Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/120329

Fitting Statistical Parent Distributions to Quantify Financial Risk in the South African Financial Index (J580)

*D. Chikobvu and O. Jakata**

Department of Mathematical Statistics and Actuarial Sciences, Faculty of Natural and Agricultural Sciences, University of the Free State, Bloemfontein, South Africa

Received: 24 Sep. 2022, Revised: 4 Jan. 2023, Accepted: 15 Jan. 2023.
Published online: 1 Sep. 2023.

Abstract: The purpose of this study is to investigate and describe the riskiness of an investment in the South African Financial Index (J580) using four relatively heavy tailed statistical parent distributions, viz: the Exponential, Weibull, Gamma and Burr distributions. The statistical distributions describe the Index returns, and quantify the riskiness of the monthly South African Financial Index (J580) for the period 1995-2018. The Maximum Likelihood Estimation (MLE) method is used to estimate the distribution parameters. The heavier-tailed Burr distribution in the heavy tailed Frétchet domain distribution is the best fitting statistical parent distribution for losses as evidenced by the AIC, BIC and other graphical measures of goodness of fit. The lighter tailed Exponential distribution is the best fitting statistical parent distribution for the positive returns (gains). The Exponential distribution is in the light tailed Gumbel domain distribution. Summary measures of financial risk, such as the Value at Risk (VaR) and Expected Shortfall (ES) are calculated using the two best fitting distributions. Financial risk (VaR and ES) quantification and risk mitigation is topical in light of the failure of the Normal distribution-based risk models, which under estimated risk in leading up to the Global Financial Crisis (GFC) of 2008-2009. The practical implications are that the Normal distribution-based risk measures ought to be replaced with other statistical parent distributions and even extreme value distributions (EVD) in order to accurately estimate financial risk. Given the limited empirical investigations on the South Africa Financial Index (J580), the results from this research provide additional and valuable information for both investors and practitioners on how to accurately estimate and assess financial risk. The study extends the empirical literature on more accurate financial risk assessment, more specifically in the context of the Financial Index in South Africa.

Keywords: Burr Distribution, Gamma Distribution, Exponential Distribution, Value at Risk, Expected Short Fall

1 Introduction

In finance and insurance, one area of interest is the statistical distributions of financial returns. It is generally assumed that financial return variables follow certain statistical distributions. The Normal distribution was once popular, but does little to carter for some extremes in returns that are often associated with financial data. The skewed and heavier-tailed distributions are the most appropriate to fit when working with financial returns data, since they account for skewness and excess kurtosis. Many statistical parent distributions have been considered in many different situations, and these include the Gamma distribution [1], Log-normal distribution [2], and the Log-logistic distribution [3]. The aim is to find a way of quantifying statistically, the riskiress associated with financial returns.

Financial returns, according to [4], have relatively heavy tails when compared to the Normal distribution. According to [5], financial returns provide a rich source of variable information with a variety of properties, ranging from Normally distributed variables to distributions with varied degrees of skewness and kurtosis. [6] concluded that some statistical distributions possess thick tails that are better suited to modeling financial losses. This paper quantifies and describes the South African Financial Index (J580) returns using four relatively heavy-tailed statistical parent distributions, viz: the Exponential, Weibull, Gamma and the Burr distributions. The chosen parent distributions are able to capture varied degrees of skewness and kurtosis and provide for the varying degrees of relatively heavy tailedness.

The statistical parent distributions use the full dataset in modelling and they concentrate the fit on the main body of the financial returns data, which is around the mean, mode and/or the median. The four relatively heavy-tailed statistical parent distributions are used to fit the positive (gains) and negative(losses) returns separately. There are many standard theoretical distributions according to [7], and the Weibull, Exponential, Logistic, Generalised Logistic, Gompertz, Normal, Extreme value, and Uniform distributions are special cases or limiting cases of the Burr distribution. The four distributions are sufficient to cater for varying degrees of relatively heavy tailedness.

*Corresponding author e-mail: owenjakata@rocketmail.com

Extreme value distributions are statistical distributions of the largest values drawn from a sample of a given size. The two main models of extreme value distributions are the Generalised Extreme Value Distributions (GEVD) and Generalised Pareto Distribution (GPD). These distributions however only cater for the very extreme losses or gains whilst ignoring the main body of the data. The statistical parent distributions on the other hand, strike a fine balance in modelling the main body of the data and some extreme values associated with the dataset. According to [8] , the GEVD unifies the Gumbel, Fréchet, and the negative Weibull class statistical distributions.

Statistical parent distributions provide a description of risk exposure, according to [9]. Key risk metrics like Value at Risk (VaR) and Expected Shortfall (ES) can be used to describe the extent of risk exposure. These important risk metrics are used by investors to gauge how exposed their firms are to risk, which can result from shifts in underlying variables like stock prices, interest rates, and exchange rates.

The main objective of this study is to identify the best-fitting statistical parent distributions for describing the risk in the return distribution of the South African Financial Index (J580). This study suggests some key alternatives to the Normal distribution when fitting to the Index returns. This aids in quantifying the financial risk using key risk metrics, such as VaR and ES. In the context of the South African stock market, a few studies on statistical parent distributions have been conducted. This study fills in the gap by determining the best-fit statistical parent statistical distributions for modelling the South African Financial Index (J580) returns data and quantifying the Index's risk level.

1.1 Statement of the Problem

This study analyses the empirical performance of four proposed statistical parent distributions in describing and quantifying risk in the monthly financial returns of the Index. Fitting the distributions provides a risk description by selecting a suitable statistical parent distribution. The four relatively heavy-tailed distributions, viz: the Exponential, Weibull, Gamma and Burr distributions are compared in terms of goodness of fit to find the best-fitting statistical parent distribution. The statistical parent distributions concentrate their fit on the main body of the sample financial returns data, which is around the mean, mode and the median, unlike the Extreme Value distributions which concentrate their fit exclusively at the extremities of the returns-distribution. The best-fitting statistical parent distributions for both the losses and gains returns are identified separately using the AIC and BIC criteria. The heavier-tailed statistical parent distributions strike a balance in accommodating the location of the bulk of the data and the thicker tails often found in financial data.

1.2 Significance of the Study

The financial sector is an important player in the South African economy; it promotes economic growth by promoting trade and commercial activities. It is crucial to modelling and quantifying the risk associated with the returns from the financial sector. The performance of the financial sector often affects other Industrial sectors. This study focuses specifically on the South African Financial Index (J580) returns.

1.3 Objectives of the Study

The main objectives of the study are to:

- Model the main body of the monthly South African Financial Index (J580) returns using four relatively heavy-tailed statistical parent distributions.
- Determine the best-fitting model for the Index returns.
- Quantifying the risk associated with the Index and comparing the risk of gains and losses.

The contribution of this study is the identification of the most appropriate statistical parent distributions for the South African Financial Index (J580) returns and the quantifying of VaR and ES as proxies for future risk. This will help investors considering investing in the Index in comprehending the risks/rewards associated with the Index, and also allowing for the quantification of the capital required to meet any regulatory requirements.

1.4 Summary of literature on statistical parent distributions

There are many studies that have been done on statistical parent distributions and a few are discussed in this section.

[10] employed the Normal Inverse Gaussian (NIG) distribution for a Vector Auto Regressive (VAR) valuation in the South African and USA stock market. The paper compared the NIG distribution with the Normal distribution, Skew Student's t-distribution and Student t-distribution, each capturing different/varying features of the financial returns. The three distributions gave a better fit than the Normal distribution. The VAR estimates from the Normal distribution proved to be inferior to the estimates obtained from the other distributions; large negative returns were accommodated and were more likely to occur in the other three distributions than in the Normal distribution case.

[11] used a class of semi-parametric Generalised long-memory models with Fractionally Integrated Asymmetric Power Autoregressive Conditional Heteroskedastic (FIAPARCH) errors. They used daily stock market Indices in eight MENA countries, namely, Bahrain, Egypt, Jordan, Kuwait, Oman, Qatar, Saudi Arabia and United Arab Emirates (UAE) over the period from May 31, 2005 to April 15, 2015. The researchers used the wavelet-based maximum likelihood estimator to estimate the proposed models. Their model proved to be a better fit than the traditional long-memory models; and their findings also showed that past prices can be used to forecast future prices.

[12] used a two-parameter Weibull distribution for modelling the financial return distributions and estimating tail-related risk measures. The model was fitted to financial returns from Global stock market indices, viz: the S&P 500 (US), FTSE 100 (UK), the All Ordinaries Index (Australia), and the HANG SENG Index (Hong Kong); as well as two exchange rate series, viz: the Australian (AU) dollar to the United States of America (US) dollar and the European Euro to the US dollar; and a single asset series: IBM. The findings revealed that the two-parameter Weibull performed most favourably for VaR estimation, before and after the Global Financial Crisis (GFC) (2007-2008).

[13] analysed the daily Japanese Nikkei 225 Index returns for the period 1984 - 2002. The study investigated if the Exponential distribution was suitable to model the distributions of returns, volatility and calm-time interval distribution of the volatility. A graphical diagnostic semi-log plot was utilised to assess the goodness-of- fit. The results revealed that empirical distribution of the returns, the volatility and the calm-time interval for the volatility can be described by the Exponential distribution. The Exponential distribution was found to be more suitable compared to the Normal distribution.

According to [14], the NIG distribution is a flexible distribution with four parameters enabling it to capture skewness and hence heavy-tailedness in financial time series data. The study used seven stocks listed on the Norwegian Stock Exchange and one listed on the New York Stock Exchange. In their study on risk quantification, the authors compared the NIG to a Normal distribution and a non-parametric model. The researchers demonstrated that the NIG distribution outperforms the Normal distribution and fits the log returns of eight stocks well in both the tails (upper and lower) and in the centre.

There many other studies done on statistical parent distributions, but not discussed in this study, which include [15], [16], [17], [18], [19], [20], [21] and [22]. To the best of authors' knowledge, this is one of the few studies that adopts the Exponential, Weibull, Gamma, Burr distributions in modelling the return distribution and forecasting the risk level of the of the South African Financial Index (J580). This provides important information to local and international investors who wish to improve portfolio diversification and reduce contagion in a globalised world economy. Developing economies stocks are less correlated with other Global markets. This study is organised as follows: section 2 presents the Research models, and section 3 presents the Results and the discussion. Section 4 presents Conclusion and Discussion.

2. Research Models

This study applies the statistical parent distributions approach for modelling non-Normal returns distributions in the context of the South African Stock market. Four statistical parent distributions, viz: the Exponential, Weibull, Gamma and Burr distributions are proposed. The chosen distributions do cater for the relatively heavy-tailed returns data and are able to capture various degrees of skewness and kurtosis. The four statistical parent distributions which are used in this financial returns distribution analysis, are discussed in this section.

2.1 Exponential distribution

[23] describes a heavy-tailed distribution as having a tail that is heavier than an Exponential distribution. The Exponential distribution gives a good starting point relative to our presumption on the nature of the returns data. The Probability Density Function (PDF) and Cumulative Distribution Function (CDF) of the Exponential distribution are respectively denoted as:

$$f(x;\lambda) = \lambda e^{-\lambda x} \qquad (1)$$

$$F(x;\lambda) = 1 - e^{-\lambda x} \qquad (2)$$

where, x represents the log returns and $\lambda > 0$ is the rate parameter.

The Maximum Likelihood Estimation (MLE) parameter for the Exponential distribution is given in the following theorem:

Theorem 1

If X is exponentially distributed with the pdf given in equation (1) where $\lambda > 0$ then, the maximum likelihood estimate of λ is given as:

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i}$$

n is the total number in the sample data set.

2.2 Weibull distribution

The Weibull distribution adds a shape parameter to the Exponential distribution, hence making it more flexible. The PDF and CDF of the Weibull distribution are respectively denoted as:

$$f(x;\lambda,k)=\frac{kx^{k-1}}{\lambda^{x}}e^{-\left(\frac{x}{\lambda}\right)^{k}} \quad (3)$$

$$F(x;\lambda,k)=1-\ e^{-\left(\frac{x}{\lambda}\right)^{k}} \quad (4)$$

where, $x>0$ represents the log returns, $\lambda>0$ and $k>0$ represent the scale and shape parameters respectively. The parameters for the Weibull as estimated by the MLE method are given in the following theorem.

Theorem 2

If X follows a Weibull distribution with parameters $\lambda>0$ and $k>0$ and a pdf given in equation (3) then, the MLE of λ is given as:

$$\hat{\lambda}=\frac{n}{\sum_{i=1}^{n}x_i{}^{k}}$$

where, k is estimated by solving equations7 numerically.

$$n+k\sum_{i=1}^{n}ln\,x_i=\frac{nk\sum_{i=1}^{n}x_i{}^{k}ln x_i}{\sum_{i=1}^{n}x_i{}^{k}}$$

2.3 Gamma distribution

The Gamma distribution is another two-parameter distribution from the Exponential family of distributions. The PDF and CDF of the Gamma distribution are respectively denoted as:

$$f(x;\alpha,\beta)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x} \quad (5)$$

$$F(x;\alpha,\beta)=\frac{1}{\beta^{\alpha}\Gamma(\alpha)}\int_0^x t^{\alpha-1}e^{-\frac{t}{\beta}}dt \quad (6)$$

With the continuous random variable x represents the log returns, α and β represents the shape and scale parameters respectively. Theorem 3 gives the Gamma parameters as estimated by the MLE method.

Theorem 3

If X follows a Gamma distribution with a PDF given in equation (5), where the parameters $\alpha>0$ and $\beta>0$ then, the maximum likelihood estimate of β is given as:

$$\hat{\beta}=\frac{n\alpha}{\sum_{i=1}^{n}x_i}$$

Substitute the solution of equation 11 into $\frac{\delta}{\delta\alpha}ln\{L(\alpha,\beta)\}$ to obtain a nonlinear equation for the maximum likelihood estimate of α. Numerical methods are then used to solve the resultant equation.

2.4 Burr distribution

According to [24], the Burr distribution was first introduced in 1942 by I. W. Burr and it and is known as Burr Type XII distribution. The Burr distribution is a three-parameter heavy-tailed distribution. The additional parameter makes the distribution more flexible and gives a better fit if the log returns data is heavy-tailed.

The PDF and CDF of the Burr distribution are respectively denoted as:

$$f(x;\alpha,k,\beta)=\frac{\alpha k}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1}\left(1+\left(\frac{x}{\beta}\right)^{\alpha}\right)^{-(k+1)} \quad (7)$$

$$F(x;\alpha,k,\beta)=1-\left(1+\left(\frac{x}{\beta}\right)^{\alpha}\right)^{-k} \quad (8)$$

The Burr distribution has two shape parameters, $\alpha>0$ and $k>0$. $\beta>0$ represents the scale parameter. Theorem 4 presents the maximum likelihood estimates of the Burr distribution.

Theorem 4

If X follows a Burr distribution with parameters α > 0, k > 0 and β > 0, and a PDF given in equation 7 then, the maximum likelihood estimate of k is given as:

$$\hat{k} = \frac{n}{\sum_{i=1}^{n} ln\left[1+\left(\frac{x_i}{\beta}\right)^{\alpha}\right]}$$

The maximum likelihood estimate of α is obtained by solving the following equation:

$$\frac{n}{\alpha} - nln\beta - \sum_{i=1}^{n} ln(x_i) + (k+1)\left\{\sum_{i=1}^{n}\left[\frac{\left(\frac{x_i}{\beta}\right)^{\alpha}}{1+\left(\frac{x_i}{\beta}\right)^{\alpha}}\right] ln\left(\frac{x_i}{\beta}\right)\right\} = 0$$

The maximum likelihood estimate of β is obtained by solving the equation above and the one below.

$$(k+1)\left\{\sum_{i=1}^{n}\left[\frac{\frac{\alpha}{\beta}\left(\frac{x_i}{\beta}\right)^{\alpha}}{1+\left(\frac{x_i}{\beta}\right)^{\alpha}}\right]\right\} - \frac{n\alpha}{\beta} = 0$$

2.5 Risk Measures

In this section, the formulas used to calculate and quantify risk in the the South African Financial Index (J580) returns data, Value at Risk (VaR) and the Expected Shortfall (ES) for the proposed parent distributions are presented.

Exponential distribution VaR and ES equations

$$VaR_p(\mathrm{X}) = -\frac{1}{\lambda}\log(1-\mathrm{p}) \quad (9)$$

$$ES_p(\mathrm{X}) = -\frac{1}{p\lambda}\{\log(1-p)\,p - p - \log(1-p)\} \quad (10)$$

for x > 0 , 0 < p < 1 , and λ > 0, the scale parameter.

Weibull distribution VaR and ES equations

$$VaR_p(X) = \lambda[-log(1-p)]^{\frac{1}{k}} \quad (11)$$

$$ES_p(X) = \frac{\lambda}{p}\left(1+\frac{1}{k}, -log(1-p)\right) \quad (12)$$

for x > 0 , 0 < p < 1 , k > 0, the shape parameter, and λ > 0, the scale parameter.

Gamma distribution VaR and ES equations

$$VaR_p(X) = \frac{1}{\beta}Q^{-1}(a, 1-p) \quad (13)$$

$$ES_p(X) = \frac{1}{\beta p}\int_0^p Q^{-1}(\alpha, 1-p)dv \quad (14)$$

for x > 0 , 0 < p < 1 , β > 0, the scale parameter, and $\alpha > 0$, the shape parameter.

Q $(\alpha;$ x) = denotes the regularised complementary incomplete gamma function.

Burr distribution VaR and ES equations

$$VaR_p(X) = [(1-p)^{\frac{1}{k}} - 1]^{\frac{1}{\alpha}} \quad (15)$$

$$ES_p(X) = \frac{1}{p}\int_0^p [(1-p)^{\frac{1}{k}} - 1]^{\frac{1}{\alpha}} \quad (16)$$

for x > 0 , 0 < p < 1 , α > 0, the first shape parameter, and k > 0, the second shape parameter.

The models for VaR and ES for the four relatively heavy-tailed distributions are adopted from R Statistical software package [25]. The VaR and ES for the best-fitting distributions: Exponential distributions and Weibull distributions are forecasted using the VaRES R-statistical software package.

2.6 Testing for Stationarity, Normality, Heteroscedasticity and Autocorrelation

The Augmented Dickey-Fuller (ADF) Test is used to test whether the Index return series is stationary. The Normality of the Index is tested using the Andersen Darling Test. The presence of heteroscedasticity in the residuals is tested using the Lagrange Multiplier (LM) test proposed by [26]. A Box-Ljung test statistic tests is used to test for autocorrelation in the data set.

2.7 Research Data

The monthly South African Financial Index (J580) secondary data (years 1995-2018) obtained (with permission) from the website iress expert: https://expert.inetbfa.com was used in this study. The South African financial sector is one of the three main sub-indices of the South African All Share Index (ALSI). It is defined as the banking, insurance and securities industries [27]. [28], stated that the financial sector contributes a quarter of the total economic growth and employs over 220,000 people in South Africa. According to the IMF Report (2014), South Africa's financial sector is large and sophisticated and the assets amount to 29.8 per cent of the GDP. Therefore, it is important to model the returns from the financial sector in order to quantify the riskiness of the index.

In this study, losses are positive since the loss function in period t for an index log return is:

$$X_t = -r_t = -\ln\left(\frac{M_t}{M_{t-1}}\right) \tag{17}$$

r_t is the monthly log returns in month t, M_t represents the monthly index in month t and ln represents the natural logarithm. When using a loss function, the losses (negative returns multiplied by -1) are positive.

3. Results

The data is analysed in the R-programming environment using packages fitdistrplus, actuar, ReIns and VARES is used to quantify VaR and ES.

3.1 Descriptive Statistics

Table 1: Monthly log returns of the South African Financial Index (J580) descriptive statistics.

Description	Values	Description	Values
Mean	-0.835	Variance	0.36512
Median	-0.0101	Standard Deviation	0.06042
Maximum	0.51195	Skewness	2.194
Minimum	-0.21652	Kurtosis	19.440

Source: Authors' own work.

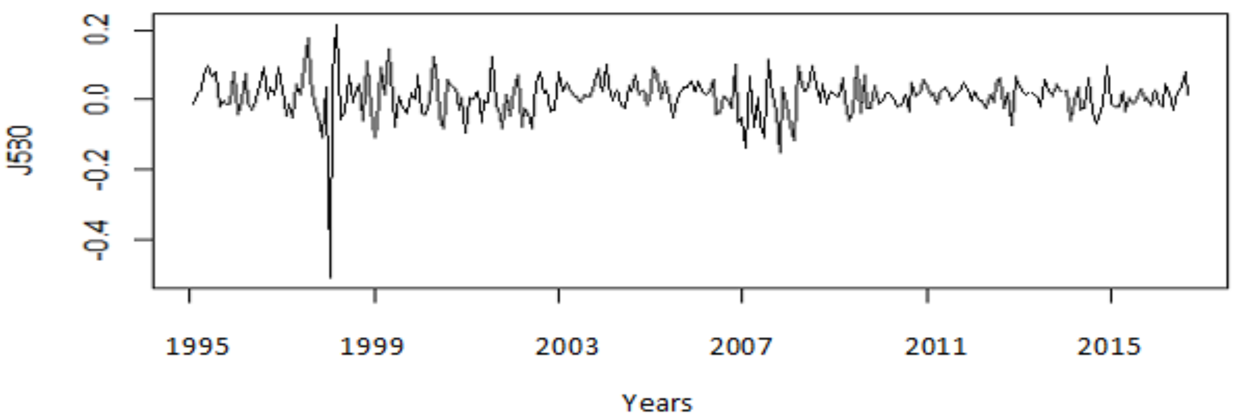
Table 1 shows that monthly South African Financial Index (J580) returns descriptive statistics. Results suggest the distribution of returns has a heavy-tail. The mean is –0.835 with a small standard deviation of 0.06042. The dataset is significantly skewed to the right and also exhibit excess kurtosis with a positive skewness coefficient (2.194>0) and large kurtosis (19.440>3). The results suggest that the skewed and fatter-tailed statistical parent distributions are the most appropriate to fit this type of data since they account for skewness, excess kurtosis and hence the heavier-tails (Hakim et al 2018).

3.2 The graphical plot of monthly Index values M_t

The graph (Figure 1) shows the monthly movement pattern of the index over the past 23 years and clearly shows an upward trend (Figure 1)

Fig.1: Time series graph of the Index's monthly values, M_t.

In Figure 1, the financial crises which had a negative impact on the South African stock market are indicated on the time series plot by sharp down turns of the index levels. The logarithmic return series data (x_t) is in Figure 2.



Source: Authors' own work.

Fig. 2: Time series graph of the Index's monthly returns, x_t.

3.3 Testing for Stationarity

The p-value for the ADF test is 0.01 which is less than the significance level of 0.05, and hence the null hypothesis of a unit root is rejected. The conclusion is that the returns data is stationary

3.4 Test for Normality.

The p-value for the Andersen Darling Normality test is 6.903e-08 which is less than 0.05. This implies that the hypothesis of Normality is rejected. It is concluded that the return data series is not Normally distributed. This again suggests the returns follow a fat-tailed distribution.

3.5 Test for Heteroscedasticity

The p-value for the ARCH LM test is 1, which is greater than 0.05. This indicates no presence of significant ARCH effects in the returns data. This reveals that there is no persistence of variance and no evidence of volatility clustering in the returns data.

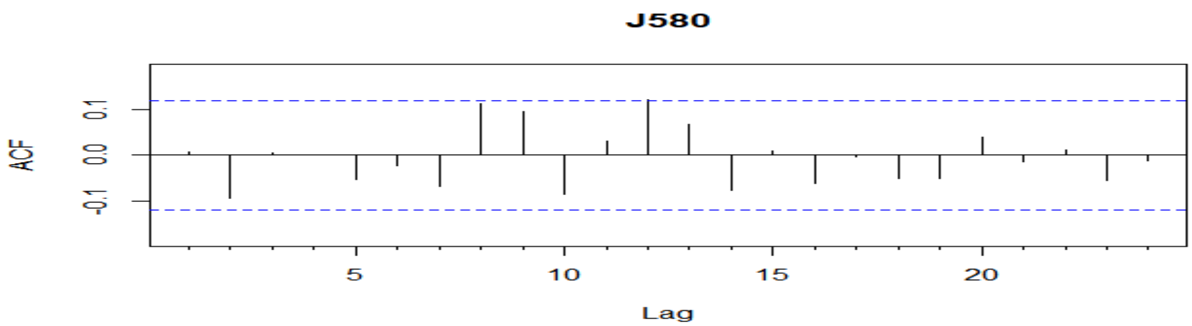


Fig 3: ACF diagram

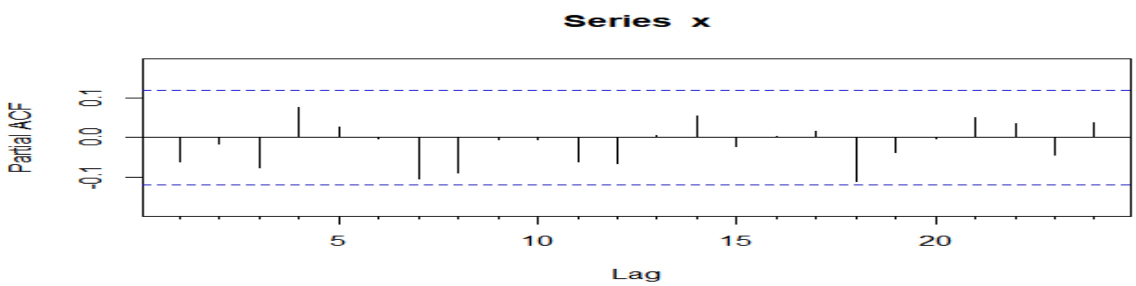


Fig 4: PACF Diagram

3.6 Test for Autocorrelation

The autocorrelation function (ACF) and partial autocorrelation function (PACF) plots indicate that there is no significant auto-correlations in the returns data. The Box-Ljung test has a p-value = 0.8998 is greater than 0.05, implies that the null hypothesis of no autocorrelation is not rejected. This means that the returns are independently distributed.

3.7 Fitting and selecting distributions

In this section, the gains and the losses are separated out and analysed separately by fitting the four statistical parent distributions, namely the Exponential, the Weibull, the Gamma and the Burr statistical distributions in the stated order.

3.7.1 Exponential Distribution

The Exponential distribution fit is depicted in Figure 5 (for losses) and Figure 6 (for gains) using diagnostic plots. The MLE method was used to estimate parameters and their standard errors.

a) Exponential Losses

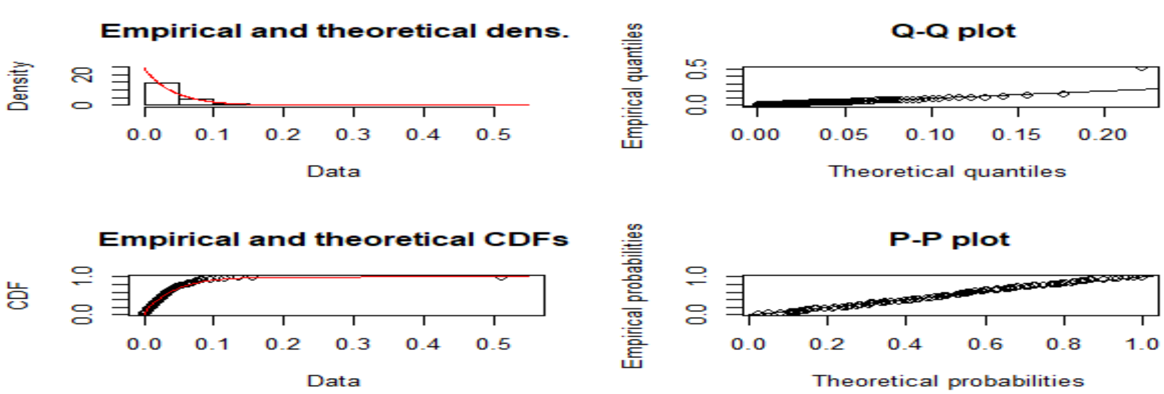


Fig 5: Diagnostic plots for the Exponential loss returns

b) Exponential Gains

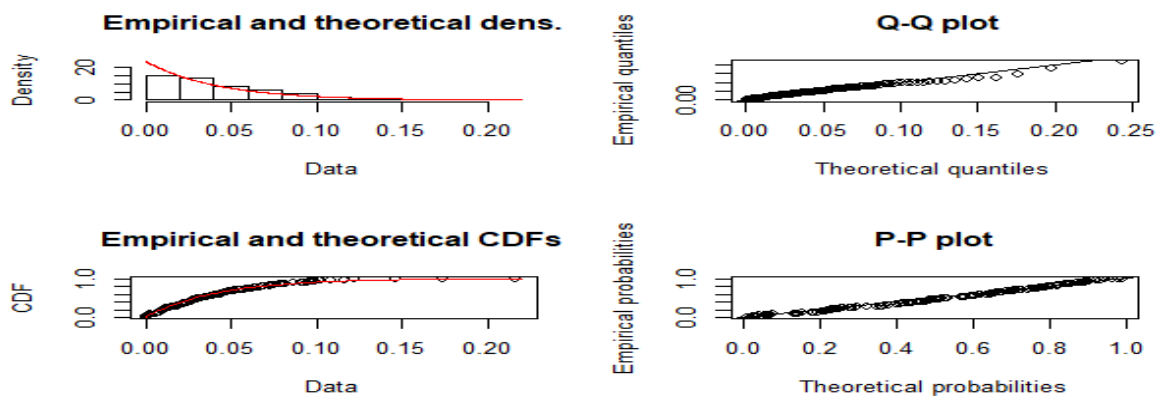


Fig. 6: Diagnostic plots for the Exponential gain returns.

The data points exhibit linearity on the P-P and Q-Q plots, and show insignificant deviation from the 45°line in Figure 5 and Figure 6. From the diagnostic plots, the Exponential distribution is a good fit for the data.

Table 2: Parameter estimates for the Exponential fit

	Parameter Estimate	Standard Error
Estimate of rate parameter (λ) for losses	0.0412269639	0.432920447
Estimate of rate parameter (λ) for gains	0,0422276076	0.535808339

The Maximum Likelihood method estimate parameters with their respective standard errors are given in Table 2 for both the losses and the gains.

3.7.2 Weibull Distribution

The Weibull distribution fit is depicted in Figure 7 (for losses) and Figure 8 (for gains) using diagnostic plots. The MLE method was again used to estimate the parameters and their standard errors.

a) Weibull losses

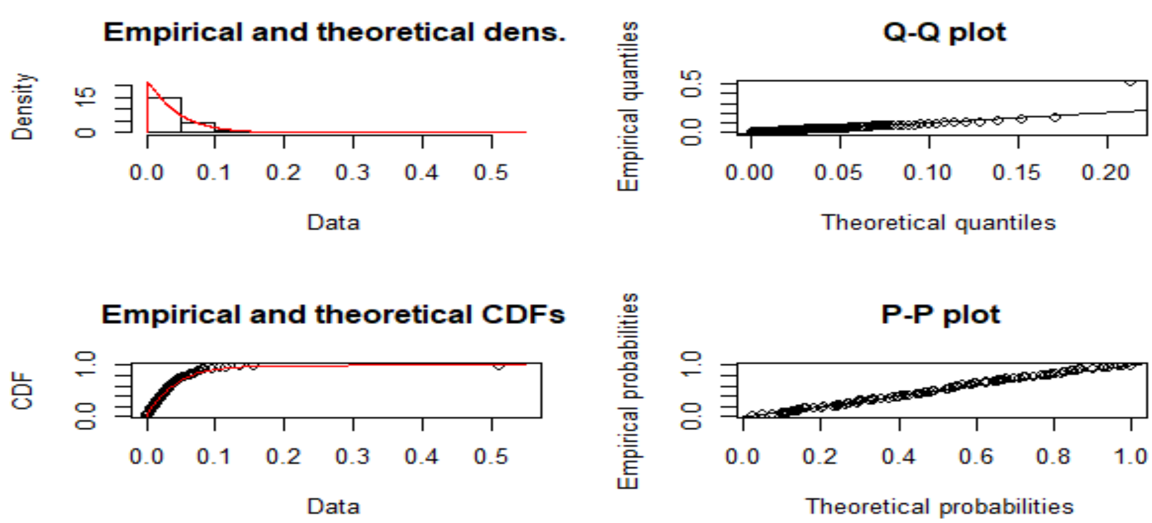


Fig. 7: Diagnostic plots for the Weibull losses

b) Weibull gains

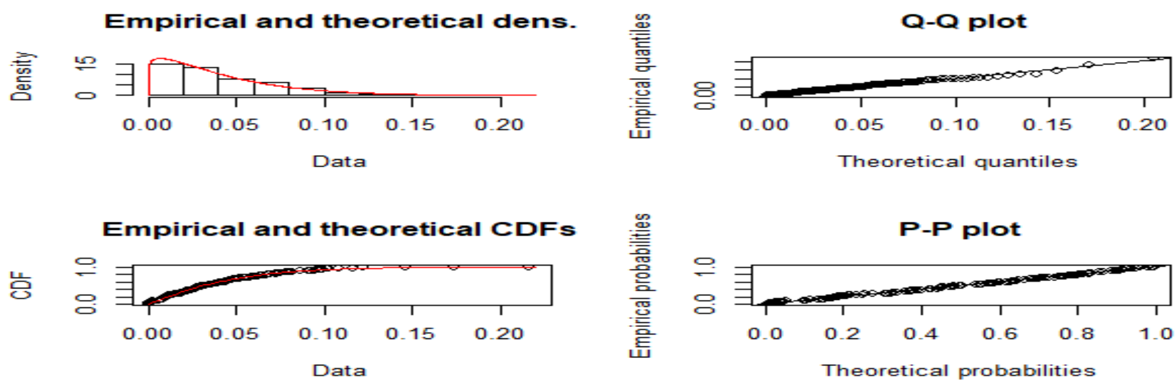


Fig. 8: Diagnostic plots for the Weibull gains

On the P-P and Q-Q plots in Figure 7 and Figure 8, there is no deviation from the 45° line, indicating that the Weibull is a good fit for the data. The error values for the two parameters are relatively low for both the losses and gains, as shown in Table 3.

Table 3: Parameter estimates for the Weibull fit

	Parameter Estimate	Standard Error
Estimate of shape (k) for losses	1.03177619	0.068555192
Estimate of scale (λ) for losses	0.04183508	0.004083269
Estimate of shape (k) for gains	1.1384825	0.07136430
Estimate of scale (λ) for gains	0.0441436	0.00320305

The Maximum Likelihood method estimate parameters with their respective standard errors are given in Table 3 for both the losses and the gains.

3.7.3 Gamma Distribution

The Gamma distribution fit is depicted in Figure 9 (for losses) and Figure 10 (for gains) using diagnostic plots. The MLE method was also used to estimate the parameters and their standard errors.

a) Gamma losses

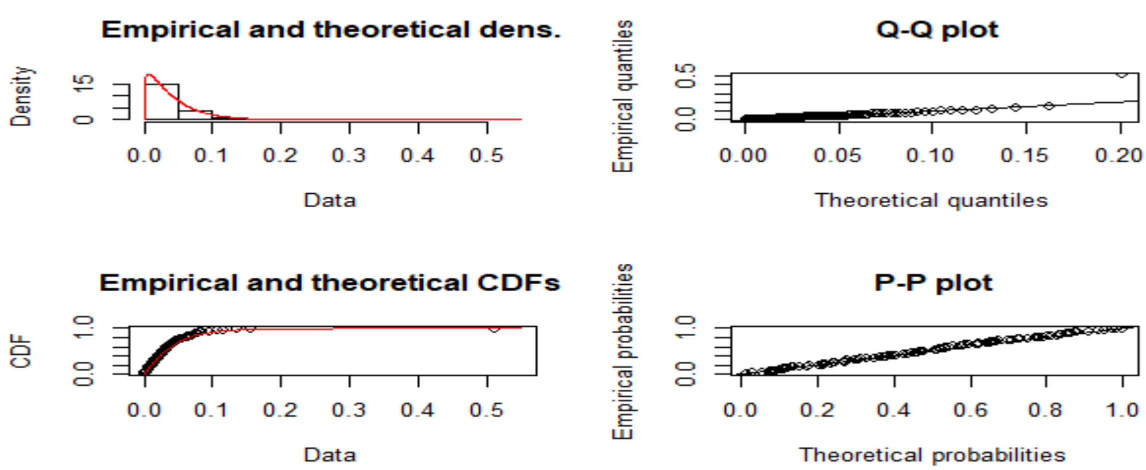


Fig.9: Diagnostic plots for the Gamma losses

b) Gamma gains

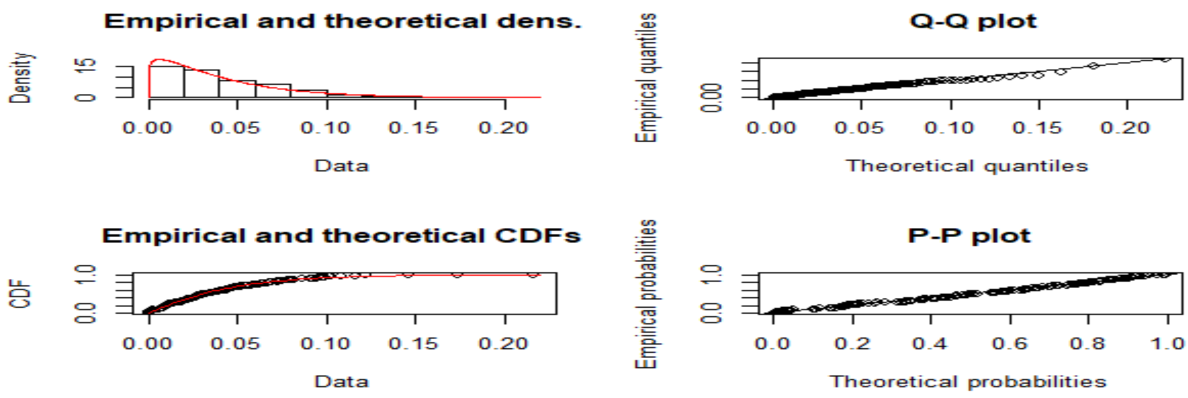


Fig. 10: Diagnostic plots for the Gamma gains

In Figure 9 and Figure 10, on the P-P and Q-Q plots, there is insignificant deviation from the reference line with the probabilities matching for most of the sample points for both the losses and gains.

Table 4: Parameter estimates for the Gamma fit

	Parameter Estimate	Standard Error
Estimates of shape (α) for losses	1.19493	0.1440603
Estimates of scale (β) for losses	0.03444962359	0.023182
Estimates of shape (α) for gains	1.178297	0.117282
Estimates of scale (β) for gains	0.0358372626	0.029089

The parameter estimates and their respective standard errors are given in Table 4.

3.7.4 Burr Distribution

The Burr distribution fit is depicted in Figure 11 (for losses) and Figure 12 (for gains) using diagnostic plots. The MLE method was also used to estimate the parameters and their standard errors.

a) Burr losses

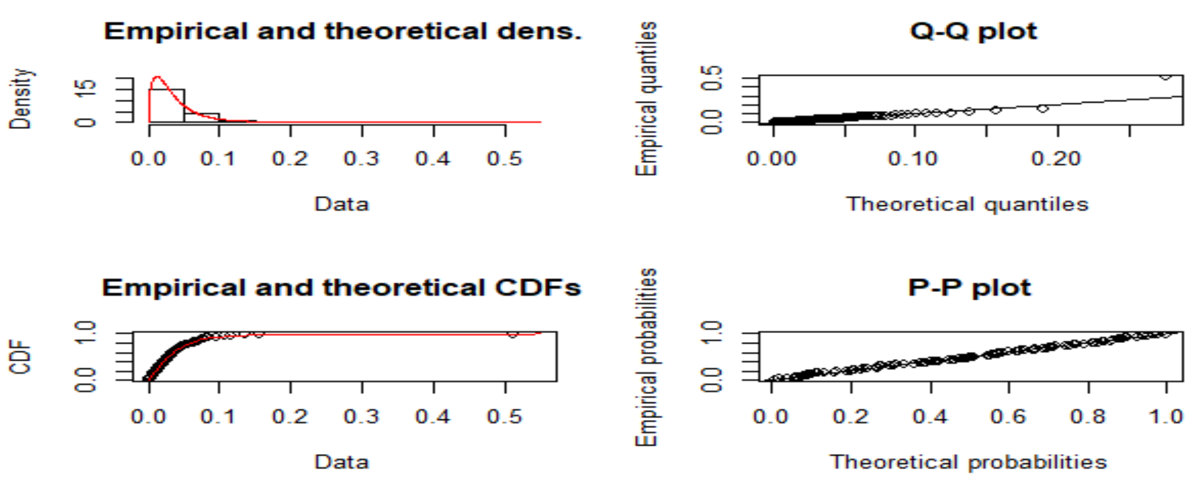


Fig. 11: Diagnostic plots for the Burr losses

b) Burr gains

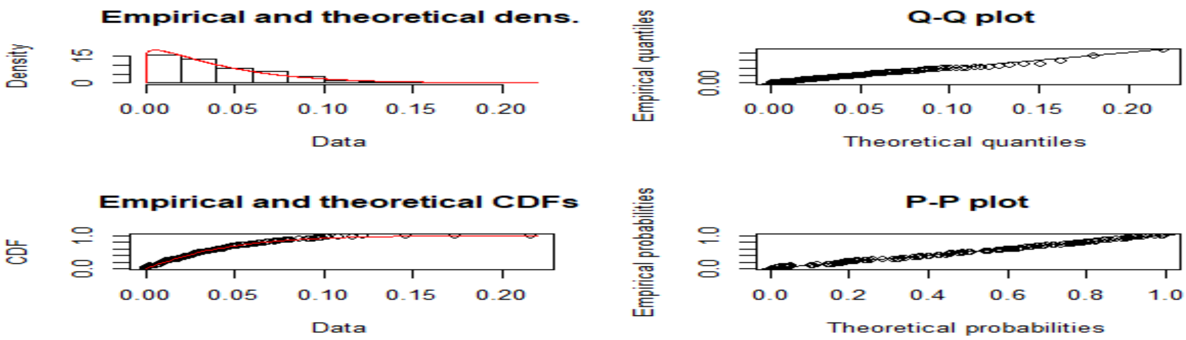


Fig. 12: Diagnostic plots for the Burr gains

On the P-P and Q-Q plots in Figure 11 and Figure 12, there is minimal deviation from the 45°line for most of the probability points. There is insignificant divergence showing that the Burr distribution is a good fit for the data. The estimates and their respective standard errors are given in Table 5.

Table 5: Parameter estimates for the Burr fit

	Parameter Estimate	Standard Error
Estimate of shape(α) for losses	2.267892	1.0568732
Estimate of shape(k) for losses	1.448459	0.1724818
Estimate of scale (β) for losses	0.0573157	0.0123455
Estimate of shape(α) for gains	87.684139	3.0122222
Estimate of shape (k) for gains	1.1087849	0.1583244
Estimate of scale (β) for gains	0.4019639	0.0040196

The parameter estimates and their respective standard errors are given in Table 5.The shape and scale parameters reveal that the all-parent distributions are a good fit to the sample data. The study employed AIC and BIC criteria to determine the best fitting model both for the losses and gains in returns**.**

3.8: BIC and the AIC for the fitted distributions

The BIC and AIC are used to select the best model and the results are shown in Table 6 below.

Table 6: AIC and BIC values for the fitted parent distributions

Distribution	**AIC**	**BIC**
Losses		
Exponential	-479.5058	-476.8054
Weibull	-477.7239	-472.3229
Gamma	-479.5886	-474.1877
Burr	-488.4884	-480.387
Gains		
Exponential	-695.0273	-691.9459
Weibull	-697.039	-690.8762
Gamma	-695.6178	-689.455

Burr	-694.5598	-685.3156

The heavier-tailed Burr distribution gives the best fit for the losses since it gives the lowest AIC and BIC values in Table 6. According to [24], the Burr distribution can be used to model the financial returns data because it can captures the properties of the distribution and the tail-behaviour of the returns data. In Table 6, the Exponential distribution gives the best fit for the gains. Following the principle of parsimony, it is not necessary to add the extra parameters, the one parameter model is sufficient.

Table 7: The VaR and ES are estimated for the gains and losses

Confidence Level	Burr distribution- losses	
	Value at Risk (VaR)	Expected Shortfall
0.95	0.1265	0.1621
0.99	0.1945	0.2348
0.995	0.2237	0.2648
	Exponential distribution-gains	
	Value at Risk (VaR)	Expected Shortfall
0.95	0.1121	0.1463
0.99	0.1628	0.2010
0.995	0.1928	0.2317

In the case of the gains, with a 95 % confidence level, the Burr distribution gives VaR and ES estimates of 12.65 % (0.1265) and 16.21% (0.1621) respectively (Table 7). This means: the expected market gain is not expected go above 12.65 % (0.1265) at the confidence level; if it goes beyond, it will average 16.21% (0.1621). The interpretation is the same for all the other estimates. The two best fitting distributions give a good starting point.

4. Conclusion, Discussion and future possible research

This study investigated the monthly South African Financial Index (J580) returns using four relatively heavy-tailed statistical parent distributions: namely the Exponential, Weibull, Gamma and Burr distributions.

4.1 Conclusion and Discussion

The four relatively heavy-tailed statistical parent distributions are able to model the non-Normal distribution of the Index returns. This approach is consistent with studies by [9], [10], [11], [12], [13] and [14].

The results confirm that the Index losses and gains are non-Normal, following the Burr and Exponential distributions, respectively. The two distributions are used to calculate VaR and ES as financial risk metrics.

Comparing losses to gains for the Index, the results indicate that the prospects of potential losses, modelled by the heavier tailed Burr distribution, are greater than the prospects of potential gains, modelled by the lighter tailed Exponential distribution.

The higher expected losses than the gains in the Index, suggests that a short position (selling the Index today hoping to buy it back at a later date) is a better investment strategy rather than a long position (holding the Index, hoping to sell it on a later date at a higher price).

The study provides additional and valuable information for both investors and practitioners on how to better accurately estimate financial risk using non-Normal statistical distributions.

4.2 Future possible research

Future possible research could include more than the four statistical parent distributions, and comparing the results with those from extreme value distributions.

References

[1] J. H. Stagge, L. M. Tallaksen, L. Gudmundsson, A. F. Van Loon and K. Stahl . Candidate distributions for climatological drought indices (SPI and SPEI). *International Journal of Climatology*, *35*(13), 4027-4040, (2015)

[2] J. Suhaila and A. A Jemain. Fitting daily rainfall amount in Peninsular Malaysia using several types of exponential distributions. *Journal of applied sciences research*, *3*(10), 1027-1036, (2007)

[3] D. L. Fitzgerald . Analysis of extreme rainfall using the log logistic distribution. *Stochastic Environmental Research and Risk Assessment*, *19*(4), 249-257, (2005)

[4] S. Y. Choi and J. H. Yoon. Modelling and risk analysis using parametric distributions with an application in equity-linked securities. *Mathematical Problems in Engineering*, 1-20, (2020)

[5] J. B. McDonald, Probability distributions for financial models. *Handbook of Statistics*, *14*, 427-461, (1996)

[6] Z. Ahmad, E Mahmoudi and M. Alizadeh. Modelling insurance losses using a new beta power transformed family of distributions. *Communications in Statistics-Simulation and Computation*, *51*(8), 4470-4491, (2022)

[7] Z Tondpour and G. H. Yari. The New Burr Distribution with Application to Bladder Cancer Susceptibility. *Reliability Theory and its Applications*, 170, (2016)

[8] A F Jenkinson. The Frequency Distribution of the Annual Maximum (or Minimum) of Meteorological Elements', *Quarterly Journal of the Royal Meteorological Society*, 81(348),158–171, (1955)

[9] A. Z. Afify, A. M Gemeay and N. A. Ibrahim. The heavy-tailed exponential distribution: risk measures, estimation, and application to actuarial data. *Mathematics*, *8*(8), 1276, (2020)

[10] Mabitsela L, Maré E, Kufakunesu R. Quantification of VaR: A note on VaR valuation in the South African equity market. *Journal of Risk and Financial Management* ;8(1), 103-126, (2015).

[11] H Boubaker and N. Sghaier, N. Semiparametric generalized long-memory modelling of some MENA Stock Market returns: A wavelet approach. *Economic Modelling*, *50*, 254-265, (2015)

[12] Q Chen and R. H. Gerlach. The two-sided Weibull distribution and forecasting financial tail risk. *International Journal of Forecasting*, *29*(4), 527-540, (2013)

[13] T Kaizoji and M Kaizoji. Empirical laws of a stock price index and a stochastic model. *Advances in Complex Systems*, *6*(03), 303-312, (2003)

[14] E Bølviken and F. E, Benth . Quantification of risk in Norwegian stocks via the normal inverse Gaussian distribution. In *Proceedings of the AFIR 2000 Colloquium, Tromsø, Norway* (Vol. 8798). AFIR (2000)

[15] P. R. Tadikamalla. A look at the Burr and related distributions. *International Statistical Review/Revue Internationale de Statistique*, 337-344, (1980)

[16] E. Eberlein and U. Keller. Hyperbolic distributions in finance. *Bernoulli*, 281-299, (1995)

[17] E Eberlein. Application of generalized hyperbolic Lévy motions to finance. In *Lévy processes*,319-336, Birkhäuser, Boston, MA, (2001)

[18] W Hu and A. N. Kercheval. The skewed t. In *Econometrics and risk management*. Emerald Group Publishing Limited, (2008)

[19] Das J, Nath DC. Burr distribution as an actuarial risk model and the computation of some of its actuarial quantities related to the probability of ruin. *Journal of mathematical finance*;6(1), 213-231, (2016)

[20] C. O. Omari, S. G. Nyambura and J. M. W Mwangi. Modelling the frequency and severity of auto insurance claims using statistical distributions, (2018)

[21] I. Hoffmann and C. J. Börner. Tail models and the statistical limit of accuracy in risk assessment. *The Journal of Risk Finance*, *21*(3), 201-216, (2020)

[22] E Afuecheta, A. Semeyutin, S Chan, S Nadarajah and D. Andrés Pérez Ruiz . Compound distributions for financial returns. *Plos one*, *15*(10), e0239652, (2020)

[23] M. C. Bryson. Heavy-tailed distributions: properties and tests. *Technometrics*, *16*(1), 61-68, (1974)

[24] A. R. Hakim, I Fithriani and M Novita. Properties of Burr distribution and its application to heavy-tailed survival time data. In *Journal of Physics: Conference Series*, 012016, (2021)

[25] S. Chan, S Nadarajah and E. Afuecheta. An r package for value at risk and expected shortfall. *Communications in Statistics-Simulation and Computation*, *45*(9), 3416-3434, (2016)

[26] R. F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the econometric society*, 987-1007, (1982)

[27] P. Hawkins. South Africa's financial sector ten years on: Performance since democracy. *Development Southern Africa*, *21*(1), 179-204, (2004)

[28] B. Butterworth and S. Malherbe. The South African financial sector: background research for the Seattleround. In *Trade and Industrial Policy Secretariat Annual Forum, Muldersdrift*, (1999)