anderson and Darling (1954) test is a refinement of the KS test. Being a goodness-of-fit test, it also evaluates if the given data sample is drawn from the proposed probability function. Unlike the KS test, the AD test gives more attention to the tails, and is more potent than the KS test. It similarly defines in the null the hypothesis that the empirical data follows the theoretical distribution. The test statistic is defined in this case as

$$A^2 = -N - S, \text{ where } \quad S = \sum_{i=1}^{N} \frac{(2i-1)}{N} [\ln F(Y_i) + \ln 1 + F(Y_{N+1-i})] \tag{3.7}$$

and F is the cumulative distribution function of the specified distribution.

Again, the null hypothesis is rejected if the test statistic, A, is greater than the critical value. Here, the *AD test* makes use of the proposed distribution to evaluate the critical values. This is beneficial for sensitivity but disadvantageous since it is necessary to evaluate each distribution's critical values. The tables for the critical values of each distribution are not provided here. However, the test is implemented in MatLab through *adtest()* for all the Normal and the GPD. For the SGT distribution, the *adtest()* was modified to assume this distribution since MatLab does not have an SGT probability distribution object. Therefore, the modified code *adtestsgt()* is annexed. It should also be noted that the SGT tends to look more like a Normal distribution for a large number of samples. Hence its critical values are estimated according to the Normal's critical values.

3.2.3. Kuiper Test

The Kuiper (1960) goodness-of-fit test is a modification of the Kolmogorov-Smirnov goodness-of-fit test. It also tests the null hypothesis that the given data comes from a population with a continuous distribution function. Again, all of our distributions are continuous, having no problem in the implementation of this test. It is test also considers in the null the hypothesis that the empirical data follows the theoretical distribution. The respective test statistic is defined as

$$V = \sup(F_N(x) - F(x)) = -\inf(F_N(x) - F(x)) \quad \text{for} \quad -\infty < x < \infty \tag{3.8}$$

where F is the hypothesised distribution function and, F_N is the sample distribution function. Similar to the Kolmogorov-Smirnov statistic, the distribution is independent of $F(x)$. This test is also comparable with the *Anderson-Darling test*. Both are equally sensitive at the tails and the median. However, the *Kuiper test* is invariant under cyclic variations. This means that if some occurrences take place in cyclic times, both the KS and the *Kuiper test* will miss it.

For its implementation, we use BenSaïda (2021)'s *Kuiper test* function *kuipertest()*, based on the exact distribution of Stephens (1965), where significance level percentages tables are presented.

4. Empirical Study

In this section, we present the data description with all the preliminary statistics of the empirical data. Each data series is presented in the respective table with all the information for a better analysis.

Right after, the fitting process is also done with the respective tables of the estimated parameters as well as the goodness-of-fit tables and graphic representation. Not all the studies are presented here, instead, the less marked are present in the appendix.

4.1. Data Description

In this dissertation, the financial series used consist of daily, weekly, and monthly observations on the S&P500 stock market index for the United States, the FTSE100 stock market index for the United Kingdom, the DAX30 stock market index for a representative European Union stock market, represented by the Deutsch index, and Nikkei225 stock market index for a significant Asian portion, represented by the Japanese index. The data for each series is retrieved from Yahoo Finance and Investing.com. The collected data correspond to 21 years, starting on January 1st of 2000 and ending on December 31st of 2020, with different data frequencies being considered.

Different sub-periods are defined within the whole period to analyse the results obtained in stressed market conditions. We define a stress market condition according to the definition of Bear Market presented in Investopedia (2021). A Bear Market is a prolonged period characterised by a downward stock price trend. Typically, we assume that the price declines in a Bear market when there is a fall of at least 20% from the most recent high (a similar but symmetric definition applies to a Bull market, characterised by a rise of 20% from the most recent low). A Bear Market should not be mistaken with a Correction, which is a short-term downturn with duration no longer than two months and characterised by a decline of at least 10%. We define the sub-periods with and without stress according to the Bear Markets related to the S&P500 index.

The first decade of the 2000s was marked by two prolonged bear markets, with great declines in the S&P 500 index value. This decade represented the worst period for the US stock market since the Great Depression. Since many investors assume different beginnings and ends for the corrections, we define the bear market periods suggested in SPGlobal (2021). The first bear market began on March 24th of 2000 and ended on October 9th of 2002, and registered a cumulative loss of 49.1%. The dot-com bubble triggered this bear market or also known as the "dot-com crash", which followed the decade-long bull market and from the higher level of uncertainty and panic caused by the 9/11 terrorist attack. The second bear market started on October 9th of 2007, and ended on March 9th of 2009, leading to a loss of 56.8%. The collapse of the subprime mortgage was boosted due to severe growth in the debt and housing bubbles.

Therefore, we get four partitions, three related to stress market conditions represented by the bear market periods, and one for the remaining observations along the original 21-years period. The stressed sub-periods are carried out only in daily returns with the intention of having enough samples. For this case, it does not make sense to consider monthly and yearly frequencies since we need a reasonable sample size to reach significant and accurate estimates.

To do a portfolio analysis, presented below in Table 4.1, we also create a portfolio with the indices presented. Given the importance of all indices, the portfolio consists of 25% of each index (all data is annexed and freely available in .csv format).

The level series were transformed into continuously compounded returns. Using the logarithmic changes, the series are a stationary process. This is demonstrated by using the augmented version of the Dickey and Fuller (1979) test, where the null hypothesis states that a unit root is present in the autoregressive time series, and the alternative hypothesis is that the time series is a stationary process.

Finally, for the analysis of the GP distribution, the fitting process will be performed in the left tail, under a threshold of 5% and 20%. All the evaluation and representation of the GP distribution is performed having in mind the distribution for the loss values. Both cases are analysed, and the best choice to opt for a threshold is made.

The preliminary statistics are presented. In a first broad analysis, the majority of the empirical means for the stress subperiods are negative. Only for the FTSE100, presented in Table 4.2, all the three series representing the entire period of 21 years did capture negative values for the mean. This may result from distinct reasons. Environmental effects as 2005 hurricanes, the Sub-Prime Housing Crisis, the Global Recession, the Brexit Vote and results and the COVID-19 Pandemic did a great impact on the U.K. economy. Adding to this, the monthly series for the S&P500, presented in Table 4.3, also captured a negative mean. Although the frequent negative means, all of them, except for the third stress partition, assume an almost null absolute value. The third stress partition representing the COVID-19 Pandemic Bear Market is a short duration effect which implies an outstanding mean and standard deviation. This should not represent symmetric distributions. Instead, the skewness factor, which expresses the distortion of the data set, is analysed.

The skewness assumes overall negative values. The weekly returns give overall the most negative values. They assume values lower than -1, except for the DAX30 in Table 4.4 and S&P500 in Table 4.3, respectively -0.90 and -0.96, which can be almost considered a highly skewed data set. The remaining return series are overall moderately skewed to the left side. The

Table 4.1.: Portfolio preliminary statistics

	Daily	Weekly	Monthly	P1	P2	P3	P4
# of observations	5327	1097	252	646	357	24	4296
mean	0.004	0.020	0.093	-0.119	-0.232	-1.916	0.049
std. dev.	1.262	2.785	4.709	1.476	2.108	3.703	1.072
skewness	-0.346	-1.11	-0.745	-0.08	0.17	-1.053	-0.195
kurtosis	10.059	12.722	4.565	4.266	6.396	3.935	6.193
excess kurtosis	7.059	9.722	1.565	1.266	3.396	0.935	3.193
minimum	-12.39	-24.16	-17.38	-5.285	-7.219	-12.39	-7.335
maximum	9.928	13.435	13.027	5.828	9.594	3.038	6.723
5th percentile	-1.966	-4.102	-9.571	-2.613	-3.942	-9.697	-1.682
95th percentile	1.906	3.903	6.426	2.31	3.094	2.419	1.73
Dickey-Fuller	2641.12	624.29	98.49	301.85	185.43	20.210	22.236
Jarque-Bera	11166.8	4545.18	49.024	43.811	173.26	5.312	1851.7

(1) The Augmented-Dickey-Fuller and Jarque-Bera statistics are present.

(2) Unless otherwise noted, all statistics reject the null hypothesis at one percent level of significance.

(3) Statistics whose value is higher than the critical values, implying the non rejection of the null hypothesis, are in bold.

Table 4.2.: FTSE100 preliminary statistics

	Daily	Weekly	Monthly	P1	P2	P3	P4
# of observations	5304	1096	253	641	358	24	4278
mean	-0.001	-0.006	-0.031	-0.088	-0.171	-1.629	0.032
std. dev.	1.199	2.487	4.082	1.408	2.172	3.362	0.995
skewness	-0.332	-1.266	-0.694	-0.232	0.087	-1.168	-0.132
kurtosis	10.864	15.282	4.122	4.480	6.651	4.390	6.130
excess kurtosis	7.864	12.283	1.122	1.480	3.651	1.390	3.130
minimum	-11.512	-23.631	-14.858	-5.589	-9.265	-11.512	-6.199
maximum	9.384	12.583	11.647	4.878	9.384	2.7538	5.904
5th percentile	-1.887	-3.771	-7.538	-2.386	-3.422	-9.053	-1.593
95th percentile	1.793	3.285	6.19	2.284	2.851	2.525	1.610
Dickey-Fuller	2848.81	623.27	127.21	322.68	209.69	17.570	33.315
Jarque-Bera	13763.9	7182.36	33.546	64.265	199.24	7.386	1759.16

(1) The Augmented-Dickey-Fuller and Jarque-Bera statistics are present.

(2) Unless otherwise noted, all statistics reject the null hypothesis at one percent level of significance.

(3) Statistics whose value is higher than the critical values, implying the non rejection of the null hypothesis, are in bold.

Table 4.3.: SP500 preliminary statistics

	Daily	Weekly	Monthly	P1	P2	P3	P4
# of observations	5284	1096	251	638	356	24	4263
mean	0.018	0.086	-0.398	-0.106	-0.233	-1.707	0.063
std. dev.	1.255	2.536	4.417	1.444	2.396	5.011	0.985
skewness	-0.393	-0.964	0.711	0.162	-0.057	0.02	-0.1465
kurtosis	13.963	10.808	4.456	4.315	6.779	3.012	7.678
excess kurtosis	10.964	7.808	1.456	1.315	3.779	0.012	4.678
minimum	-12.77	-20.084	-11.942	-6.005	-9.47	-12.77	-6.895
maximum	10.957	11.424	18.564	5.574	10.957	8.881	6.837
5th percentile	-1.915	-4.015	-7.244	-2.402	-4.628	-10.83	-1.55
95th percentile	1.723	3.559	8.34	2.226	3.628	6.74	1.527
Dickey-Fuller	3313.26	631.27	106.78	314.11	243.17	52.29	21.716
Jarque-Bera	26501	2953.7	43.33	48.78	211.97	0.0017*	3901.70

(1) The Augmented-Dickey-Fuller and Jarque-Bera statistics are present.

(2) Unless otherwise noted, all statistics reject the null hypothesis at one percent level of significance.

(3) Statistics whose value is higher than the critical values, implying the non rejection of the null hypothesis, are in bold.

S&P500 is the index that showed more positive skewness. S&P500 monthly returns spotted a moderately high positive skew with a value of 0.71. In a fair analysis, it is advisable to invest daily or monthly rather than weekly, since weekly returns show a high negative skewness. The first and the third partitions gave the other two positive skews, with 0.16 and 0.02, respectively. We consider that the respective distributions are fairly symmetrical. To sum up, the weekly returns are mostly highly negative skewed while the monthly returns have mostly a moderate-high negative skew. For the daily returns, the skewness tends to be moderately low, standing at about -0.35. The third subperiod, representing the COVID-19 Pandemic, exhibits a highly negative skew in almost all cases. These values result from a bear market of just one month that had a massive impact on the economy. The other two bear markets have, overall, an almost null skewness, showing symmetry. A comprehensive way to interpret this stands in the frequent small gains against the few substantial losses.

In terms of kurtosis, almost all the empirical data is leptokurtic. It has a positive excess kurtosis. The only exception of this trait is the covid crisis period for Nikkei 225, presented in Table 4.5, the same period that the Dickey-Fuller test rejected.

The majority of the samples returns present a higher maximum loss than the maximum gain. Only for the second partition, the maximum overcome in absolute value the minimum of the data set. Showing what most researchers claim, there are more abrupt drops than rises in stock market prices. The Nikkei225 daily returns also presented a small outstanding higher maximum gain above the absolute value of the a maximum loss. The higher loss of the entire 21-years period is equivalent to the higher loss of the third partition, meaning that the Covid-19

Table 4.4.: DAX30 preliminary statistics

	Daily	Weekly	Monthly	P1	P2	P3	P4
# of observations	5327	1097	252	646	357	24	4297
mean	0.013	0.064	0.268	-0.168	-0.216	-1.867	0.065
std. dev.	1.493	3.262	6.129	1.887	2.209	3.683	1.297
skewness	-0.171	-0.857	-0.895	-0.118	0.495	-1.241	-0.125
kurtosis	8.735	9.385	5.890	4.675	8.269	4.807	6.035
excess kurtosis	5.735	6.385	2.890	1.675	5.269	1.807	3.035
minimum	-13.05	-24.35	-29.33	-8.875	-7.434	-13.05	-7.067
maximum	10.798	14.942	19.374	7.553	10.798	3.633	7.086
5th percentile	-2.398	-5.028	-10.04	-3.124	-4.205	-9.710	-2.05
95th percentile	2.253	4.662	8.886	2.855	2.751	2.648	2.092
Dickey-Fuller	2728	583.8	111.12	331.61	196.73	16.048	51.597
Jarque-Bera	7326	1998	121.35	77.013	427.60	9.424	1660.5

(1) The Augmented-Dickey-Fuller and Jarque-Bera statistics are present.

(2) Unless otherwise noted, all statistics reject the null hypothesis at one percent level of significance.

(3) Statistics whose value is higher than the critical values, implying the non rejection of the null hypothesis, are in bold.

Table 4.5.: NIKKEI225 preliminary statistics

	Daily	Weekly	Monthly	P1	P2	P3	P4
# of observations	5145	1097	252	630	348	22	4147
mean	0.007	0.036	0.138	-0.133	-0.256	-1.442	0.057
std. dev.	1.492	3.076	5.653	1.689	2.668	2.227	1.297
skewness	-0.387	-1.008	-0.721	0.100	-0.252	-0.483	-0.367
kurtosis	9.377	11.06	4.446	4.307	7.251	2.379	7.159
excess kurtosis	6.377	8.061	1.446	1.307	4.251	-0.622	4.159
minimum	-12.11	-27.88	-27.22	-7.234	-12.11	-6.274	-11.153
maximum	13.235	15.817	14.013	7.222	13.24	2.003	7.731
5th percentile	-2.344	-5.137	-10.067	-2.644	-4.847	-5.628	-2.099
95th percentile	2.206	4.217	8.010	2.648	3.323	1.449	2.055
Dickey-Fuller	2730.2	572.32	97.111	344.52	196.43	6.275*	52.872
Jarque-Bera	8847	3156.2	43.842	45.887	261.82	1.208*	3081.72

(1) The Augmented-Dickey-Fuller and Jarque-Bera statistics are present.

(2) Unless otherwise noted, all statistics reject the null hypothesis at one percent level of significance.

(3) Statistics whose value is higher than the critical values, implying the non rejection of the null hypothesis, are in bold.

stress period caused, in fact, one of the highest losses of all time during a short period.

All the statistics presented are obtained with the logarithmic returns. The second last row gives the Dickey-Fuller statistic. For all the data sets except for the third partition of the Nikkei225 return series, these values are greater than their critical value at a 1-per cent level. This provides support to the alternative hypothesis that the series of logarithmic returns is a stationary process. This is an essential hypothesis that, as said before, reduces the chances of producing estimation errors.

The last row gives the Jarque-Bera statistic for detecting withdraws of the data from the normality. The Jarque-Bera statistic is calculated in MatLab using $JB = n(\frac{s^2}{6} + \frac{(k-3)^2}{24})$, where n is the sample size, and s and k represent the sample skewness and kurtosis, respectively. Therefore, the normality assumption is not rejected only for the third partition of the Nikkei225 and the S&P500. Although rejected, the statistics for the remaining indices are very low, almost not rejected. Since most statistics are greater than their critical value, the logarithmic changes series are non-normal.

4.2. Estimation results

We held the estimation of each distribution's parameters through the log-likelihood maximisation process. As aforementioned, the SGT distribution parameter estimation is performed, resorting to BenSaïda's MatLab code. For the parameter estimation process of the GPD, we use Matlab's Distribution Fitter app, which can also display the comparison of the empirical observations with the fitted distribution. In addition, a QQ-plot is also presented in other works of literature to analyse the fitting process. The results for the estimated parameters are presented in the aforementioned tables.

Albeit most results were successfully estimated, four estimations did not generate reliable results. Only for the parameter estimation of the S&P500 monthly SGT distribution (Table 4.6), the Maximum likelihood has converged to a boundary point of the parameter space. On the other hand, for the second period of DAX30 at 5% tail data (Table 4.7), the FTSE100 monthly returns at 5% tail data (Table 4.8) and the portfolio second subperiod at 5% tail (Table 4.9), the generalised Pareto distribution did converge to an estimate out of the boundary, leading to unreliable data. Out-of-boundary estimated parameters may not represent entirely the real data behaviour. The generalised Pareto boundary approximated results are due to the small sample size. This shows that, even if the choice of 5% tail information is better to evaluate the tail behaviour, we also need to pay attention to the amount of information present in that 5% of the empirical distribution. The 5% of the daily returns did not present any problems, suggesting that around 260 observations is data enough to estimate the generalised Pareto accurately.

The SGT showed almost null standard errors for the estimators, as included in parentheses under the estimated values presented in the tables. On the other hand, the GP estimated parameters showed some significant standard errors.

The values for k and n jibe with BenSaïda and Slim (2016) expectations, complying with the convergence of $k \to 2$ and $n \to \infty$. Remember that k and n are shape parameters that control the height and tails of the distribution. Their combination indicates a kurtosis higher than the permitted by the Normal distribution. Additionally, the skewness parameter λ suggests a possible symmetry in some cases, while in others, a high value close in absolute terms to 1 indicates a highly skewed distribution. The estimated skewness and kurtosis are calculated using equations 2.7 and 2.8.

The estimated skewness does not match the empirical skewness completely. Most of the estimated skewness values are, in absolute terms, lower than the actual empirical skewness. That is to say that the fitted distribution assumes a more symmetric distribution than the actual observed. These differences are not very significant since, as early mentioned, the empirical skewness is highly distorted due to the excess of kurtosis. Another cause for the excessive empirical skewness may be the atypical observations that exceed the standard deviation of the observed data. As mentioned earlier, all the empirical samples are leptokurtic, which overestimates the skewness. Despite the difference in the values, the estimated skewness and the empirical skewness signal are almost always the same. Therefore, the SGT is a worthwhile distribution that does not overestimate the skewness, letting to a rigorous interpretation of the actual behaviour of the market. Moreover, the Nikkei225 shows a very simillar skewness (Tables 4.10 and 4.5).

For the GP estimated parameters, most of ξ values are lower than 1/4 due to the low amount of information when choosing the lower 5% of the total data. This leads us to the impossibility to evaluate both the skewness and the kurtosis of each data series. The estimated skewness for the left tails is relatively high. This means that even considering only the tails, the analysed tail is skewed to the loss region. All the estimated parameters here presented for the GP distribution fit are evaluated into the series of loss returns; This means that all the data from the series is given with absolute values.

4.3. Goodness of fit results

The goodness-of-fit inspected numerically is also presented. Laio et al. (2009) affirms that the lower the value of the criterion, the better the fit. The goodness-of-fit is also inspected graphically. The returns are plotted as a normalised histogram along with the normal and the SGT fitted distribution. Both 5% and 20% thresholds are represented against the fitted GPD over the same point for the left tail. All tests are implemented with the null hypothesis that the data comes from the hypothesised distribution at a 1% significance level.Although not present in this dissertation, all the graphic representations are annexed.

For the Nikkei 225 return series, the SGT was utterly rejected by all tests in the weekly returns on the Table 4.11. In a graphic interpretation, it is also possible to check the non-relation of the estimated SGT with the empirical distribution. The estimated Normal distribution is also plotted, and again it does not describes the empirical distribution. In Figure 4.1(a),

Table 4.6.: SP500 estimated parameters

	Daily	Weekly	Monthly	P1	P2	P3	P4
SGT							
k	1.272	1.678	1.066	1.741	0.914	1.142	1.159
	(0.000)	(0.000)	(0.016)	(0.005)	(0.000)	(0.000)	(0.000)
n	4.416	4.000	200.	9.878	91.218	1.2E+31	9.138
	(0.000)	(0.000)	(0.000)	(0.04)	(0.000)	(0.000)	(0.000)
λ	-0.055	-0.001	0.221	0.023	-0.018	0.17	-0.02
	(0.007)	(0.000)	(0.012)	(0.380)	(0.001)	(121.5)	(0.005)
μ	0.019	0.237	-0.38	-0.11	-0.03	-1.74	0.06
	(0.005)	(0.000)	(0.003)	(0.079)	(0.000)	(5.932)	(0.001)
σ	1.26	2.529	4.497	1.444	2.383	5.099	0.986
	(0.000)	(0.000)	(0.001)	(0.006)	(0.000)	(2.022)	(0.000)
skewness	-0.531	-0.011	0.812	0.067	-0.089	0.579	-0.122
kurtosis	43.057	4.2E+7	5.946	4.62	7.213	5.260	8.438
GPD - 5%							
ξ	0.194	0.307	-0.213	0.244	-0.263	6.478	-0.027
	(-0.139)	(-0.36)	(-0.65)	(-0.53)	(-0.76)	(-9.79)	(-0.12)
σ	0.945	1.671	1.808	0.546	2.354	0.000	0.849
	(0.593)	(0.762)	(0.992)	(0.047)	(1.235)	(-6.48)	(0.736)
μ	1.933	4.107	6.854	2.299	4.113	8.287	1.582
skewness	4.467	20.627	1.147	6.638	1.019	(a)	1.849
kurtosis	63.988	(b)	4.056	792.41	3.590	(b)	7.857
GPD - 20%							
ξ	0.134	0.160	-0.213	-0.076	-0.025	-1.141	-0.012
	(-0.07)	(-0.15)	(-0.30)	(-0.14)	(-0.26)	(0.000)	(-0.06)
σ	0.838	1.645	2.806	0.910	1.940	8.851	0.795
	(0.633)	(1.191)	(2.085)	(0.806)	(1.394)	(9.992)	(0.737)
μ	0.686	1.501	3.122	1.134	1.367	3.303	0.559
skewness	3.239	3.688	1.148	1.616	1.857	-0.116	1.9298
kurtosis	24.98	35.218	4.059	6.339	7.918	1.786	8.453

(1) All the estimates are significant at the 1% level.

(2) The standard errors for the estimators are included in parentheses.

skewness and kurtosis measures are calculated using equations (2.3) and (2.4) for the SGT distribution and

(a) skewness for the generalised Pareto distribution is only evaluated for $\xi < 1/3$.

(b) kurtosis for the generalised Pareto distribution is only evaluated for $\xi < 1/4$

Table 4.7.: DAX30 estimated parameters

	Daily	Weekly	Monthly	P1	P2	P3	P4
SGT							
k	1.249	2.105	2.218	2.289	1.372	2.268	1.117
	(0.016)	(0.000)	(0.000)	(0.001)	(0.161)	(0.000)	(0.032)
n	7.191	4.000	4.000	5.07	2.000	4.000	18.265
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)
λ	-0.068	-0.007	-0.001	-0.014	-0.100	0.002	-0.052
	(0.043)	(0.000)	(0.002)	(0.119)	(2.223)	(0.035)	(0.033)
μ	0.011	0.241	0.751	-0.155	-0.261	-1.376	0.064
	(0.000)	(0.000)	(0.000)	(0.016)	(0.424)	(0.000)	(0.027)
σ	1.492	3.257	6.146	1.900	2.278	3.618	1.298
	(0.01)	(0.000)	(0.000)	(0.003)	(0.049)	(0.000)	(0.003)
skewness	-0.376	-0.045	-0.005	-0.052	-1.080	0.010	-0.229
kurtosis	9.460	2.8E+7	2.6E+7	7.304	7.1E+7	2.5E+7	6.543
GPD - 5%							
ξ	0.071	0.283	-0.156	-0.131	-1.096	6.619	-0.150
	(-0.13)	(-0.35)	(-0.61)	(-0.30)	(0.000)	(-9.98)	(-0.11)
σ	1.139	2.067	6.634	1.554	3.468	0.000	1.208
	(0.882)	(1.060)	(3.098)	(1.119)	(4.563)	(-6.62)	(1.165)
μ	2.410	5.105	10.317	2.953	4.053	6.410	2.083
skewness	2.518	11.133	1.318	1.402	-0.080	(a)	1.336
kurtosis	14.043	(b)	4.780	5.178	1.786	(b)	4.864
GPD - 20%							
ξ	0.061	0.134	0.068	0.080	0.098	-0.064	-0.075
	(-0.06)	(-0.14)	(-0.33)	(-0.21)	(-0.37)	(-1.62)	(-0.06)
σ	1.040	1.997	4.588	1.080	1.446	3.442	1.104
	(0.894)	(1.522)	(2.923)	(0.744)	(0.835)	(0.608)	(1.084)
μ	0.929	2.234	3.958	1.396	1.261	2.223	0.766
skewness	2.432	3.237	2.493	2.601	2.785	1.668	1.62
kurtosis	13.062	24.936	13.745	15.034	17.463	6.653	6.362

(1) All the estimates are significant at the 1% level.

(2) The standard errors for the estimators are included in parentheses.

skewness and kurtosis measures are calculated using equations (2.3) and (2.4) for the SGT distribution and

(a) skewness for the generalised Pareto distribution is only evaluated for $\xi < 1/3$.

(b) kurtosis for the generalised Pareto distribution is only evaluated for $\xi < 1/4$

Table 4.8.: FTSE100 estimated parameters

	Daily	Weekly	Monthly	P1	P2	P3	P4
SGT							
k	1.60	2.036	1.379	1.409	1.699	12.492	1.480
	(0.000)	(0.000)	(0.001)	(0.200)	(0.000)	(0.000)	(0.035)
n	2.206	2.000	118.74	15.753	3.833	4.000	6.357
	(0.000)	()	(0.000)	(0.009)	(0.000)	(0.010)	(0.004)
λ	-0.056	-0.132	-0.351	-0.059	-0.028	-0.003	-0.040
	(0.000)	(0.00)	(0.000)	(0.356)	(20.66)	(0.043)	(0.000)
μ	-0.001	-0.005	-0.001	-0.087	-0.177	-1.161	0.031
	(0.000)	(0.030)	(0.028)	(0.120)	(0.053)	(0.000)	(0.000)
σ	1.219	2.394	4.131	1.408	2.211	2.556	1.002
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.205)
skewness	-0.508	-0.903	-0.882	-0.195	-0.240	-0.008	-0.193
kurtosis	4.7E+7	3.3E+7	4.737	4.997	4.1E+7	3.2E+7	8.328
GPD - 5%							
ξ	0.1545	0.3612	-0.5658	-0.3795	-0.3743	1.9972	-0.0244
	(-0.140)	(-0.377)	(-0.788)	(-0.337)	(0.4629)	(-4.237)	(-0.131)
σ	0.893	1.554	4.629	1.478	2.716	0.212	0.790
	(0.588)	(0.330)	(2.436)	(1.292)	(1.820)	(-1.99)	(0.680)
μ	1.885	3.766	7.512	2.295	3.321	6.310	1.594
skewness	3.578	(a)	0.470	0.770	0.780	(a)	1.862
kurtosis	32.401	(b)	2.234	2.851	2.876	(b)	7.950
GPD - 20%							
ξ	0.121	0.240	-0.147	-0.040	0.013	2.202	-0.003
	(-0.066)	(-0.156)	(-0.318)	(-0.196)	(-0.269)	(-4.261)	(-0.065)
σ	0.784	1.303	3.690	1.037	1.605	0.193	0.738
	(0.596)	(0.825)	(2.577)	(0.838)	(1.112)	(-2.20)	(0.675)
μ	0.706	1.597	2.902	0.968	1.251	2.238	0.621
skewness	3.068	6.364	1.346	1.782	2.083	(a)	1.980
kurtosis	21.864	448.97	4.910	7.394	9.686	(b)	8.844

(1) All the estimates are significant at the 1% level.

(2) The standard errors for the estimators are included in parentheses.

skewness and kurtosis measures are calculated using equations (2.3) and (2.4) for the SGT distribution and

(a) skewness for the generalised Pareto distribution is only evaluated for $\xi < 1/3$.

(b) kurtosis for the generalised Pareto distribution is only evaluated for $\xi < 1/4$

Table 4.9.: Portfolio estimated parameters

	Daily	Weekly	Monthly	P1	P2	P3	P4
SGT							
k	1.458	2.020	1.871	1.295	1.736	2.964	1.328
	(0.000)	(0.000)	(0.000)	(0.082)	(0.000)	(53.37)	(0.004)
n	4.980	4.000	2.000	129.2	3.999	2.059	8.697
	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)	(0.000)	(0.000)
λ	-0.063	-0.047	-0.334	0.022	0.001	-0.74	-0.053
	(0.000)	(0.000)	(0.000)	(0.146)	(0.006)	(642.0)	(0.015)
μ	0.005	0.214	0.001	-0.123	-0.208	-1.989	0.049
	(0.002)	(0.000)	(0.020)	(0.013)	(0.000)	(159.0)	(0.033)
σ	1.266	2.726	5.148	1.477	2.137	3.993	1.074
	(0.000)	(0.000)	(0.000)	(0.004)	(0.000)	(118.9)	(0.002)
skewness	-0.412	-0.325	-2.298	0.068	0.009	-2.838	-0.239
kurtosis	15.434	3.0E+7	5.9E+7	4.466	3.9E+7	7.2E+7	7.088
GPD - 5%							
ξ	0.099	0.235	0.473	-0.491	-1.115	6.207	-0.074
	(-0.13)	(-0.31)	(-1.24)	(-0.40)	(0.000)	(-9.42)	(-0.11)
σ	1.016	2.069	1.467	1.470	3.629	0.000	0.952
	(0.752)	(1.155)	(-0.09)	(1.388)	(4.744)	(-6.21)	(0.878)
μ	1.969	4.132	9.718	2.490	3.734	6.627	1.695
skewness	2.806	6.071	(a)	0.579	-0.095	(a)	1.622
kurtosis	17.751	291.033	(b)	2.425	1.786	(b)	6.376
GPD - 20%							
ξ	0.1100	0.1886	-0.2356	-0.0656	0.0875	-1.1686	-0.246
	(-0.07)	(0.1408)	(-0.27)	(-0.20)	(-0.34)	(65535)	(-0.06)
σ	0.8305	1.5286	5.2521	1.0713	1.3709	9.7477	0.8403
	(0.650)	(1.074)	(3.842)	(0.887)	(0.825)	(65535)	(0.792)
μ	0.771	1.953	2.849	1.014	1.269	2.132	0.655
skewness	2.926	4.321	1.086	1.661	2.678	-0.137	1.8611
kurtosis	19.549	57.117	3.828	6.610	16.015	1.787	7.945

(1) All the estimates are significant at the 1% level.

(2) The standard errors for the estimators are included in parentheses.

skewness and kurtosis measures are calculated using equations (2.3) and (2.4) for the SGT distribution and

(a) skewness for the generalised Pareto distribution is only evaluated for $\xi < 1/3$.

(b) kurtosis for the generalised Pareto distribution is only evaluated for $\xi < 1/4$

Table 4.10.: NIKKEI225 estimated parameters

	Daily	Weekly	Monthly	P1	P2	P3	P4
SGT							
k	1.572	2.653	1.920	2.035	1.720	2.880	1.464
	(0.000)	(0.000)	(0.000)	(4E+6)	(0.000)	(8641)	(0.000)
n	5.512	2.088	18.583	7.825	3.999	2.013	8.468
	(0.000)	(0.000)	(0.000)	(2E+6)	(0.000)	(0.000)	(0.000)
λ	-0.044	-0.175	-0.315	0.081	0.002	-0.594	-0.031
	(0.001)	(0.058)	(0.000)	(2E+6)	(0.003)	(0.000)	(0.007)
μ	0.013	0.040	0.146	-0.128	-0.151	-1.492	0.060
	(0.002)	(0.254)	(0.000)	(2E+6)	(0.000)	(0.000)	(0.001)
σ	1.481	3.051	5.621	1.688	2.698	2.136	1.291
	(0.000)	(0.000)	(0.000)	(2E+6)	(0.000)	(0.000)	(0.000)
skewness	-0.228	-0.929	-0.617	0.221	0.018	-0.488	-0.123
kurtosis	9.879	2.5E+7	3.848	4.543	4.0E+7	2.673	6.253
GPD - 5%							
ξ	0.264	0.418	1.296	0.107	-0.045	5.65	0.127
	(-0.17)	(-0.38)	(-1.74)	(-0.41)	(-0.72)	(-8.66)	(-0.14)
σ	0.907	1.329	0.465	0.895	2.238	0.000	0.835
	(0.475)	(0.433)	(-1.22)	(0.418)	(0.997)	(-5.65)	(0.559)
μ	2.3502	5.186	10.259	2.511	4.559	3.756	2.119
skewness	8.360	(a)	(a)	2.892	1.76	(a)	3.149
kurtosis	(b)	(b)	(b)	19.013	7.242	(b)	23.287
GPD - 20%							
ξ	0.094	0.121	-0.075	0.107	0.284	-1.117	0.007
	(-0.06)	(-0.13)	(-0.18)	(-0.19)	(-0.40)	(0.00)	(-0.06)
σ	0.972	1.874	4.377	0.756	1.420	3.212	0.968
	(0.798)	(1.439)	(3.238)	(0.477)	(0.610)	(4.329)	(0.880)
μ	0.958	2.153	4.582	1.419	1.690	1.956	0.816
skewness	2.748	3.061	1.620	2.886	11.399	-0.097	2.040
kurtosis	16.95	21.733	6.366	18.939	(b)	1.789	9.3255

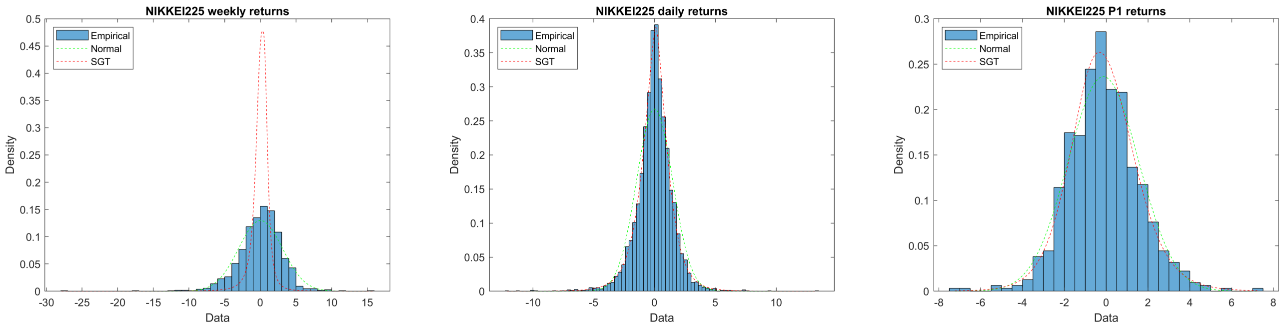
(1) All the estimates are significant at the 1% level.

(2) The standard errors for the estimators are included in parentheses.

skewness and kurtosis measures are calculated using equations (2.3) and (2.4) for the SGT distribution and

(a) skewness for the generalised Pareto distribution is only evaluated for $\xi < 1/3$.

(b) kurtosis for the generalised Pareto distribution is only evaluated for $\xi < 1/4$



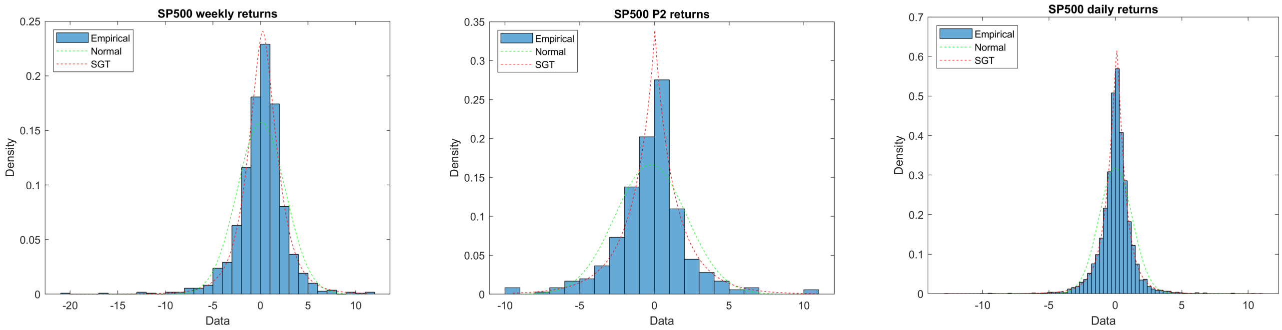
(a) (b) (c)

Figure 4.1.: Nikkei225 fitted distributions for (a) weekly returns (b) daily returns (c) p1 returns

The left skewness is visible, and the pike of information around the mean exceeds the normal distribution considerably. Therefore, neither the SGT nor the Normal distributions are good enough to represent the behaviour. On the other hand, the SGT distribution accommodates remarkably well the allocations for the remaining Nikkei 225 return series. As we can see in Figures 4.1(b) and 4.1(c), the SGT is an excellent adaptation for the real world returns behaviour. For the first partition of Nikkei returns, the Normal distribution provides a good monitor for the empirical distribution along with the SGT, but again, the SGT approximation is favoured. The skewness of the observed distribution is visible in the graphic representation. In contrast, the normal distribution may follow a tail but does not follow the other due to the skewness, while the SGT follows both tails while supporting the peak around the mean. In other words, the Normal distribution can't keep track of the curvature present in the tails.

The index that shows the best results is the S&P500, whose results are presented in Table 4.12. The KS and the Kuiper test do not reject the SGT for all samples, while the AD test rejected the weekly and the second stress subperiod. The AD test is not entirely accurate for the SGT, since it focuses on the convergence for the normal distribution to evaluate the critical values. As graphically visible, the weekly returns (Figure 4.2(a)), rejected by the AD test, accommodate very well the SGT distribution. In the meantime, the Normal distribution is rejected. It is graphically explicit that the peak is much higher than the Normal distribution peak, and the left tail presents an extreme skew. Environmental, economic, and social events and phenomenon provide financial drains on our economy. Events such as the Dot.com Bubble, Terrorist Attacks, Stock Market Crash, War, Hurricanes, and others. The second partition (Figure 4.2(c)), also rejected by the AD test, slightly repudiate the SGT distribution due to its higher peak. Despite this, the SGT preserves the tail behaviour exceptionally well. The Sub-Prime Housing Crisis and the Housing bubble were the main reasons for the excess of tail data and a lower peak around the mean. The daily returns (Figure 4.2(b)) and the fourth partition have high empirical peaks, and both welcome the SGT distribution very well.

The FTSE100 has the most rejected fit results (Table **??**). Both the daily and the weekly



(a) (b) (c)

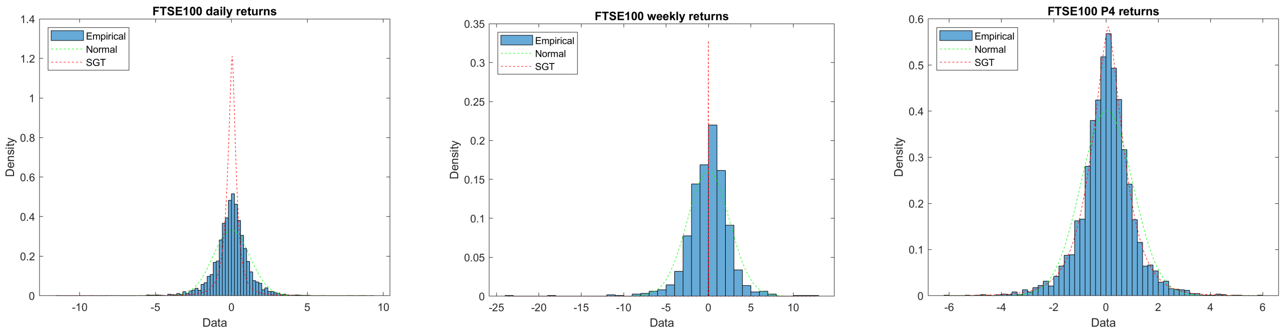
Figure 4.2.: SP500 fitted distributions for (a) weekly returns (b) daily returns (c) p1 returns

returns (Figures 4.3(a) and 4.3(b)) do not follow an SGT distribution nor a Normal distribution. Despite the exaggeration made by the SGT for peak information, it still shows arch tails, i.e. it can accompany the skewness but not entirely the data. On the other hand, the SGT distribution follows the empirical data for the fourth partition (Figure 4.3(c)) extremely well, which has a higher peak. The GP distribution holds the tail behaviour for the daily and weekly returns for both the 5% and 20% tail information (Figures 4.4(a), 4.4(b) and 4.4(c)). All the analytical tests are non-rejected, and its graphical representation shows the well-fitting. For the FTSE100 index, Theodossiou's distribution could not correspond to the real behaviour. This can be due to the specific distribution and having nothing to do with the economic and financial reason for conducting the empirical data.

The DAX30 returns follow the same analysis previously made, as revealed by Figure 4.5(c). The good-of-fit results are presented in Table 4.14. Albeit the skewness presented, the DAX30 daily returns form a balanced distribution. Analysing the timeline of gains and losses, for example, after the 2011 European dept crisis related to Spain and Italy, which implicated a remarkable drop in the DAX30 price, the gains were substantially high. The same happened after the 2020 coronavirus lock-downs. For the Portfolio (Table 4.15), the third subperiod 4.5(a) was analytically rejected in all tests for the SGT, while the Normal distribution was not dismissed. Again, the amount of data is a significant feature for fitting distributions. All the other distributions followed the SGT exceptionally well. The problem with the covid subperiod is due to the lack of information. Results like the ones obtained for the Portfolio daily returns (Figure 4.5(b)) are preferable.

The remaining graphic representations are annexed. We recommend a close look at the GP fit for the tails. The principal difficulty for this distribution is the amount of data, and it is challenging to validate the fit for each subperiod and monthly returns.

Most of the stress partitions present good fit results. Even with the small amount of data, principally in the covid-19 bear market partition, the SGT produced an acceptable fit. It is in this same partition that the Normal distribution was accepted for all the return series. The

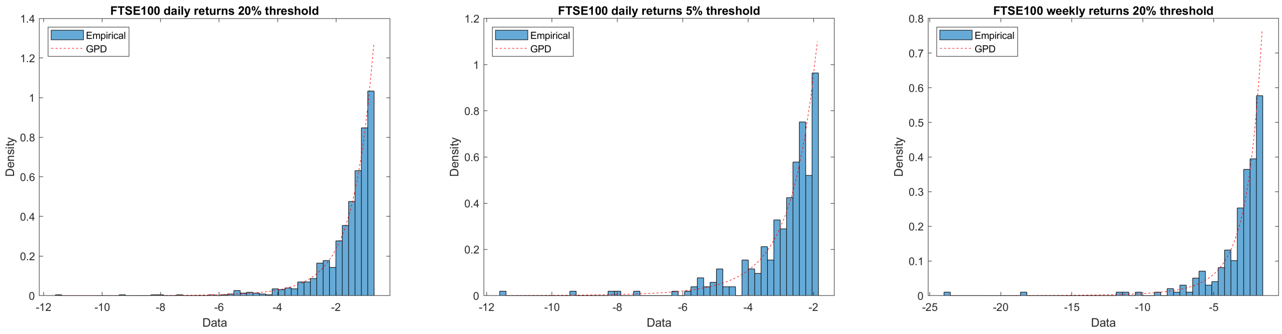


(a) (b) (c)

Figure 4.3.: FTSE100 fitted distributions for (a) daily returns (b) weekly returns (c) p4 returns

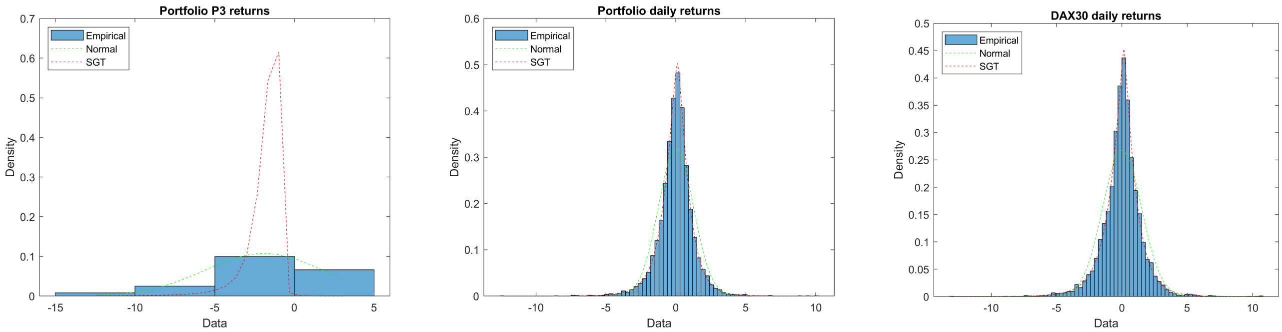
fourth partition, the bigger and without stress, was the partition with the better estimation. Therefore, for both stress and non-stress market periods, the SGT establishes a good empirical distribution adjust. It should also be noted that, even for the non-stress partition, unavoidable events exist, such as social, environmental and economic crises. Taking into attention the DAX30 case, the crisis in European countries such as Italy and Spain had a significant implication on its price. In the same way, we could observe unusual gains, which compensated for the tremendous losses.

The main skewness patterns were most observed in weekly returns. Monthly returns also showed a negative asymmetry, exhibiting the skewness pattern in high frequency series. This way, we do not support the investment on high frequency such as weekly and monthly return series, moreover we alert to the use of the Normal distribution, since it does not reflect the real behaviour of the empirical distribution. It should also be noted that long-duration periods, such as our 21-years period, are not correctly performed by a Normal distribution, being the SGT distribution an advantage. This is due to the disregard of the skewness and kurtosis parameters by the Normal distribution. On the other hand, in the tails' analysis, the GP distribution is a pretty good fit. It can follow the skewness patters and corresponds to the real behaviour of the tails of a non centralized distribution.



(a) (b) (c)

Figure 4.4.: FTSE100 GPD fitted distributions for (a) daily returns 20% (b) daily returns 5% (c) weekly returns 20%



(a) (b) (c)

Figure 4.5.: Fitted distributions for (a) Portfolio P3 returns (b) Portfolio daily returns (c) DAX daily returns

Table 4.11.: NIKKEI225 goodness-of-fit

	Daily	Weekly	Monthly	P1	P2	P3	P4
SGT							
KS test	0.009*	0.254	0.042*	0.02*	0.024*	0.084*	0.010*
	(0.023)	(0.049)	(0.102)	(0.065)	(0.087)	(0.337)	(0.025)
AD test	0.62*	289.32	0.450*	0.178*	0.322*	0.143*	0.614*
	(0.752)	(0.751)	(0.75)	(0.751)	(0.750)	(0.724)	(0.752)
Kuiper test	0.016*	0.455	0.077*	0.035*	0.047*	0.148*	0.016*
	(0.028)	(0.060)	(0.125)	(0.079)	(0.079)	(0.407)	(0.031)
GPD - 5%							
KS test	0.043*	0.051*	0.306*	0.079*	0.144*	0.331*	0.044*
	(0.101)	(0.216)	(0.433)	(0.281)	(0.371)	(0.929)	(0.112)
AD test	0.443*	0.224*	0.885*	0.524*	0.819*	0.314*	0.784*
	(3.880)	(3.888)	(3.910)	(3.887)	(3.901)	(4.07)	(3.880)
Kuiper test	0.071*	0.101*	0.425*	0.148*	0.272*	0.641*	0.088*
	(0.123)	(0.263)	$(+\infty)$	(0.341)	(0.448)	(0.995)	(0.137)
GPD - 20%							
KS test	0.03*	0.041*	0.150*	0.046*	0.12*	0.293*	0.046*
	(0.051)	(0.109)	(0.224)	(0.144)	(0.193)	(0.669)	(0.056)
AD test	1.301*	0.383*	1.618*	0.357*	1.204*	7.8319	3.818*
	(3.879)	(3.880)	(3.89)	(3.882)	(3.884)	(3.959)	(3.879)
Kuiper test	0.056*	0.072*	0.254*	0.076*	0.180*	0.493*	0.067*
	(0.062)	(0.273)	(0.273)	(0.176)	(0.236)	(0.789)	(0.069)
Normal							
KS test	0.0671	0.0506	0.0857*	0.0344*	0.0859*	0.1129*	0.0587
	(0.0227)	(0.0490)	(0.1018)	(0.0640)	(0.0874)	(0.3367)	(0.0252)
AD test	65535	6.503	1.876*	1.144*	4.741	0.29*	26.066
	(3.878)	(3.879)	(3.880)	(3.879)	(3.88)	(3.897)	(3.878)
Kuiper test	0.116	0.095	0.128	0.060*	0.158	0.196*	0.102
	(0.028)	(0.060)	(0.125)	(0.079)	(0.107)	(0.407)	(0.031)

(1) All the tests were executed at a 5% level of significance.

(2) Critical values at 5% for each test are given by the last row in parenthesis.

* corresponds to the non rejection of the null hypothesis that the empirical distribution follows the estimated distribution.

Table 4.12.: SP500 goodness-of-fit

	Daily	Weekly	Monthly	P1	P2	P3	P4
SGT							
KS test	0.0099*	0.0278*	0.0302*	0.0230*	0.0641*	0.0965*	0.0118*
	(0.0223)	(0.0490)	(0.1020)	(0.0642)	(0.0856)	(0.3229)	(0.0249)
AD test	0.578*	1.312	0.245*	0.316*	1.603	0.222*	0.706*
	(0.752)	(0.751)	(0.75)	(0.751)	(0.750)	(0.726)	(0.752)
Kuiper test	0.016*	0.039*	0.055*	0.040*	0.068*	0.166*	0.018*
	(0.028)	(0.060)	(0.125)	(0.079)	(0.079)	(0.389)	(0.031)
GPD - 5%							
KS test	0.036*	0.067*	0.123*	0.143*	0.170*	0.334*	0.037*
	(0.099)	(0.216)	(0.433)	(0.281)	(0.371)	(0.929)	(0.110)
AD test	0.484*	0.299*	0.613*	0.536*	0.783*	0.322*	0.459*
	(3.88)	(3.888)	(3.91)	(3.887)	(3.901)	(4.074)	(3.880)
Kuiper test	0.073*	0.128*	0.219*	0.211*	0.286*	0.65*	0.058*
	(0.122)	(0.263)	$(+\infty)$	(0.341)	(0.448)	(0.995)	(0.135)
GPD - 20%							
KS test	0.0185*	0.0396*	0.1141*	0.0874*	0.0593*	0.3641*	0.0513*
	(0.0499)	(0.1089)	(0.2239)	(0.1424)	(0.1890)	(0.6685)	(0.0555)
AD test	0.377*	0.401*	1.462*	1.336*	0.343*	8.612	3.488*
	(3.879)	(3.880)	(3.89)	(3.882)	(3.884)	(3.959)	(3.879)
Kuiper test	0.032*	0.067*	0.169*	0.144*	0.10*	0.564*	0.068*
	(0.061)	(0.273)	(0.273)	(0.174)	(0.231)	(0.789)	(0.068)
Normal							
KS test	0.099	0.0844	0.099*	0.042*	0.1018	0.124*	0.0830
	(0.022)	(0.049)	(0.102)	(0.064)	(0.086)	(0.323)	(0.025)
AD test	65535	16.8535	3.031*	1.706*	6.145	0.341*	55.27
	(3.878)	(3.879)	(3.880)	(3.879)	(3.88)	(3.896)	(3.878)
Kuiper test	0.185	0.162	0.159	0.087	0.19	0.234*	0.146
	(0.027)	(0.060)	(0.125)	(0.079)	(0.105)	(0.3895)	(0.031)

(1) All the tests were executed at a 5% level of significance.

(2) Critical values at 5% for each test are given by the last row in parenthesis.

* corresponds to the non rejection of the null hypothesis that the empirical distribution follows the estimated distribution.

Table 4.14.: DAX30 goodness-of-fit

	Daily	Weekly	Monthly	P1	P2	P3	P4
SGT							
KS test	0.006*	0.028*	0.049*	0.016*	0.521	0.080*	0.009*
	(0.022)	(0.049)	(0.102)	(0.064)	(0.086)	(0.323)	(0.025)
AD test	0.218*	1.862	0.801	0.191*	1916.816	0.312*	0.235*
	(0.751)	(0.751)	(0.750)	(0.751)	(0.750)	(0.726)	(0.752)
Kuiper test	0.012*	0.054*	0.092	0.031*	0.971	0.147*	0.016*
	(0.027)	(0.060)	(0.060)	(0.078)	(0.078)	(0.390)	(0.030)
GPD - 5%							
KS test	0.035*	0.060*	0.170*	0.079*	0.297*	0.334*	0.057*
	(0.099)	(0.216)	(0.433)	(0.277)	(0.371)	(0.929)	(0.110)
AD test	0.466*	0.338*	0.972*	0.398*	3.665*	0.323*	0.951*
	(3.880)	(3.888)	(3.910)	(3.887)	(3.901)	(4.074)	(3.880)
Kuiper test	0.065*	0.119*	0.298*	0.137*	0.397*	0.651*	0.096*
	(0.121)	(0.263)	$(+\infty)$	(0.336)	(0.448)	(0.995)	(0.135)
GPD - 20%							
KS test	0.026*	0.042*	0.082*	0.059*	0.091+	0.245*	0.045*
	(0.050)	(0.109)	(0.224)	(0.141)	(0.189)	(0.669)	(0.055)
AD test	0.661*	0.448*	0.283*	0.418*	0.941*	1.581*	3.495*
	(3.879)	(3.880)	(3.890)	(3.882)	(3.884)	(3.959)	(3.879)
Kuiper test	0.045*	0.076*	0.131*	0.090*	0.170*	0.388*	0.074
	(0.061)	(0.273)	(0.273)	(0.173)	(0.231)	(0.789)	(0.068)
Normal							
KS test	0.077	0.059	0.074*	0.043*	0.098	0.139*	0.074
	(0.022)	(0.049)	(0.102)	(0.064)	(0.086)	(0.323)	(0.025)
AD test	67.315	9.940	2.767*	1.658*	7.883	0.602*	43.38
	(3.878)	(3.879)	(3.880)	(3.879)	(3.880)	(3.896)	(3.878)
Kuiper test	0.142	0.115	0.143	0.074*	0.188	0.230*	0.132
	(0.027)	(0.060)	(0.125)	(0.078)	(0.105)	(0.39)	(0.031)

(1) All the tests were executed at a 5% level of significance.

(2) Critical values at 5% for each test are given by the last row in parenthesis.

* corresponds to the non rejection of the null hypothesis that the empirical distribution follows the estimated distribution.

Table 4.15.: Portfolio goodness-of-fit

		Daily	Weekly	Monthly	P1	P2	P3	P4
SGT								
KS test		0.006*	0.038*	0.567	0.019*	0.026*	0.452	0.007*
		(0.022)	(0.049)	(0.102)	(0.064)	(0.086)	(0.323)	(0.025)
AD test		0.101*	2.644	1666.5	0.267*	0.345*	29.711	0.139*
		(0.752)	(0.751)	(0.75)	(0.751)	(0.750)	(0.726)	(0.752)
Kuiper test		0.010*	0.048*	0.991	0.032*	0.050*	0.7271	0.013*
		(0.027)	(0.060)	(0.125)	(0.078)	(0.078)	(0.39)	(0.031)
GPD - 5%								
KS test		0.031*	0.074*	0.176*	0.134*	0.259*	0.334*	0.038*
		(0.099)	(0.216)	(0.433)	(0.277)	(0.371)	(0.929)	(0.110)
AD test		0.173*	0.375*	0.868*	0.865*	4.381	0.319*	0.444*
		(3.88)	(3.888)	(3.91)	(3.887)	(3.901)	(4.074)	(3.8803)
Kuiper test		0.056*	0.13*	0.344*	0.217*	0.426*	0.647*	0.057*
		(0.121)	(0.263)	$(+\infty)$	(0.336)	(0.448)	(0.995)	(0.135)
GPD - 20%								
KS test		0.0241*	0.0531*	0.0849*	0.0638*	0.1197*	0.2926*	0.0487*
		(0.0497)	(0.1089)	(0.2239)	(0.1413)	(0.1930)	(0.6685)	(0.0553)
AD test		0.44*	0.766*	0.617*	0.564*	1.204*	7.832	3.280*
		(3.879)	(3.880)	(3.89)	(3.882)	(3.884)	(3.959)	(3.879)
Kuiper test		0.042*	0.097*	0.168*	0.102*	0.180*	0.493*	0.078
		(0.061)	(0.273)	(0.273)	(0.173)	(0.236)	(0.789)	(0.068)
Normal								
KS test		0.067	0.051	0.086*	0.049*	0.085*	0.135*	0.061
		(0.023)	(0.049)	(0.102)	(0.064)	(0.087)	(0.323)	(0.025)
AD test		68.508	6.503	1.876*	2.472*	4.957	0.552*	37.665
		(3.878)	(3.879)	(3.880)	(3.879)	(3.88)	(3.896)	(3.878)
Kuiper test		0.116	0.095	0.128	0.087	0.16	0.229*	0.117
		(0.028)	(0.060)	(0.125)	(0.078)	(0.105)	(0.39)	(0.031)

(1) All the tests were executed at a 5% level of significance.

(2) Critical values at 5% for each test are given by the last row in parenthesis.

* corresponds to the non rejection of the null hypothesis that the empirical distribution follows the estimated distribution.

5. Conclusions

The Skewed Generalised t is a very compelling distribution as it nests several other distributions. For almost all cases, the Skewed Generalised t distribution provided an excellent fit to the empirical distribution of the data. Evidence from our sample proves that such model-based parametric distribution is appropriately applied. Furthermore, analytical and graphical analysis strongly favours the SGT distribution as a good fit for financial index returns.

While we expected long-horizon returns to be found normally distributed, our most extended horizon of twenty-one years did not verify the attributes, nor did the goodness-of-fit test accept the normality. In fact, the normal distribution is carried only for the third partition distribution, which has the lowest number of values. Indeed, an increased time scale does not benefit the normal distribution. This evidence jeopardises the legitimacy of the Aggregational Gaussianity.

We investigate the skewness patterns in four market indices and a hypothetical equal-weighted portfolio. We find that a negative skewness is predominant among the indices. Weekly returns are the most negative skewed, while stress-partitions are considered symmetric. Kurtosis tends to impact skewness values, although not in a direct way.

We want to draw the investors and risk managers' attention to the use of the Normal distribution. The results presented in this dissertation provide solid support for using the SGT for economic and financial models, providing a better approximation to reality. And, as mentioned before, the Normal distributions lack many possible risks incorporated in the tails. Therefore, an approximation for the Skewed Generalised t distribution on the daily returns data and its implementation, rather than the normal distribution, is highly recommended.

This dissertation provides multiple ideas for future work and research. The future research should focus on similar studies regarding individual stocks, foreign exchange rates, and alternative assets. It is also worth studying the effect of the Skewed Generalised t distribution, when working with financial models and its applications on the evaluation of the VaR. The same applies to the creation of expanded pricing models that incorporate a high flexible distribution, such as SGT. Research about how the Generalised Pareto would benefit financial models avoiding tails risk is also suggested. On top of this, restructuring the same analysis made in this work but estimating the parameters via Generalised Method of Moments (GMM) and forward goodness-of-fit test with the Log-Likelihood ratio test is a surplus-value to future research.

Bibliography

Anderson, T. W. and Darling, D. A. (1954). A test of goodness of fit. *Journal of the American statistical association*, 49(268):765–769.

Arnold, B. C. (2014). Pareto distribution. *Wiley StatsRef: Statistics Reference Online*, pages 1–10.

Ayán, M. N. R. and Díaz, M. Á. R. (2008). The reduction of skewness and kurtosis of observed variables by data transformation: Effect on factor structure. *Psicológica*, 29(2):205–227.

Bali, T. G. (2003). The generalized extreme value distribution. *Economics Letters*, 79(3):423–427.

Bali, T. G. and Theodossiou, P. (2007). A conditional-sgt-var approach with alternative garch models. *Annals of Operations Research*, 151(1):241–267.

Balkema, A. A. and De Haan, L. (1974). Residual life time at great age. *The Annals of probability*, pages 792–804.

BenSaïda, A. (2021). Kuiper test (https://www.mathworks.com/matlabcentral/fileexchange/50158-kuiper-test), matlab central file exchange. retrieved august 8, 2021.

BenSaïda, A. and Slim, S. (2016). Highly flexible distributions to fit multiple frequency financial returns. *Physica A: Statistical Mechanics and its Applications*, 442:203–213.

Box, G. E. and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, 26(2):211–243.

Chu, J., Dickin, O., and Nadarajah, S. (2019). A review of goodness of fit tests for pareto distributions. *Journal of Computational and Applied Mathematics*, 361:13–41.

Coles, S. G. and Dixon, M. J. (1999). Likelihood-based inference for extreme value models. *Extremes*, 2(1):5–23.

Corrado, C. J. and Su, T. (1996). Skewness and kurtosis in s&p 500 index returns implied by option prices. *Journal of Financial research*, 19(2):175–192.

Das, K. P. and Halder, S. C. (2016). Understanding extreme stock trading volume by generalized pareto distribution. *The North Carolina Journal of Mathematics and Statistics*, 2:45–60.

DeGroot, M. H. (2012). *Probability and statistics*. Pearson.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366a):427–431.

Dodge, Y. (2008). *The concise encyclopedia of statistics*. Springer Science & Business Media.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (2013). *Modelling extremal events: for insurance and finance*, volume 33. Springer Science & Business Media.

Gencay, R. and Selcuk, F. (2004). Extreme value theory and value-at-risk: Relative performance in emerging markets. *International Journal of forecasting*, 20(2):287–303.

Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of mathematics*, pages 423–453.

Greenwood, J. A., Landwehr, J. M., Matalas, N. C., and Wallis, J. R. (1979). Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form. *Water resources research*, 15(5):1049–1054.

Grimshaw, S. D. (1993). Computing maximum likelihood estimates for the generalized pareto distribution. *Technometrics*, 35(2):185–191.

Hair Jr, J. F., Hult, G. T. M., Ringle, C. M., and Sarstedt, M. (2021). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage publications.

Harris, R. D. and Küçüközmen, C. C. (2001). The empirical distribution of uk and us stock returns. *Journal of Business Finance & Accounting*, 28(5-6):715–740.

Hoel, P. G. et al. (1954). Introduction to mathematical statistics. *Introduction to mathematical statistics.*, (2nd Ed).

Investopedia (2021). Bear market term (https://www.investopedia.com/terms/b/bearmarket.asp) retrieved july 20, 2021.

Jenkinson, A. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological events. *Quarterly Journal of the Royal Meteorological Society*, 81:158–172.

Johansson, A. (2005). Pricing skewness and kurtosis risk on the swedish stock market.

Kucukozmen, C., Harris, R., and Yilmaz, F. (2004). Skewness in conditional distribution of daily equity returns. *Applied Financial Economics*, 14:195–202.

Kuiper, N. H. (1960). Tests concerning random points on a circle. In *Nederl. Akad. Wetensch. Proc. Ser. A*, volume 63, pages 38–47.

Laio, F., Di Baldassarre, G., and Montanari, A. (2009). Model selection techniques for the frequency analysis of hydrological extremes. *Water Resources Research*, 45(7).

Liu, H.-C., Lee, M.-C., and Chang, C.-M. (2009). The role of sgt distribution in value-at-risk estimation: evidence from the wti crude oil market. *Investment Management and Financial Innovations*, (6, Iss. 1):86–95.

MatLab (2021). Matlab website.

McDonald, J. B. and Newey, W. K. (1988). Partially adaptive estimation of regression models via the generalized t distribution. *Econometric theory*, 4(3):428–457.

Mordkoff, J. T. (2016). The assumption (s) of normality. *Dostupno na: goo. gl/g7MCwK (Pristupljeno 27.05. 2017.).*

NIST/SEMATECH (2021). e-handbook of statistical methods (https://www.itl.nist.gov/div898/handbook/eda/section3/eda35g.html), e-handbook of statistical methods. retrieved october 25, 2021.

Packham, N., Papenbrock, J., Schwendner, P., and Woebbeking, F. (2017). Tail-risk protection trading strategies. *Quantitative Finance*, 17(5):729–744.

Papastathopoulos, I. and Tawn, J. A. (2013). Extended generalised pareto models for tail estimation. *Journal of Statistical Planning and Inference*, 143(1):131–143.

Peiro, A. (1999). Skewness in financial returns. *Journal of Banking & Finance*, 23(6):847–862.

Pham, M. H., Tsokos, C., and Choi, B.-J. (2019). Maximum likelihood estimation for the generalized pareto distribution and goodness-of-fit test with censored data. *Journal of Modern Applied Statistical Methods*, 17(2):11.

Pickands III, J. (1975). Statistical inference using extreme order statistics. *the Annals of Statistics*, pages 119–131.

Sikora, G., Michalak, A., Bielak, Ł., Miśta, P., and Wyłomańska, A. (2019). Stochastic modeling of currency exchange rates with novel validation techniques. *Physica A: Statistical Mechanics and its Applications*, 523:1202–1215.

SPGlobal (2021). Bear market historic (https://www.spglobal.com/marketintelligence/en/news-insights/blog/average-bear-markets-since-1946-33-declines-that-took-two-years-to-recover) retrieved july 20, 2021.

Spiegel, M. R., Schiller, J. J., and Srinivasan, R. A. (2016). *Probabilidade e Estatística-: Coleção Schaum*. Bookman Editora.

Sprent, P. (2012). *Applied nonparametric statistical methods*. Springer Science & Business Media.

Stephens, M. (1965). The goodness-of-fit statistic v_n: Distribution and significance points. *Biometrika*, 52(3/4):309–321.

Theodossiou, P. (1998). Financial data and the skewed generalized t distribution. *Management Science*, 44(12-part-1):1650–1661.

Tolikas, K. (2008). Value-at-risk and extreme value distributions for financial returns. *The Journal of Risk*, 10(3):31–77.

Von Mises, R. (1936). La distribution de la plus grande de n valuers. *Rev. math. Union interbalcanique*, 1:141–160.

A. Moments

In this analysis we consider the definitions presented in Spiegel et al. (2016):
The r-th moment of a random variable in relation to the mean μ , also commonly called the r-th central moment, is defined as $m_r = E[(X-\mu)^r]$ for $r = 0, 1, 2, \ldots$. Overall, we have:

$$m_r = \sum_{j=1}^{n} (x_j - \mu)^r f(x_j) = \sum (x-\mu)^r f(x) \qquad \text{(discrete variable)} \tag{A.1}$$

$$m_r = \int_{-\infty}^{+\infty} (x-\mu)^r f(x)\; dx \qquad \text{(continuous variable)} \tag{A.2}$$

In each case, there are similar formulas to the ones seen before with $\mu = 0$.

The normalised n-th central moment or standardised moment is the n-th central moment divided by σ^n. Therefore, the n-th standardised moment of a variable X is

$$M_n = \frac{m_n}{\sigma^n} = \frac{E[(X-\mu)^n]}{\sigma^n} \tag{A.3}$$

Having in mind a financial interpretation where r_t represents the return of an asset at the time t.

The mean is also called the first moment and is equal to the expected value of the distribution. We define it as

$$\mu = E[r_t] \tag{A.4}$$

Similarly, the variance is the second central moment, representing the dispersion of the distribution around its mean. We define it as

$$\sigma^2 = E[r_t - \mu]^2 \tag{A.5}$$

The third and the fourth standardized moments represent the skewness and kurtosis of the alleged distribution. We define the skewness and kurtosis, respectively, according to the formulas:

$$S_k = \frac{E[(X-\mu)^3]}{\sigma^3} = \frac{m_3}{\sigma^3} \tag{A.6}$$

$$K_u = \frac{E[(X-\mu)^4]}{\sigma^4} = \frac{m_4}{\sigma^4} \tag{A.7}$$

It is common to compare both values to zero and three, the respective values for a normal distribution. It is also common to refer to the excess kurtosis instead of the kurtosis since the comparison to the normal distribution is done by subtracting three to the kurtosis value. We can characterize a distribution by its skewness and kurtosis values.

If a distribution has a kurtosis greater than 3, we call it leptokurtic. If the kurtosis is equal to 3, the same as the normal distribution, we say the distribution is mesokurtic. And it is called platykurtic when the kurtosis is lower than 3.

For the SGT distribution, given the fact that the probability density function of the transformed variable $z = x - \mu$, where μ is the mean, is given by equation 2.1. The fourth theorem of Theodossiou (1998) shows that

$$E(|z|) = E(|x|) - \lambda\mu + 2\int_{\mu}^{0} x f_1 dx - 2\mu \int_{\mu}^{0} f_1 dx \qquad \text{for } \lambda < 0$$
$$E(|z|) = E(|x|) - \lambda\mu + 2\int_{\mu}^{0} x f_2 dx - 2\mu \int_{\mu}^{0} f_2 dx \qquad \text{for } \lambda > 0$$

where $E(|x|)$ is the expected value of $|x|$ given by

$$E(|x|) = (1+\lambda^2)\frac{\sigma}{S(\lambda)} \frac{B(\frac{2}{k}, \frac{n-1}{k})}{B(\frac{1}{k}, \frac{n}{k})^{1/2} B(\frac{3}{k}, \frac{n-2}{k}^{-1/2})}$$

for $S(\lambda) = \left(1 + 3\lambda^2 - 4\lambda^2 B\left(\frac{2}{k}, \frac{n-1}{k}\right)^2 B\left(\frac{1}{k}, \frac{n}{k}\right)^{-1} B\left(\frac{3}{k}, \frac{n-2}{k}\right)^{-1}\right)^{1/2}$

For every $r \in \mathbb{N}$, the r^{th} noncentered moment function of x is, as in function (5) of Theodossiou (1998):

$$E(x^r) = ((-1)^r (1-\lambda)^{r+1} + (1+\lambda)^{r+1}) \frac{\sigma^r}{S(\lambda)^r} \frac{B(\frac{r+1}{k}, \frac{n-r}{k}) B(\frac{1}{k}, \frac{n}{k})^{-1+(r+2)}}{B(\frac{3}{k}, \frac{n-2}{k})^{r/2}}$$

Consequently, the first moment, the mean, is given replacing r by 1.

$$\begin{aligned}
\mu = E(x) &= 2\lambda \frac{\sigma}{S(\lambda)} \frac{B(\frac{2}{k}, \frac{n-1}{k})}{B(\frac{1}{k}, \frac{n}{k})^{1/2} B(\frac{3}{k}, \frac{n-2}{k})^{1/2}} \\
&= 2\lambda\sigma \frac{B(\frac{2}{k}, \frac{n-1}{k})}{((1+3\lambda^2) B(\frac{1}{k}, \frac{n}{k}) B(\frac{3}{k}, \frac{n-2}{k}) - 4\lambda^2 B(\frac{2}{k}, \frac{n-1}{k}))^{1/2}} \\
&= 2\lambda\sigma\theta^* \frac{B(\frac{2}{k}, \frac{n-1}{k})}{B(\frac{1}{k}, \frac{n}{k})}
\end{aligned}$$

for θ^* given by equation 2.6.

Next, the second and the third normalized moments are

$$m_3 = E(z^3) = E(x^3) - 3\mu\sigma^2 - \mu^3 =$$
$$= 4\lambda(1+\lambda^2)\frac{\sigma^3}{S(\lambda)^3}\frac{B(\frac{4}{k},\frac{n-3}{k})\times B(\frac{1}{k},\frac{n}{k})^{1/2}}{B(\frac{3}{k},\frac{n-2}{k})^{3/2}} - 3\mu\sigma^2 - \mu^3$$

$$m_4 = E(z^4) = E(x^4) - 4\mu m_3 - 6\mu^2\sigma^2 - \mu^4 =$$
$$= (1+10\lambda^2+5\lambda^4)\frac{\sigma^4}{S(\lambda)^4}\frac{B(\frac{5}{k},\frac{n-4}{k})\times B(\frac{1}{k},\frac{n}{k})}{B(\frac{3}{k},\frac{n-2}{k})^2} - 4\mu m_3 - 6\mu^2\sigma^2 - \mu^4$$

Therefore, the skewness and kurtosis estimated through the SGT distribution is given as:

$$\begin{aligned}
S_k = \frac{m_3}{\sigma^3} &= \frac{1}{\sigma^3}\left(4\lambda(1+\lambda^2)\frac{\sigma^3}{S(\lambda)^3}\frac{B(\frac{4}{k},\frac{n-3}{k})\times B(\frac{1}{k},\frac{n}{k})^{1/2}}{B(\frac{3}{k},\frac{n-2}{k})^{3/2}} - 3\mu\sigma^2 - \mu^3\right)\\
&= 4\lambda(1+\lambda^2)\frac{B(\frac{1}{k},\frac{n}{k})^{3/2}B(\frac{3}{k},\frac{n-2}{k})^{3/2}}{((1+3\lambda^2)B(\frac{1}{k},\frac{n}{k})B(\frac{3}{k},\frac{n-2}{k}))^{3/2}}\frac{B(\frac{4}{k},\frac{n-3}{k})B(\frac{1}{k},\frac{n}{k})^{1/2}}{B(\frac{3}{k},\frac{n-2}{k})^{3/2}} - \frac{1}{\sigma^3}(3\mu\sigma^2+\mu^3)\\
&= 4\lambda(1+\lambda^2)\frac{B(\frac{1}{k},\frac{n}{k})^4}{((1+3\lambda^2)B(\frac{1}{k},\frac{n}{k})B(\frac{3}{k},\frac{n-2}{k}))^{3/2}}\frac{B(\frac{4}{k},\frac{n-3}{k})}{B(\frac{1}{k},\frac{n}{k})} - \frac{1}{\sigma^3}(3\mu\sigma^2+\mu^3)\\
&= 4\theta^{*3}\lambda(1+\lambda^2)\frac{B(\frac{4}{k},\frac{n-3}{k})}{B(\frac{1}{k},\frac{n}{k})} - 2\theta^*\lambda\frac{B(\frac{2}{k},\frac{n-1}{k})}{B(\frac{1}{k},\frac{n}{k})}\left[3+4\theta^{*2}\lambda^2\frac{B(\frac{2}{k},\frac{n-1}{k})^2}{B(\frac{1}{k},\frac{n}{k})^2}\right]
\end{aligned}$$

where, the second part ($\frac{1}{\sigma^3}(3\mu\sigma^2-\mu^3)$) unrolls as follows:

$$\begin{aligned}
\frac{1}{\sigma^3}(3\mu\sigma^2+\mu^3) &= \frac{1}{\sigma^3}\left(3\left(2\lambda\sigma\theta^*\frac{B(\frac{2}{k},\frac{n-1}{k})}{B(\frac{1}{k},\frac{n}{k})}\right)\sigma^2 + \left(2\lambda\sigma\theta^*\frac{B(\frac{2}{k},\frac{n-1}{k})}{B(\frac{1}{k},\frac{n}{k})}\right)^3\right)\\
&= \left(3(2\lambda\theta^*)\frac{B(\frac{2}{k},\frac{n-1}{k})}{B(\frac{1}{k},\frac{n}{k})} + (2\lambda\theta^*)^3\frac{B(\frac{2}{k},\frac{n-1}{k})^3}{B(\frac{1}{k},\frac{n}{k})^3}\right)\\
&= 2\theta^*\lambda\frac{B(\frac{2}{k},\frac{n-1}{k})}{B(\frac{1}{k},\frac{n}{k})}\left[3+4\theta^{*2}\lambda^2\frac{B(\frac{2}{k},\frac{n-1}{k})^2}{B(\frac{1}{k},\frac{n}{k})^2}\right]
\end{aligned}$$

For the kurtosis we have:

$$
\begin{aligned}
K_u = \frac{m_4}{\sigma^4} &= \frac{1}{\sigma^4}\left((1+10\lambda^2+5\lambda^4)\frac{\sigma^4}{S(\lambda)^4}\frac{B(\frac{5}{k},\frac{n-4}{k})\times B(\frac{1}{k},\frac{n}{k})}{B(\frac{3}{k},\frac{n-2}{k})^2} - 4\mu m_3 - 6\mu^2\sigma^2 - \mu^4\right) \\
&= \frac{(1+10\lambda^2+5\lambda^4)}{S(\lambda)^4}\frac{B(\frac{5}{k},\frac{n-4}{k})\times B(\frac{1}{k},\frac{n}{k})}{B(\frac{3}{k},\frac{n-2}{k})^2} - \frac{1}{\sigma^4}(4\mu m_3 + 6\mu^2\sigma^2 + \mu^4) \\
&= \frac{\theta^{*4}[1+5\lambda^2(2+\lambda^2)]B(\frac{5}{k},\frac{n-4}{k})}{B(\frac{1}{k},\frac{n}{k})} - \frac{1}{\sigma^4}(4\mu m_3 + 6\mu^2\sigma^2 + \mu^4) \\
&= \frac{\theta^{*4}[1+5\lambda^2(2+\lambda^2)]B(\frac{5}{k},\frac{n-4}{k})}{B(\frac{1}{k},\frac{n}{k})} + 24\theta^{*2}\lambda^2\frac{B(\frac{2}{k},\frac{n-1}{k})^2}{B(\frac{1}{k},\frac{n}{k})^2} \\
&\times\left[6+12\theta^{*2}\lambda^2\frac{B(\frac{2}{k},\frac{n-1}{k})^2}{B(\frac{1}{k},\frac{n}{k})} - 8\theta^{*2}(1+\lambda^2)\frac{B(\frac{4}{k},\frac{n-3}{k})}{B(\frac{2}{k},\frac{n-1}{k})}\right]
\end{aligned}
$$

where the second part unrolls as follows:

$$
\begin{aligned}
-\frac{1}{\sigma^4}(4\mu m_3 + 6\mu^2\sigma^2 + \mu^4) &= -\frac{\mu}{\sigma^4}(4m_3 + 6\mu\sigma^2 + \mu^3) \\
&= -\frac{1}{\sigma^4}\left(2\lambda\sigma\theta^*\frac{B(\frac{2}{k},\frac{n-1}{k})}{B(\frac{1}{k},\frac{n}{k})}\right)\left(4\left(\frac{4\lambda(1+\lambda^2)\sigma^3}{S(\lambda)^3}\frac{B(\frac{4}{k},\frac{n-3}{k})\times B(\frac{1}{k},\frac{n}{k})^{1/2}}{B(\frac{3}{k},\frac{n-2}{k})^{3/2}}\right.\right. \\
&\left.\left. -3\mu\sigma^2 - \mu^3\right) + 6\sigma^2\left(2\lambda\sigma\theta^*\frac{B(\frac{2}{k},\frac{n-1}{k})}{B(\frac{1}{k},\frac{n}{k})}\right) + \left(2\lambda\sigma\theta^*\frac{B(\frac{2}{k},\frac{n-1}{k})}{B(\frac{1}{k},\frac{n}{k})}\right)^3\right) \\
&= \frac{2\theta^*\lambda}{\sigma^3}\frac{B(\frac{2}{k},\frac{n-1}{k})}{B(\frac{1}{k},\frac{n}{k})}\left(-4\left(4\sigma^3\lambda(1+\lambda^2)\theta^{*3}\frac{B(\frac{4}{k},\frac{n-3}{k})}{B(\frac{1}{k},\frac{n}{k})} - \frac{\sigma^3}{\sigma^3}(3\mu\sigma^2+\mu^3)\right)\right. \\
&\left. -6\sigma^3\left(2\lambda\theta^*\frac{B(\frac{2}{k},\frac{n-1}{k})}{B(\frac{1}{k},\frac{n}{k})}\right) - \sigma^3\left(2\lambda\theta^*\frac{B(\frac{2}{k},\frac{n-1}{k})}{B(\frac{1}{k},\frac{n}{k})}\right)^3\right) \\
&= 4\theta^{*2}\lambda^2\frac{B(\frac{2}{k},\frac{n-1}{k})^2}{B(\frac{1}{k},\frac{n}{k})^2}\times\left[6+12\theta^{*2}\lambda^2\frac{B(\frac{2}{k},\frac{n-1}{k})^2}{B(\frac{1}{k},\frac{n}{k})} - 8\theta^{*2}(1+\lambda^2)\frac{B(\frac{4}{k},\frac{n-3}{k})}{B(\frac{2}{k},\frac{n-1}{k})}\right] \\
&= 24\theta^{*2}\lambda^2\frac{B(\frac{2}{k},\frac{n-1}{k})^2}{B(\frac{1}{k},\frac{n}{k})^2}\times\left[6+12\theta^{*2}\lambda^2\frac{B(\frac{2}{k},\frac{n-1}{k})^2}{B(\frac{1}{k},\frac{n}{k})} - 8\theta^{*2}(1+\lambda^2)\frac{B(\frac{4}{k},\frac{n-3}{k})}{B(\frac{2}{k},\frac{n-1}{k})}\right]
\end{aligned}
$$

for θ^* given by equation 2.6.

For the case of the Generalised Pareto distribution we pay attention to the skewness and kurtosis presented in the skewness and kurtosis are given for $\xi < 1/3$ and $\xi < 1/4$, respectively, as follows:

$$
S_k = \frac{2(1+\xi)\sqrt{1-2\xi}}{(1-3\xi)} \tag{A.8}
$$

$$
K_u = \frac{3(1-2\xi)(2\xi^2+\xi+3)}{(1-3\xi)(1-4\xi)} \tag{A.9}
$$

B. Demonstration: equality of BenSaïda's and Theodossiou's formula

Let k, n, λ, μ and σ be the scaling parameters representing the height, tails, skewness, mean and variance of the distribution. And θ^* given by equation 2.6. We demonstrate the veracity of BenSaïda and Slim's probability density function starting with it and aiming to get Theodossiou's function by replacing parameters and specifying the passages.

$$\frac{k}{2\theta^*\sigma B\left(\frac{1}{k},\frac{n}{k}\right)\left[1+\frac{|x-\mu^*|^k}{(1+sign(x-\mu^*)\lambda)^k\sigma^k\theta^{*k})}\right]^{(n+1)/k}}$$

$$= \frac{k\sqrt{(1+3\lambda^2)B\left(\frac{1}{k},\frac{n}{k}\right)B\left(\frac{3}{k},\frac{n-2}{k}\right)-4\lambda^2 B\left(\frac{2}{k},\frac{n-1}{k}\right)^2}}{2\sigma B\left(\frac{1}{k},\frac{n}{k}\right)^2\left[1+\frac{|x-\mu^*|^k}{(1+\lambda sign(x-\mu^*))^k\sigma^k\theta^{*k}}\right]^{(n+1)/k}}$$

$$= \frac{k\sqrt{(1+3\lambda^2)\frac{B\left(\frac{3}{k},\frac{n-2}{k}\right)}{B\left(\frac{1}{k},\frac{n}{k}\right)}-4\lambda^2\frac{B\left(\frac{2}{k},\frac{n-1}{k}\right)^2}{B\left(\frac{1}{k},\frac{n}{k}\right)^2}}}{2\sigma B\left(\frac{1}{k},\frac{n}{k}\right)\left[1+\frac{|x-\mu^*|^k}{(1+\lambda sign(x-\mu^*))^k\sigma^k((\frac{k}{n-2})^{-1/k}\theta)^k}\right]^{(n+1)/k}}$$

$$= \frac{k\sqrt{(1+3\lambda^2)\frac{B\left(\frac{3}{k},\frac{n-2}{k}\right)}{B\left(\frac{1}{k},\frac{n}{k}\right)}-4\lambda^2\frac{B\left(\frac{2}{k},\frac{n-1}{k}\right)^2}{B\left(\frac{1}{k},\frac{n}{k}\right)^2}}}{2\sigma B\left(\frac{1}{k},\frac{n}{k}\right)\left[1+\frac{(\frac{k}{n-2})|\frac{x-\mu^*}{\sigma}|^k}{(1+\lambda sign(x-\mu^*))^k\theta^k}\right]^{(n+1)/k}}$$

$$= C\frac{1}{\left(1+\left(\frac{k}{n-2}\right)\frac{1}{\theta^k}\frac{|\frac{x-\mu^*}{\sigma}|^k}{(1+\lambda sign(x-\mu^*)^k}\right)^{(n+1)/k}}$$

where we can formulate C, replacing $S(\lambda)$ by its corresponding formula:

$$
\begin{aligned}
C &= \frac{k}{2\sigma} B\left(\frac{1}{k}, \frac{n}{k}\right)^{-3/2} B\left(\frac{3}{k}, \frac{n-2}{k}\right)^{1/2} S(\lambda) \\
&= \frac{k}{2\sigma} \frac{B(\frac{3}{k}, \frac{n-2}{k})^{1/2}}{B(\frac{1}{k}, \frac{n}{k})^{3/2}} \left(1 + 3\lambda^2 - 4\lambda^2 \frac{B(\frac{2}{k}, \frac{n-1}{k})^2}{B(\frac{3}{k}, \frac{n-2}{k}) B(\frac{1}{k}, \frac{n}{k})}\right)^{1/2} \\
&= \frac{k}{2\sigma} \frac{1}{B(\frac{1}{k}, \frac{n}{k})} \left((1 + 3\lambda^2) \frac{B(\frac{3}{k}, \frac{n-2}{k})}{B(\frac{1}{k}, \frac{n}{k})} - 4\lambda^2 \frac{B(\frac{2}{k}, \frac{n-1}{k})^2}{B(\frac{1}{k}, \frac{n}{k})^2}\right)^{1/2}
\end{aligned}
$$

C. Kolmogorov-Smirnov table

n\α	0.001	0.01	0.02	0.05	0.1	0.15	0.2
1		0.99500	0.99000	0.97500	0.95000	0.92500	0.90000
2	0.97764	0.92930	0.90000	0.84189	0.77639	0.72614	0.68377
3	0.92063	0.82900	0.78456	0.70760	0.63604	0.59582	0.56481
4	0.85046	0.73421	0.68887	0.62394	0.56522	0.52476	0.49265
5	0.78137	0.66855	0.62718	0.56327	0.50945	0.47439	0.44697
6	0.72479	0.61660	0.57741	0.51926	0.46799	0.43526	0.41035
7	0.67930	0.57580	0.53844	0.48343	0.43607	0.40497	0.38145
8	0.64098	0.54180	0.50654	0.45427	0.40962	0.38062	0.35828
9	0.60846	0.51330	0.47960	0.43001	0.38746	0.36006	0.33907
10	0.58042	0.48895	0.45662	0.40925	0.36866	0.34250	0.32257
11	0.55588	0.46770	0.43670	0.39122	0.35242	0.32734	0.30826
12	0.53422	0.44905	0.41918	0.37543	0.33815	0.31408	0.29573
13	0.51490	0.43246	0.40362	0.36143	0.32548	0.30233	0.28466
14	0.49753	0.41760	0.38970	0.34890	0.31417	0.29181	0.27477
15	0.48182	0.40420	0.37713	0.33760	0.30397	0.28233	0.26585
16	0.46750	0.39200	0.36571	0.32733	0.29471	0.27372	0.25774
17	0.45440	0.38085	0.35528	0.31796	0.28627	0.26587	0.25035
18	0.44234	0.37063	0.34569	0.30936	0.27851	0.25867	0.24356
19	0.43119	0.36116	0.33685	0.30142	0.27135	0.25202	0.23731
20	0.42085	0.35240	0.32866	0.29407	0.26473	0.24587	0.23152
25	0.37843	0.31656	0.30349	0.26404	0.23767	0.22074	0.20786
30	0.34672	0.28988	0.27704	0.24170	0.21756	0.20207	0.19029
35	0.32187	0.26898	0.25649	0.22424	0.20184	0.18748	0.17655
40	0.30169	0.25188	0.23993	0.21017	0.18939	0.17610	0.16601
45	0.28482	0.23780	0.22621	0.19842	0.17881	0.16626	0.15673
50	0.27051	0.22585	0.21460	0.18845	0.16982	0.15790	0.14886
OVER 50	$\frac{1.94947}{\sqrt{n}}$	$\frac{1.62762}{\sqrt{n}}$	$\frac{1.51743}{\sqrt{n}}$	$\frac{1.35810}{\sqrt{n}}$	$\frac{1.22385}{\sqrt{n}}$	$\frac{1.13795}{\sqrt{n}}$	$\frac{1.07275}{\sqrt{n}}$

Figure C.1.: Kolmogorov-Smirnov table

D. MatLab Code

The code used to the analysis and representation of the studied data may heavy this dissertation content. Therefore, all the code used to gather the data, graphics and econometric methodologies is available on the open source repository: https://github.com/DanieSofia/patternskewness.git. There is also code from another author (BenSaïda and Slim, 2016) that is not publicly shared in the aforementioned repository. Some fundamental code is chosen from several types of public code according to the interest of the programmer.