Thailand Statistician
July 2025; 23(3): 598-614
http://statassoc.or.th
Contributed paper

Forecasting Financial Risk using Statistical Parent Distributions in the South African Industrial Index (J520) Returns

Owen Jakata* and Delson Chikobvu
Department of Mathematical Statistics and Actuarial Science, University of the Free State, Bloemfontein, South Africa.
*Corresponding author; e-mail: owenjakata@rocketmail.com

Received: 8 November 2021
Revised: 2 October 2022
Accepted: 8 January 2023

Abstract

This study investigates suitable models for forecasting financial risk of the monthly South African industrial index (J520) returns. Financial returns are leptokurtic (heavy-tailed) relative to the normal distribution. This study proposes some important alternatives to the normal distribution often used in fitting to financial returns and to forecast financial risk more accurately than suggested by the normal distribution. The study identifies the best-fitting parent distributions to the South African industrial index (J520) returns data and quantifies the financial risk of this index. South Africa is a stable developing country and such information is crucial in developing a diversified portfolio for the international investor inclusive of a developing country's assets. This study uses four relatively heavy-tailed parent distributions, viz: the exponential, Weibull, gamma and the Burr distributions in contrast to past studies, if any, used to describe this particular index returns. The exponential and Weibull distributions are the best fitting parent distributions for the gains and losses respectively. The exponential and the Weibull distributions are in the light tailed Gumbel distribution domain. The study provides the framework on how parent distributions are used to quantify and forecast the financial risk, inclusive of the value-at-risk (VaR) and the expected shortfall (ES) measures. The results reveal that the prospects of potential gains are greater than the prospects of potential losses for one invested in the index. This is useful information for investors who wish to participate in the South African stock market.

Keywords: Value-at-Risk, exponential distribution, Weibull distribution, gamma distribution, Burr distribution.

1. Introduction

Forecasting financial risk is crucial for controlling the risk and expected losses for banks and other organizations. One of the main objectives in finance has been to improve the efficiency and flexibility of financial risk assessments. According Hu and Kercheval (2008), the choice of risk measures is less important for portfolio management than the choice of the appropriate statistical distribution family. The best choice of return distribution is a critical issue for risk management. Statisticians, risk analysts and investors are often faced with two challenges when modelling financial return distributions: selecting appropriate statistical distributions for the returns data, whilst determining how well the selected statistical distributions fit the returns data on one hand, and selecting the appropriate risk measure to quantify the risk. Hoffmann and Börner (2020) stated that in many disciplines, there is often a need to fit a suitable statistical model to existing data to be able to make statements regarding uncertain future outcomes.

The normal or Gaussian distribution is a widespread distribution generally used for modelling stock returns in finance. According McDonald (1996), the normal distribution may provide an adequate representation for some financial returns but other series may not be so conveniently modelled. In general, financial return distributions are not normal, and it is one of the stylized facts of stock returns. Stock returns are generally characterised by skewness, kurtosis, and heavy-tails. A number of different statistical distributions have been used as an alternative to the normal distribution. The various types of distributions can be both symmetric and skewed, and their tails are often heavier than those of the normal distribution. Eberlein and Keller (1995) fitted the generalised hyperbolic distributions to produce a suitable fit to the stock returns. Madan and Seneta (1990) introduced the variance gamma distribution used in option pricing and collateralised debt obligations pricing. Choi and Yoon (2020) fitted 12 distributions, viz: the Cauchy, Laplace, normal, Student’s t, skew normal, skew Cauchy, skew Laplace, skew Student’s t, hyperbolic, normal inverse Gaussian (NIG), variance gamma, and general hyperbolic to four stock indices namely the HSCEI, KOSPI 200, S&P 500, and EURO STOXX 50.

Chen and Gerlach (2013) used a two-parameter Weibull distribution for modelling the conditional financial return distribution for the purposes of forecasting tail-related risk measures. The model was fitted to the daily return series from four international stock market indices, viz: the S&P 500 (US), FTSE 100 (UK), the AORD 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-sided Weibull performed most favourably for conditional VaR forecasting, prior to the crisis as well as during and after it.

According to Choi and Yoon (2020), financial returns contain heavy-tails relative to the normal distribution. Many studies in literature reveal that financial returns provide a rich source of variables with a variety of characteristics, ranging from normally distributed variables to distributions with various degrees of skewness and kurtosis (McDonald 1996). According Ahmad et al. (2020), statistical distributions that are uni-modal and possess thick right tails are more useful in modeling financial losses. To improve on the modelling process, many researchers have been searching for physical and distributional properties of financial returns using empirical data. One area of interest, is the use of parent distributions (other than the normal) in modelling financial returns. Parent distributions concentrate their fit where the bulk of the data is located. This is around the mean, mode or median. It is generally accepted and assumed that financial return variables follow certain assumed statistical distributions.

This study uses the exponential, Weibull, gamma and Burr distributions in modelling index returns since financial returns are expected to be heavy-tailed. The chosen distributions do cater for the relatively heavy-tails and are able to capture various degrees of skewness and kurtosis. This study investigates the performance of four relatively heavy-tailed parent distributions in fitting to the South African industrial index (J520) returns. The relatively more heavy-tailed Burr distribution was included in the four proposed distributions of this study as it is able to cater for even heavier tails. According to Yari and Tondpour (2017) many standard theoretical distributions, including the Weibull, exponential, logistic, generalized logistic, Gompertz, normal, extreme value, and uniform distributions, are special cases or limiting cases of the Burr distribution. The authors are of the opinion that the four distributions are sufficient for this study to avoid clogging the work.

The main objective of this study, is to determine the best fitting parent statistical distribution that is appropriate for describing and quantifying the Index return distribution. This study proposes some important alternatives to the normal distribution which can be used to fit financial returns and to quantify and forecast financial risk. A few studies on parent distributions have been carried out in the context of the South African stock market. This study covers this gap and determines the best-fit parent distribution suitable for modelling the South African industrial index (J520) returns data and to forecast the risk level of the Index. The left (losses) and the right (gains) tail of the monthly South African industrial index (J520) returns data are considered separately.

1.1. Statement of the Problem

Statisticians, risk analysts and investors have been interested in modelling financial return distributions for many years. One of the reasons for the interest, is that the financial return distributions can be modelled using parent distributions, which can be used to forecast the riskiness of the financial returns. The normal distribution may provide an adequate representation for some financial returns but other series may not be so conveniently modelled. The relative performance of four proposed parent distributions is investigated using the monthly financial returns of the index for the purposes of choosing the most suitable parent distribution to fit the Index returns data. The best fitting models for the negatives (losses) and positive (gains) returns are used to forecast risk and can be used as a good starting point for developing better models for risk management.

1.2. Justification of the Study

The industrial sector is viewed as a fundamental activity in the economy of any country as it brings economic growth and capital formation. It is a leading sector for economic development. In terms of market capitalisation, the industrial sector is the most represented sector on the South African stock market. Therefore, it is important to model the returns from the industrial sector, so as to quantify and forecast the risk level of the index. The forecasted quantitative risk indicators have practical applications in financial risk management.

1.3 Objectives of the Study

The main objective of the study is to compare the relative performance of the exponential, Weibull, gamma and Burr distributions in fitting the financial returns of the South African industrial index (J520) index.

The specific objectives are to:

- Determine the best fitting models for the negative returns (losses) and positive returns (gains) of the Index using the AIC and BIC criteria and other criteria.
- Forecasting risk associated with the index using risk measures of VaR and ES.

The study provides the framework and some useful information for investors on how statistical distribution models are used to quantify and forecast the risk level of a stock/index. This will assist investors, who are considering to invest in the index, to fully understand the risk and returns associated with the index, which in turn helps them to determine the amount of capital needed to be set aside to meet regulatory requirements.

The main contribution of this study lies in the identification of the most suitable parent statistical distribution for the index returns, and more accurate VaR and ES estimates. All the information is proxy for future risk. The quantitative risk indicators also provide useful information to policy makers deciding to take measures to promote economic and market growth. This study is organised as follows: Section 2 presents a review of literature, Section 3 presents the research methodology, and Section 4 presents results and the discussion. Section 5 gives the conclusion and research limitations/implications.

2. Review of Literature

This section gives an overview of some of the articles in literature on modelling return distribution data with a brief discussion on the fitted models using the parent distributions. There are many studies that have been done on parent distributions including some discussed in this section.

Kaizoji and Kaizoji (2003) analysed the daily Nikkei 225 Index returns for the period 1984-2002. The research investigated if the exponential distribution was suitable in modelling the distributions of returns, volatility and calm-time interval distribution of the volatility. They applied a linear regression model fit to both the positive (gains) and negative (losses) returns to estimate the parameters of the exponential distribution. A graphical diagnostic semi-log plot was used to assess the goodness-of-fit. The exponential distribution was suitable in all three situations.

Chae et al. (2006) studied how the Korean composite stock price index return distribution has evolved for the period 1995 to 2003. The Exponential distribution best fitted the Index returns before the Asian crisis of 1997. However, with time, the distribution of the returns became narrower than an exponential distribution.

Mittnik and Rachev (1993) compared the Weibull distribution to the Laplace, Stable Paretian, Max-Stable, Min-Stable, Geometric Stable and the log-normal distributions by fitting to the S&P 500 daily stock index returns for the period 1982-1986. The maximum likelihood estimation (MLE) method was used to estimate the parameters. The Kolmogorov-Smirnov goodness of fit test was used to compare the models. The Weibull distribution outperformed the Stable Paretian in all the cases considered.

Ivanov (2022) used monetary data from the bank and applied the variance-gamma distribution with stochastic linear drift coefficient in order to calculate the basic monetary risk measures using monetary data. The VaR was prescribed to financial institutions for the estimation of the portfolio losses by the regulations Basel I and Basel II. The researcher obtained the distribution function, the probability density function and the lower partial expectation for the considered process in closed forms. The results were successfully applied to the estimation of the VaR and the ES of an investment portfolio.

Das and Nath (2016) used an algorithm to fit the Burr XII distribution to a set of insurance data. The findings were that, the probability of ultimate ruin was obtained as a solution to an integro-differential equation and in this case, the claim severity was distributed as Burr XII distribution. The equation was solved numerically to obtain an approximation to the probability of ultimate ruin.

Omari et al. (2018) used nine proposed parent distributions to model insurance claims. The distributions are namely: Poisson, geometric, negative binomial, Pareto, Weibull, exponential, and gamma, log-logistic, and log-normal distributions. MLE method was used to estimate the parameters. Their findings showed that the lognormal distribution gave the best fit distribution to model the claim size. The negative binomial and geometric distributions gave the best fit distributions for claim frequency data compared to other statistical distributions.

In their study, Afuecheta et al. (2020) proposed six Student-t-based parent distributions, viz: the half normal, Fréchet, Lomax, Burr III, inverse gamma and the generalized gamma innovations. GARCH models are fitted to the data using the maximum likelihood method. Six financial return series were used: S&P 500, Dow Jones Industrial Average (DJI), Ultra-Low Sulphur (Diesel), Texas

Propane (Propane), Bitcon (BTC) and Litecoin (LTC). The study revealed that the generalized gamma innovations gave the best-fit distribution and the Burr innovations gave the second best-fit distribution, the Frétchet innovations gave the third best-fit distribution, the inverse gamma innovations gave the fourth best-fit distribution, the AST innovations give the fifth best-fit distribution, the Lomax innovations give the sixth best-fit distribution and the GHYP innovations gave the seventh best-fit distribution.

Mabitsela et al. (2015) employed the NIG distribution for vector auto regressive (VAR) valuation in the South African and USA stock market using the FTSE/JSE40 and S&P 500 indices, respectively. The researchers compared the NIG distribution with the normal distribution, skew Student's t-distribution and the Student-t-distribution, each capturing different features of the financial returns. The skew Student's t and the Student-t-distributions had heavier tails than the NIG distribution which had semi-heavy tails. The three distributions gave a better fit than the normal distribution.

This study differs from these other studies in that it adopts the exponential, Weibull, gamma, Burr distributions in modelling the return distribution and forecasting the risk level of the of the South African industrial index (J520).

3. Research Models

This study applies the parent distributions approach for modelling non-normal distributions in the context of the South African stock market. Four parent distributions: the exponential, Weibull, gamma and Burr distributions since financial returns are expected to be heavy detailed. The chosen distributions do cater for the relatively heavy-tails and are able to capture various degrees of skewness and kurtosis. These distributions are commonly used in financial return distribution analysis, but not yet on the South African industrial index returns. The probability distribution functions along with their parameter estimations and their respective properties are discussed in this section. Losses and gains for the Industrial index are analysed separately. The negative returns are converted into positive losses by multiplying by a negative one.

3.1. Exponential distribution

Bryson (1974) described a fat-tailed distribution as having a tail that is fatter than an exponential distribution. The exponential distribution gives a good starting point relative to our presumption on the nature of our data. According to Omari et al. (2018), the probability density function (PDF) and cumulative distribution function (CDF) of the exponential distribution are respectively denoted as:

$$f(x;y) = \lambda e^{-\lambda x},$$

$$F(x;y) = 1 - e^{-\lambda x},$$

where x represents the log returns and $\lambda > 0$ is the rate parameter. The exponential parameters are estimated by the MLE method.

3.2. Weibull distribution

The Weibull distribution adds a shape parameter to the exponential distribution, hence making it more flexible. According to Nielsen (2011), the PDF and CDF of the Weibull distribution are respectively denoted as:

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

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

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

3.3. Gamma distribution

The gamma distribution is another two-parameter distribution from the exponential family of distributions. According to Omari et al. (2018), the PDF and CDF of the gamma distribution are respectively denoted as:

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

$$F(x;\alpha,\beta)=\frac{1}{\beta^{\alpha}\Gamma(\alpha)}\int_{0}^{x}t^{a-1}e^{\frac{t}{\beta}}dt,$$

with the continuous random variable x represents the log returns, α and β represents the shape and scale parameters, respectively. The gamma parameters are estimated by the MLE method.

3.4. Burr distribution

According to Hakim et al. (2021), 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 fat-tailed distribution. The additional parameter makes the distribution more flexible and gives a better fit if the log returns data is fat-tailed.

According to Tadikamalla (1980), the PDF and CDF of the Burr distribution are respectively denoted as:

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

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

The Burr distribution has two shape parameters, $\alpha>0,$ $k>0$ and $\beta>0$ represents the scale parameter. The Burr parameters as estimated by the MLE method.

3.5. Risk Measures

In this section, the models for the VaR and the ES for the proposed parent distributions are presented.

Exponential distribution VaR and ES equations (Chan et al. 2016)

$$VaR_{p}(x)=-\frac{1}{\lambda}\log(1-p),$$

$$ES_{p}(x)=-\frac{1}{p\lambda}(\log(1-p)p-p-\log(1-p)),$$

for $x>0,$ $0<p<1,$ and $\lambda>0,$ the scale parameter.

Weibull distribution VaR and ES equations (Chan et al. 2016)

$$VaR_{p}(x)=\lambda[-\log(1-p)]^{\frac{1}{k}},$$

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

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

Gamma distribution VaR and ES equations (Chan et al. 2016)

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

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

for $x > 0,\ 0 < p < 1,\ \beta > 0,$ 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 (Chan et al. 2016)

$$VaR_p(x) = [(1-p)^{\frac{1}{k}} - 1]^{\frac{1}{\alpha}},$$

$$ES_p(x) = \frac{1}{p}\int_0^p [(1-p)^{\frac{1}{k}} - 1]^{\frac{1}{\alpha}} dv,$$

for $x > 0,\ 0 < p < 1,\ \alpha > 0,$ the first shape parameter, and $k > 0,$ the second shape parameter. The VaR and ES for the best fitting distributions: exponential distributions and Weibull distributions are forecasted using the VaRES R-statistical software package.

3.6. Test for normality, stationarity, heteroscedasticity and auto-correlation

To test for normality of the monthly South African industrial index (J520) return distribution, the Anderson-Darling and the QQ graphical plot are used. Stationarity tests are applied to determine whether the data is stationary. The augmented Dickey-Fuller (ADF) test is used to test for stationarity of the monthly South African industrial index (J520) returns. The ARCH LM test is used to test for the presence of heteroscedasticity in the monthly South African industrial index (J520) return series. The auto-correlation function (ACF) and the partial auto-correlation function (PACF) method and Box-Ljung test are used to test for auto-correlation of the return distribution.

4. Results and Discussion

This study compares the relative performance of the four proposed distributions, quantifying and forecasting the risk level associated with the best-fit distribution for the losses and the gains.

4.1. Software used and research data

Data analysis was done using different packages that deal with statistical distributions in the R programming environment: fitdistrplus, ReIns, actuar and VaRES (Chan et al. 2016). VaRES is the statistical package which is used to forecast the VaR and ES of the best fitting distributions. The monthly South African Industrial Index (J520) returns (years 1995-2018) are used in this study and are obtained from the website iress expert: https://expert.inetbfa.com (with permission). For this Index, it consists of industrial firms listed on the South African stock market. The industrial firms are from the following sub-categories, construction and materials (J235), aerospace and defence (J271), general industrials (J272), electronic and electrical equipment (J273), industrial engineering (J275), commercial vehicles and trucks industrial transportation (J277) and support services (J279). The industrial sector is the largest in terms of market capitalisation and it measures the overall performance for this sector Therefore, it is of critical importance to forecast the riskiness of this index.

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

$$x_t^- = -r_t^- = -\ln\frac{M_t}{M_{t-1}},$$

r_t are the monthly log returns for the month t, M_t represents the monthly index in month t and ln represents the natural logarithm (Velasco and Lapuz 2018). When using a loss function, the losses (minimum returns) are positive.

4.2. Exploratory data analysis

A preliminary analysis was done prior to fitting the parent distributions to investigate the characteristics of the of the monthly South African industrial index (J520) returns.

Table 1 Descriptive statistics

Description	Values of r_t	Description	Values for r_t
Mean	–0.009366	Variance	0.003302
Median	–0.010478	Standard Deviation	0.057467
Maximum	0.328471	Skewness	1.016932
Minimum	–0.140273	Kurtosis	4.420852

Table 1 gives descriptive statistics for r_t. The table shows that monthly South African Industrial Index (J520) returns have a positive skew (1.016932 > 0) and a kurtosis greater than 3 (4.420852 > 3). It shows the distribution of returns has a fat-tails. The descriptive statistics show that the data is asymmetrical, right-skewed and fat-tailed. This confirms the widely expressed opinions that financial returns data have fat-tails.

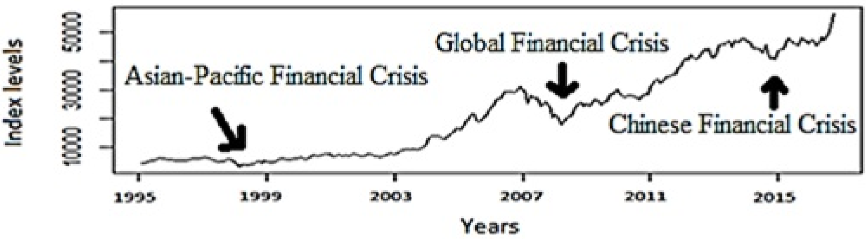

Figure 1 Time series plot of the monthly South African Industrial Index (J520), M_t

The time series plot of the monthly South African Industrial Index (J520) levels (Figure 1) shows a significant increase from 1995 to 2018 with a lot of oscillating fluctuations, and the graph of monthly log return data (r_t) (Figure 2) confirms the volatility of the South Africa stock market. The Asian financial crisis (1997 to 1998), the global financial crisis (2007 to 2008) and the Chinese stock market crash which are indicated in Figure 1, had a negative impact on the South African stock market which is confirmed by sharp down turns of the index levels. Global stock markets have become increasingly involved in global trade; this results in high volatility of the Index. Therefore, it is important to model the returns from the industrial sector in order to describe the riskiness of the index.

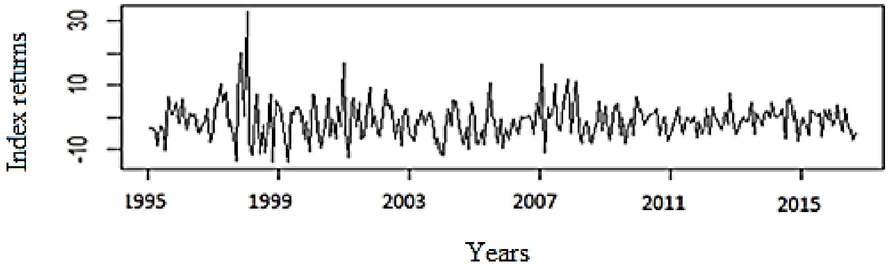

Figure 2 Time series plot of monthly returns for the South African Industrial Index (J520), r_t

4.3. Tests for stationarity, normality, heteroscedasticity and autocorrelation

4.3.1 Test for stationarity

The following results were obtained for the ADF Test: Dickey-Fuller = –6.7391, Lag order = 6, p-value = 0.01. The results show that a p-value < 0.05 was obtained: therefore, it is concluded that the returns data is stationary.

4.3.2 Test for normality

To test for normality, two methods are used: i) the Andersen-Darling test and ii) QQ plot for the monthly log returns.

i) Andersen-Darling Test

To test for normality under the Andersen-Darling normality test shows a p-value of 0.01 which implies that the monthly data series is not normally distributed and that the returns data has fat-tails.

ii) QQ plot for the monthly log returns

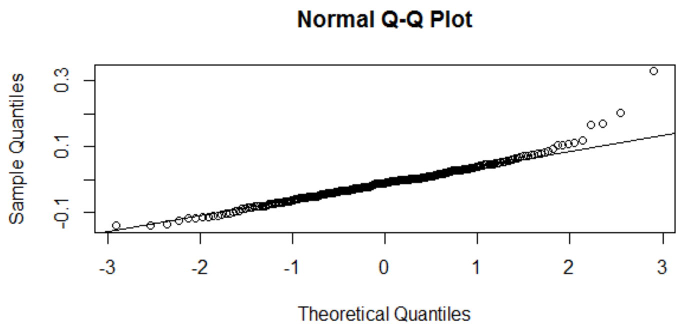

Figure 3 The QQ plot for r_t, the monthly South African industrial index (J520) returns

In Figure 3, some points are above the straight line which implies that the positive returns (gains) are not normally distributed, suggesting that the data has fat-tails.

4.3.3 Test for heteroscedasticity

The ARCH LM test indicates that there are no significant ARCH effects that exist in the returns data (χ^2 = 8.366974, df = 12, p-value = 0.7558355). This implies that the heteroscedasticity is insignificant and therefore the unconditional (static) models are adopted in this study.

4.3.4 Test for autocorrelation

Two autocorrelation tests were carried out using two methods:

i) The ACF and PACF

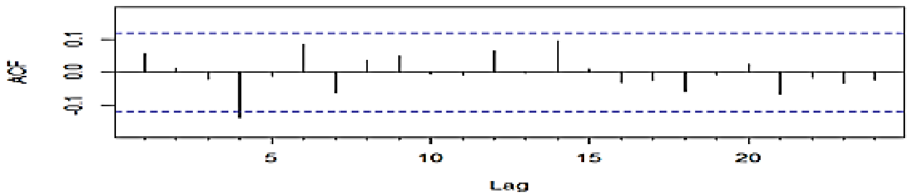

Figure 4 ACF diagram for r_t

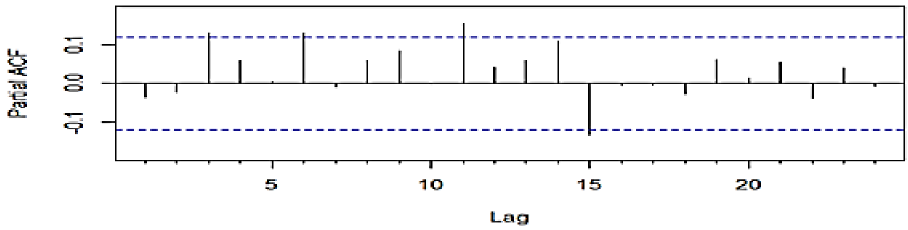

Figure 5 PACF Diagram for r_t

In Figures 3 and 4, the ACF and PACF diagrams reveal that there are no significant autocorrelations in the returns data.

ii) The Box-Ljung test for autocorrelation ad the results showed that: χ^2 = 0.8806, df = 1, p-value = 0.348.

The p-value > 0.05 is obtained which indicates that the autocorrelation is insignificant at this level of significance. This confirms that the return distribution is independently distributed.

4.4. The histogram of monthly South African Industrial Index (J520) log returns data.

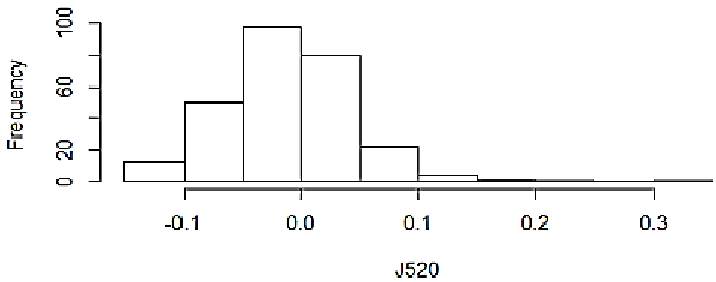

Figure 6 The histogram of r_t the monthly South African industrial index (J520) returns

The histogram in Figure 6 reveals that returns data is relatively asymmetrical about the mean. This is consistent with the positive skewness value obtained. The large kurtosis value obtained implies there exist some high peaked observations in the returns data, implying that the data has fat-tails.

4.5. Fitting and selecting the best-fit distribution

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

4.5.1 Exponential distribution

Figures 7 and 8 show the exponential fit to the South African industrial index returns.

a) Exponential gains

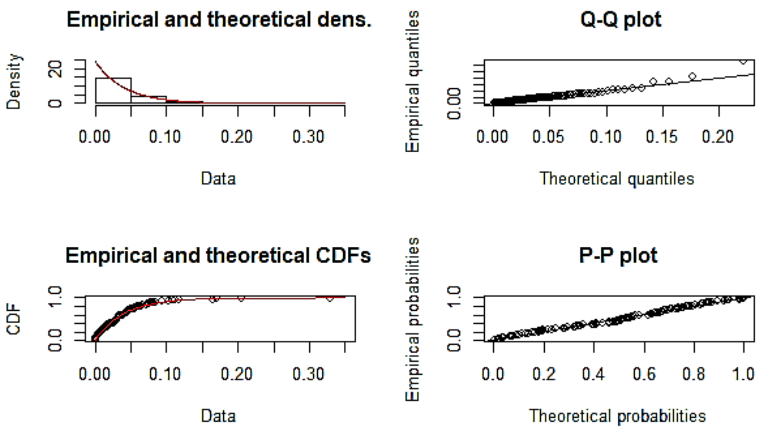

Figure 7 Diagnostic plots for the exponential gains

b) Exponential losses

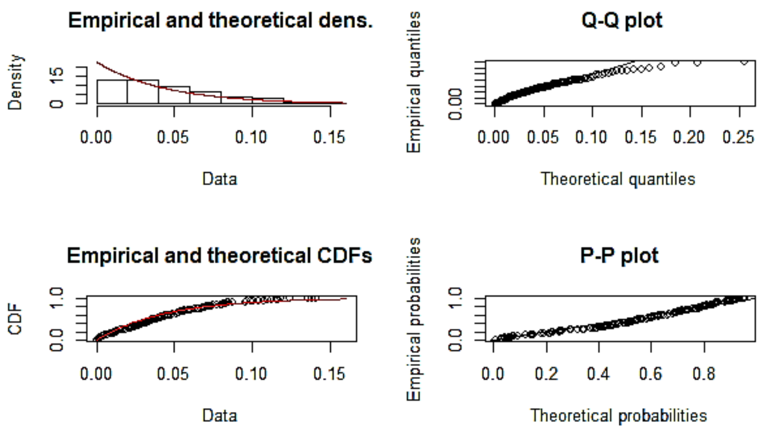

Figure 8 Diagnostic plots for the exponential losses

The data points exhibit some linearity on the P-P plot and slight deviation from the 45°line in Figures 7 and 8. From the diagnostic plots, it shows that the exponential distribution is a moderate fit for the data. The Maximum Likelihood method was used to estimate parameters for both the losses and the gains. The results revealed that the Exponential distribution rate parameters for the gains and losses are 0.0412 and 0.0445, respectively. The rate parameters for the gains and the losses are approximately 0.04 which implies that the parent distribution is a good fit to the sample data.

4.5.2 Weibull distribution

Figures 9 and 10 show the Weibull fit to the South African industrial index returns.

a) Weibull gains

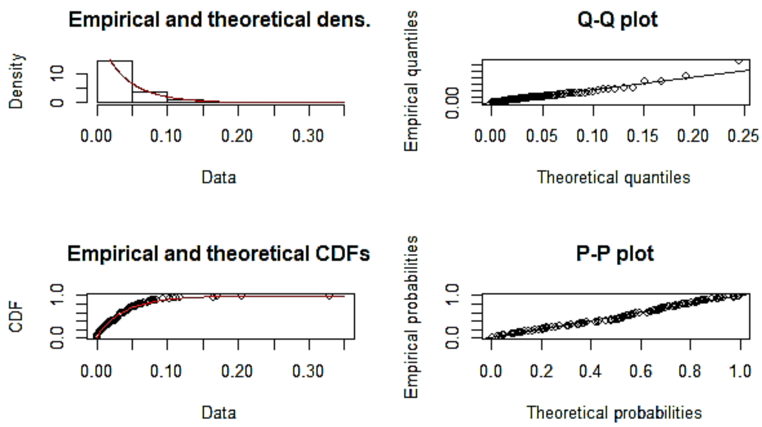

Figure 9 Diagnostic plots for the Weibull gains

b) Weibull losses

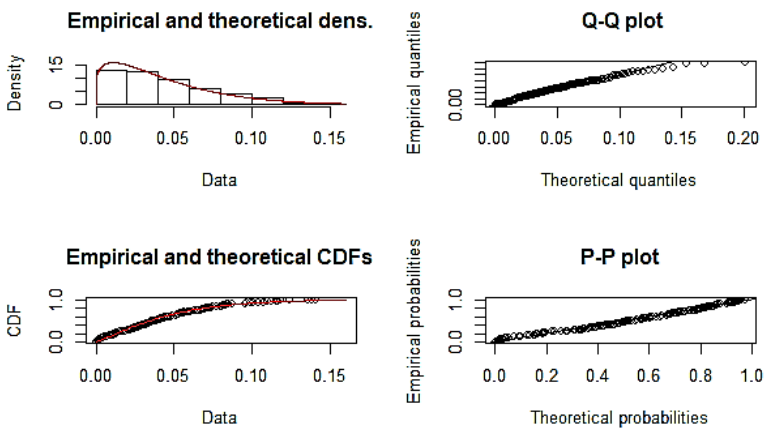

Figure 10 Diagnostic plots for the Weibull losses

On the PP plot in Figures 9 and 10, there is minimal deviation from the 45° line. Most of the probabilities match therefore indicating that the Weibull is a good fit for the data. The density, QQ, CDF plots confirm the goodness of fit. The two parameters are estimated for both the gains and losses using the MLE method. The results revealed that the Weibull distribution shape and scale parameters for the gains are 0.9299 and 0.0399, respectively. The results also showed that the shape and scale parameters for the losses are 1.2059 and 0.0470, respectively. The shape and scale parameters reveal that the parent distributions are a good fit to the sample data since they are approximately 1 and 0.04, respectively.

4.5.3 Gamma distribution

Figures 11 and 12 show the gamma fit to the South African industrial index returns.

a) Gamma gains

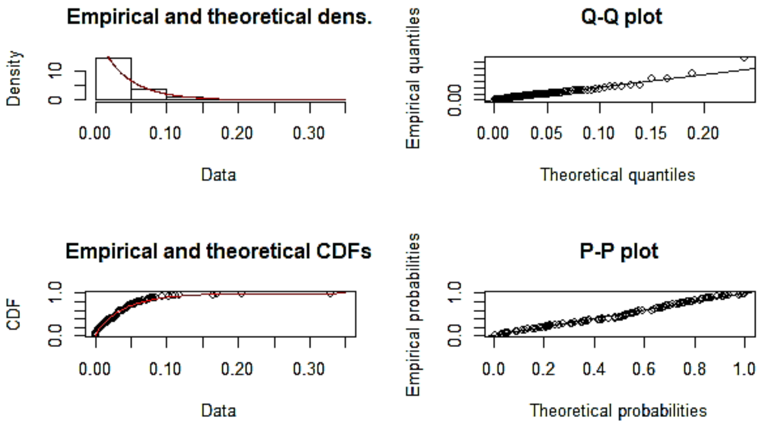

Figure 11 Diagnostic plots for the gamma gains

b) Gamma losses

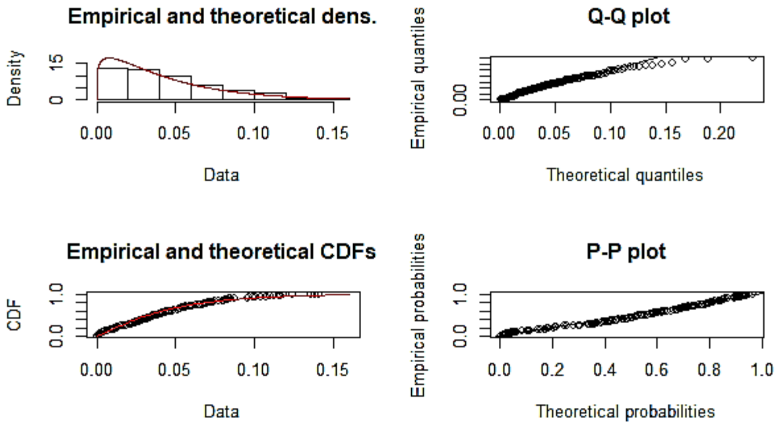

Figure 12 Diagnostic plots for the gamma losses

In Figures 11 and 12, on the PP plots, there is a marginal deviation from the reference line with the probabilities matching for most of the sample points for both the losses and gains. There is a strong indication of linearity. The PP plot corroborated by the QQ plot shows that the gamma distribution is a good fit for both the losses and gains. The parameters were estimated using the method of maximum likelihood estimate. The results revealed that the Gamma distribution shape and scale parameters for the gains are 0.0882 and 0.0467, respectively. The results also revealed that the shape and scale parameters for the losses are 1.2166 and 0.0365, respectively. The shape and scale parameters reveal that the parent distributions are a good fit to the sample data.

4.5.4 Burr distribution

Figures 13 and 14 show the gamma fit to the South African industrial index returns.

a) Burr gains

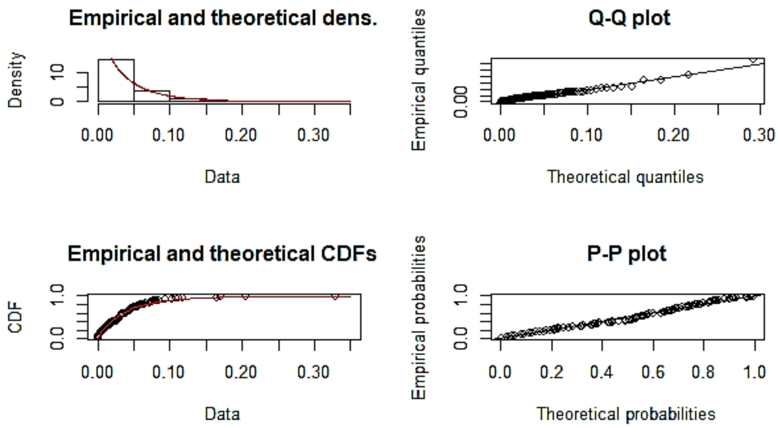

Figure 13 Diagnostic plots for the Burr gains

b) Burr losses

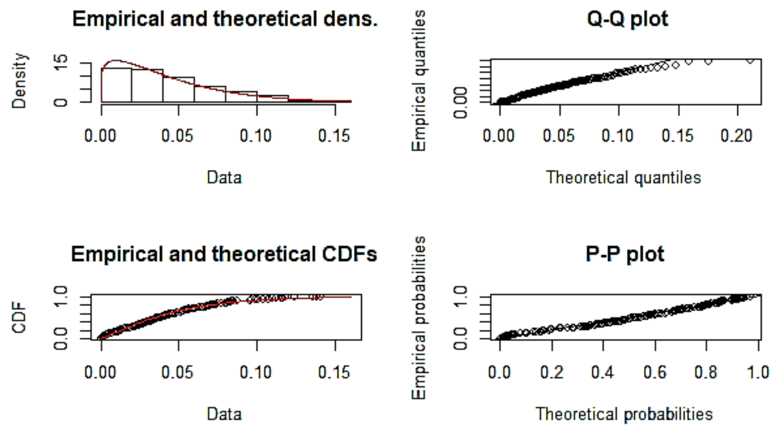

Figure 14 Diagnostic plots for the Burr losses

In the PP plot in Figures 13 and 14, there is minimal deviation from the 45° line for most of the probability points. Most of the probabilities match. There is slight but insignificant divergence showing that the Burr distribution is a good fit for the data. The QQ plot and other plots confirm the same. Parameters were estimated using the method of maximum likelihood estimate.

The results revealed that the Burr distribution shape 1, shape 2 and scale parameters for the gains are 28.6937, 0.9526 and 0.7592, respectively. The results also revealed that the shape 1, shape 2 and scale parameters for the losses are 96.1062, 1.1821 and 0.4474, respectively. The shape and scale parameters that are approximately between 1.00 and 0.40, respectively, which reveals that the 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 negative (losses) and positive (gains) returns.

4.5.5 Summary of the BIC and the AIC estimates for the fitted distributions

A comparison of the AIC and BIC for all distributions fitted on the gains and losses are shown in Table 2. The minimum AIC/BIC denotes the best model.

Table 2 AIC and BIC values for the fitted parent distributions

Distribution	Gains		Losses	
	AIC	BIC	AIC	BIC
Exponential	–483.8899	–481.1796	–674.1782	–671.1030
Weibull	–482.8953	–477.4763	–679.7998	–673.6494
Gamma	–483.0689	–477.6499	–675.8123	–669.6620
Burr	–480.9770	–472.8484	–677.2161	–667.9905

The light-tailed exponential distribution gives the best fit for the gains of the Industrial Index when compared to other distributions. The Weibull distributions gives the best fit for losses when compared to the other distributions. This means that both the exponential and Weibull distributions are good models for the gains and the losses, respectively. The exponential and the Weibull distributions are in the Gumbel domain of attraction distributions. The maxima (extreme returns) for the exponential and Weibull parent distributions will converge to the Gumbel domain class distribution. The distributions give better fits than the distribution in the Frétchet-Pareto domain (the heavy-tailed Burr distribution) indicating that the gains and the losses are somewhat lighter-tailed.

4.6. Forecasting VaR and ES for the best fit distributions

The VaR and ES for the exponential and Weibull distributions are forecasted by the VaRES R-statistical software package. The results are shown in Table 3.

Table 3 The VaR and ES are estimated for the gains and losses

Confidence level	Exponential distribution for gains		Weibull distribution for losses	
	VaR	ES	VaR	ES
95.0	0.1235016	0.15822737	0.1168371	0.15557605
99.0	0.1898521	0.22916026	0.1668908	0.20954817
99.5	0.2184277	0.25855593	0.1874674	0.23079504

At higher quantiles, the gains, with a 99.5 % confidence level, the exponential distribution gives VaR and ES estimates of 21.84% (0.2184277) and 25.86% (0.25855593), respectively (Table 3). This means: the expected market gain is not expected go above 21.84 % (0.2184277); if it goes beyond, it will average 25.86% (0.25855593). The interpretation is the same for all the other estimates. The results also reveal that for an investment in the South African industrial index (J520), the prospects of potential gains are greater than the prospects of potential losses.

4.7. Discussion

This study applied four relatively heavy-tailed parent distributions, viz: the exponential, Weibull, gamma and Burr distributions to describe financial risk in the South African industrial index (J520) returns. The ability of the four parent distributions to model the non-normal distribution of the the South African industrial index (J520) returns is consistent with the studies by Kaizoji and Kaizoji (2003), Chae et al. (2006), Mittnik and Rachev (1993), Ivanov (2022), Das and Nath (2016), Omari et al. (2018), Afuecheta et al. (2020) and Mabitsela et al. (2015). The findings show that the index gains and losses are non-normal and follow the exponential and the Weibull distributions respectively. The two distributions both provided adequate modelling for the gains and losses, and are then used to estimate VaR ad ES as financial risk measures. The findings of this paper have important implications for investment, diversification and hedging decisions for both local and international investors. The model framework contained in this study is of use to academic researchers, investors, managers and policymakers who would want to quantify and forecast the risk of investing. The full distribution modeling using parent distributions does not fully consider the tails of the distribution therefore they tend to underestimate risk measures. This gives a good starting point for developing better models that represent an adequate and reliable framework for risk forecasting and management. Further investigations may include the estimation of financial risk using the extreme value distributions and other parent distributions and compare with the current findings.

5. Summary, Conclusion and Future Possible Research

This section summaries the research, and concludes. Future possible research is also discussed.

5.1. Summary and Conclusion

This study set out to investigate the alternative statistical distributions which can be used to model financial return distributions. The study had two objectives: to determine the best-fit distribution and to quantify financial risk associated with the index. The study analysed the monthly South African industrial index (J520) returns using four relatively fat-tailed parent distributions, viz: the exponential, Weibull, gamma and Burr distributions. The findings of this study provide important implications for both local and international investors, risk managers, regulators, policymakers and researchers. The exponential distribution (also described by Bryson (1974)) was the best fitting parent distribution for the gains. Therefore, it is the appropriate model for the relatively light-tailed South African industrial index (J520) gains. For the negative returns (losses), the Weibull distribution gave the best fit when compared to the other distributions. The maxima for the two best-fit parent distributions will converge to the Gumbel domain (light tailed) class distribution.

The Exponential and the Weibull distributions were used to describe the risk level and the results revealed that: the prospect of potential gains is greater than the prospect of potential losses for one invested in the South African Industrial Index (J520). The ability to select the best-fit statistical distribution for the financial returns data gives individual investors, corporate planners, and government policy maker's opportunities to forecast the risk level accurately when making their investment decisions. This study provides the framework and some useful information to local and international investors seeking a well-diversified portfolio inclusive of developing countries assets in their portfolio. This implies that the study will provide a framework that can be used to determine the correct statistical distribution models that are used to describe the financial risk level.

5.2. Future possible research

Currently, the research is limited to modelling the main body of the financial returns, however for future possible research, one might consider fitting the more specialised extreme value models to the tails of the positive and negative returns, since parent distributions tend to underestimate the risk at the tails. The EVT allows us to quantify the stochastic behaviour of an event found in the right and the left tails.

Acknowledgements

The authors would also like thank the Department of Mathematical Statistics and Actuarial Science at the University of the Free State in south Africa for supporting this research.

References

Afuecheta E, Semeyutin A, Chan S, Nadarajah S, Andrés Pérez Ruiz D. Compound distributions for financial returns. PLoS One. 2020; 15(10): 1-25.

Ahmad Z, Mahmoudi E, Alizadeh M. Modelling insurance losses using a new beta power transformed family of distributions. Commun Stat Simul Comput. 2020; 51(8): 4470-4491.

Bryson MC. Heavy-tailed distributions: properties and tests. Technometrics. 1974; 16(1): 61-68.

Chae S, Jung W, Yang J, Moon H. Temporal evolution of the return distribution in the Korean stock market. J Korean Phys Soc. 2006; 48(2): 313-317.

Chan S, Nadarajah S, Afuecheta E. An R package for value at risk and expected shortfall. Commun Stat Simul Comput. 2016; 45(9): 3416-3434.

Chen Q, Gerlach RH. The two-sided Weibull distribution and forecasting financial tail risk. Int J Forecast. 2013; 29(4): 527-540.

Choi SY, Yoon JH. Modelling and risk analysis using parametric distributions with an application in equity-linked securities. Math Prob Eng. 2020; 2020(1): 1-20.

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. J Math Finance. 2016; 6(1): 213-231.

Eberlein E, Keller U. Hyperbolic distributions in finance. Bernoulli. 1995; 1(3): 281-299.

Hakim AR, Fithriani I, Novita M. Properties of Burr distribution and its application to heavy-tailed survival time data. J Phys Conf Ser. 2021; 1725(1): 1-9.

Hoffmann I, Börner CJ. Tail models and the statistical limit of accuracy in risk assessment. J Risk Finance. 2020; 21(3): 201-216.

Hu W, Kercheval AN. The skewed t distribution for portfolio credit risk. Adv Econometrics. 2008; 22: 55-83.

Ivanov RV. The risk measurement under the variance-gamma process with drift switching. J Risk Financ Manag. 2022; 15(1): 1-22.

Kaizoji T, Kaizoji M. Empirical laws of a stock price index and a stochastic model. Adv Complex Syst. 2003; 6(3): 303-312.

Mabitsela L, Maré E, Kufakunesu R. Quantification of VaR: a note on VaR valuation in the South African equity market. J Risk Financ Manag. 2015; 8(1): 103-126.

Madan DB, Seneta E. The variance Gamma (V.G.) model for share market returns. J Bus. 1990; 63(4): 511-524.

McDonald JB. Probability distributions for financial models. Handb Stat. 1996; 14: 427-461.

Mittnik S, Rachev ST. Modelling asset returns with alternative stable distributions. Econom Rev. 1993; 12(3): 261-330.

Nielsen MA. Parameter estimation for the two-parameter Weibull distribution. MSc [dissertation], Provo (UTAH): Brigham Young University; 2011.

Omari CO, Nyambura SG, Mwangi JMW. Modelling the frequency and severity of auto insurance claims using statistical distributions. J Math Finance. 2018; 8(1): 137-160.

Tadikamalla PR. A look at the Burr and related distributions. Int Stat Rev. 1980; 48(3): 337-344.

Velasco A, Lapuz D. Extreme value modelling for measuring financial risk with application to selected Philippine stocks. J Comput Appl Math. 2018; 7(3): 1-13.

Yari G, Tondpour Z. The new Burr distribution and its application. Math Sci. 2017; 11(1): 47-54.