Strathmore
UNIVERSITY

A Statistical Analysis of the Log Returns of Cryptocurrencies

Ndegwa Kevin Irungu - 110189

Submitted in partial fulfilment of the requirements for the Degree of Bachelor of Business Science in Actuarial Science at Strathmore University

Strathmore Institute of Mathematical Science
Strathmore University
Nairobi, Kenya

March 2022

DECLARATION

I declare that this work has not been previously submitted and approved for the award of a degree by this or any other University. To the best of my knowledge and belief, the Research Project contains no material previously published or written by another person except where due reference is made in the Research Project itself.

Ndegwa Kevin Irungu

............KevN.................................... [*Signature*]

............04/03/2022............................. [*Date*]

This Research Project has been submitted for examination with my approval as the Supervisor.

Kevin Otieno

.....KEVIN OTIENO.............................. [*Signature*]

...................07/03/22......................... [*Date*]

Strathmore Institute of Mathematical Sciences

Strathmore University

Abstract

There has been an increase in interest and demand for cryptocurrencies and thus understanding their statistical properties is important for it implies their risk. Understanding the risk involved in investing in the cryptocurrencies allows one to evaluate the same risk against their own risk tolerance and thus determine whether it is worthwhile to venture into cryptocurrencies and if so, the optimal weight of the investment in the portfolio. This study seeks to find the statistical distribution from a family of fat tailed distributions that best explains the log returns of cryptocurrencies. It was conducted in Nairobi between May 2021 and February 2022. The data used was obtained from Yahoo Finance. The results suggested that the Generalized Hyperbolic Distribution gives the best fit for the large cryptocurrencies ranked by market capitalization.

Acknowledgements

I would like to thank God for the grace to push through with this research.

I acknowledge the support received from family and friends.

I finally acknowledge my supervisor, Kevin Otieno, for his guidance. His great knowledge and experience in statistics mixed with his advice and friendliness played a big role in the completion of this research.

Table of Contents

CHAPTER 1 : INTRODUCTION

Background of the Study

Cryptocurrencies

According to the Forbes advisor, cryptocurrency is a decentralized digital money that is based on blockchain technology (Forbes, 2020). Digital in this case implies that the currency is virtual and is secured by cryptography in which the information is recorded and shared in form of codes which only the persons for whom the information is intended can only decode. It is decentralized in that it is not under the control of governments and central authorities like the central Bank. According to Euromoney learning, a blockchain is a digital ledger that is duplicated and distributed across the entire network of computer systems on the blockchain.[1] The system records the information in a way that is difficult to change, hack or cheat the system. The first and most popular cryptocurrency based on the blockchain technology was Bitcoin, launched in 2009. According to Statista's statistics, there were over 4000 cryptocurrencies as of 2021 which represents a huge increase from the few digital coins that existed as of 2013.[2]

The demand and interest in cryptocurrencies are increasing. There is thus a need for them to be treated as more than just a novelty. Some cryptocurrencies have seen more growth than others. There is much discussion as whether cryptocurrencies should be classed as currencies, assets or investment vehicles, and this is a key topic in itself (Chu, Chan, Nadarajah, & Osterrieder, 2017). The gap exists where the volatility of cryptocurrencies is important to study especially about financial investments like pricing or hedging instruments.

Blockchain and cryptocurrencies have found a wide spectrum of application scenarios in various types of industries, ranging from the underlying techniques of data storage, encryption, and verification, to the middle level of finance and asset management, and to a variety of high-level business models. Blockchain can be considered as the next generation of cloud computing and is expected to radically reshape the behaviour model of individuals and organizations, and thus realize the transition from the Internet of Information today to the future Internet of Value (Yuan & Wang, 2018).

---

[1] https://www.euromoney.com/learning/blockchain-explained/what-is-blockchain
[2] https://www.statista.com/statistics/863917/number-crypto-coins-tokens/

People who deal with cryptocurrencies, especially investors, require understanding the distribution of the cryptocurrencies and the statistical properties which then helps in understanding the risk involved and the risk they will experience in the portfolio. They always put this into consideration while finding the best ratios with which to invest in different financial assets and in this case, cryptocurrencies. This is known as portfolio optimization. A good optimization entails selecting a good distribution that fits the cryptocurrency data.

In Kenya, there is a known restaurant in Nyeri County that offers customers the option of paying their bills using cryptocurrencies and in particular, Bitcoin. The restaurant has set a minimum trading volume of Sh100. Clients use a blockchain wallet which allows sending and receiving the Bitcoin.[3] The Central Bank of Kenya is however not impressed with the Cryptocurrency introduction into the country. Citing Security concerns, it blacklisted the currencies in 2018 and warned against dealing in them. A few other establishments such as an ATM in Nairobi and a lounge in Kenyatta University are said to trade in Bitcoin. As at May 2021, the 4 most popular Cryptocurrencies in Kenya were Bitcoin, Ethereum, Ripple and XRP.[4]

Several studies have been done on the statistical properties of the cryptocurrencies. (Chan, Chu, Nadarajah, & Osterrieder, 2017) Analyse the statistical properties of the largest cryptocurrencies, determined by market capitalization. They also explore the exchange rate of Bitcoin versus the US Dollar using fifteen of the most popular parametric distributions in finance, the most comprehensive collection of distributions ever fitted to any exchange rate data. The statistical distributions are not exhaustive and especially those that are used to fit fat tailed data. It is noted in the literature review that the Cryptocurrencies data exhibit fat tails. Finance data which includes cryptocurrencies have many outliers. This means that the extreme observations are significantly more extreme than would be expected for normal distributions. Kurtosis is sensitive to tail weight and therefore, high kurtosis is nearly synonymous with having a heavy tailed distribution (David & David, 2015).

We use returns in this case. Returns are used for the purpose of normalization. They allow measuring of all variables in a comparable metric. Despite unequal or very

[3] https://www.standardmedia.co.ke/entertainment/local-news/2001287484/betty-s-bitcoin-nyeri-restaurant-accepts-cryptocurrency-payments-despite-cbk-warning
[4] https://citizentv.co.ke/news/4-popular-cryptocurrencies-in-kenya-11080780/

different prices, the returns allow for analysis of the relationship between such variables. The returns desired in this study are the log returns.

Log Returns

The log return is also known as the continuously compounded return. The advantage of using log returns when analysing data is that the relative changes in the variables become noticeable and comparable to other variables with different base values. The properties that are common for financial log-return series are the sample mean of the series is close to zero and the marginal distribution is roughly symmetric. It also has a peak at zero and is heavy tailed. The sample autocorrelations of the series are small at almost all lags, but the sample autocorrelations of the absolute values and squares of the series are significant for many lags. Volatility is clustered. This implies that days of either large or small movements are followed by days of similar characteristics.

Statistical Distributions

A distribution may be thought of as a function that describes the relationship between observations in a sample space. Most available data conform to well-known and well understood mathematical functions. With a modification of the parameters of a function, it can fit data. Good examples of parameters would be the mean and the standard deviation in the case of the Gaussian distribution. Once a distribution function is known, it can be used as a shorthand for describing and calculating related quantities, such as likelihoods of observations, and plotting the relationship between observations in the domain.[5]

The student's t distribution and the Generalized hyperbolic distribution are the most common distributions to be fitted to cryptocurrencies data. The normal inverse Gaussian, generalized t and Laplace distributions have also been fitted. These distributions are mainly fitted because they are heavy tailed and heavy tails are common in financial data. (Chan, Chu, Nadarajah, & Osterrieder, 2017) find that there is no one best fitting distribution jointly for the sample of cryptocurrencies that they use. They suggest the generalized hyperbolic distribution for Bitcoin, the normal inverse Gaussian distribution for Dash, Monero and Ripple and the Generalized t and Laplace distributions for Dogecoin and MaidSafeCoin respectively. They also point out the surprise that the Laplace distribution which is a light tailed distribution gives

[5] https://machinelearningmastery.com/statistical-data-distributions/

the best fit for the return of a cryptocurrency data. They thus conclude that distribution of tails of cryptocurrencies range from light tailed via semi-heavy tailed to heavy tailed. Given that financial instruments have heavy tails, this is surprising.

The distributions are however not exhaustive and there remains more distributions that would fit the cryptocurrencies data. This calls for further research by fitting more distributions that have been fitted before to data with the similar characteristics to those of the cryptocurrencies data, and especially the heavy tail.

Problem Statement

Investors who deal with Financial Portfolios that include cryptocurrencies require to understand the distribution of the cryptocurrencies and the statistical properties that imply the risk properties. This is in finding the best ratios in the optimization. A good optimization will entail selecting a good distribution that fits the cryptocurrency data. This is because investors always consider the risk and the returns. They look forward to the lowest possible risk for the highest possible returns. If data follows a certain distribution, it aids in knowing the returns to expect and the risk attached. Cryptocurrencies especially exhibit very high risk from studies already done, and it is important to understand the pattern of the data before considering it in the portfolio. It may also help to determine the weight or portion of the investment given to the cryptocurrencies in the portfolio.

A huge number of the empirical studies have been done on Financial Stocks. This is especially on the indices like the New York Stocks Exchange Composite Index and company Stocks like Boeing, Citigroup and General Motors. For high frequencies, the returns of the Financial Stocks data studied was not normal enough to accept the Gaussian distribution to model it. The returns have heavy tails and so the generalised Pareto distribution is more acceptable to capture the greater and more frequent extremes that the returns show (Quigley, 2008).

This paper seeks to fit fat tailed distributions from literature to the cryptocurrencies data of 6 best ranked cryptocurrencies data by market capitalization which will help estimate the risk and returns thus making informed decision in portfolio optimization.

Research Objective

This study seeks to find out whether the log returns of the sampled cryptocurrencies data will bring out certain properties of the stated statistical distributions. The specific objectives are:

1. To fit fat tailed distributions to sampled cryptocurrencies data.
2. To determine the best fit distribution of the log returns of the cryptocurrencies.

Research Questions

1. Which of the fat-tailed distributions best fit the log returns of the different cryptocurrencies?
2. How well does the distribution fit the cryptocurrency data?

Significance of the Study

This study would be useful especially to portfolio and risk managers. This would especially be those who are risk seekers and are looking for a way to invest or enter the technology market. Financial institutions would also be major users of the information. Banks have been traditionally heavily regulated in the lending space and have used complex infrastructure for simple transactions, like sending money to borrowers' bank account or debit cards. Blockchain technology could allow financial institutions to create direct links between each other on a frictionless peer to peer basis and avoid correspondent banking. This would improve automation (Forbes, 2019). Blockchain is safe, secure, decentralized, and reduces the chances of human error or malicious fraud. Some IT specialists believe future blockchains will become practically un-hackable, making it a very attractive option for financial institutions, which must keep customers' money and personal information safe. Fewer data breaches boost customer confidence and keeps banks in line with their regulators.

Most users are trading them for investment purposes: either as a long-term investment in new technology or looking to make a short-term profit. Investigating the volatility of cryptocurrencies is important in terms of financial investment like hedging or pricing instruments. Bitcoin possesses some of the same hedging abilities as gold and can be included in the variety of tools available to market analysts to hedge market specific risk.$

CHAPTER 2 : LITERATURE REVIEW

Introduction

In this chapter, the study outlines the key theoretical issues in my study which include the Statistical analysis of cryptocurrencies, Heavy tailed data in relation to cryptocurrencies and the log returns of Financial Assets. The chapter concludes with an empirical literature review on the distributions suggested for study.

Theoretical Literature

Statistical Analysis of Cryptocurrencies

(Chan, Chu, Nadarajah, & Osterrieder, 2017) Analyse the statistical properties of the largest cryptocurrencies, determined by market capitalization, of which Bitcoin is the most prominent example and characterize their exchange rates versus the U.S. Dollar by fitting parametric distributions to them. They use seven cryptocurrencies and having briefly examined the summary statistics for both the exchange rates and the log returns of exchange rates of the cryptocurrencies, they provide a visual representation of the distribution of the log returns. They use the method of maximum likelihood. It is shown that returns are clearly non-normal, however, no single distribution fits well to all the cryptocurrencies analysed. They find that for the most popular currencies, such as Bitcoin and Litecoin, the generalized hyperbolic distribution gives the best fit, while for the smaller cryptocurrencies the normal inverse Gaussian distribution, generalized t distribution, and Laplace distribution give good fits.

(Chu, Nadarajah, & Chan, 2015) Explore the exchange rate of Bitcoin versus the US Dollar using fifteen of the most popular parametric distributions in finance, the most comprehensive collection of distributions ever fitted to any exchange rate data. Again, the method of maximim likelihood is used. These distributions are the normal, Student t, logistic, Laplace, exponential power, skew normal, skew t, generalized t, skewed exponential power, asymmetric exponential power, skewed Student's t, asymmetric Student's t, normal inverse gamma, hyperbolic and generalized hyperbolic distributions. They include heavy tailed and light tailed distributions. The generalized hyperbolic distribution gives the best fit, as assessed by the log-likelihood value, AIC value, AICc value, BIC value, HQC value, CAIC value, probability plot and the density plot.

Heavy Tailed Data

In certain studies, it has been realized that some indices like the S&P 500 do not fall into a normal distribution curve. It is however also true that they fit a normal distribution better than bitcoin. The reasons might be that bitcoin represents a different asset class and follows different assumptions from the broader market. Second, bitcoin, at this time, is relatively an immature market and will be tamed by institutional investment. Or third, the Black-Scholes model is becoming less reliable overall and volatility, as well as unpredictable distributions of such volatility, will become the 'new normal' (Swift, 2019).

The current stand is that cryptocurrencies exhibit heavy tails. In Finance, stock returns, interest rates changes, foreign exchange rates changes and other data like cryptocurrencies have many outliers than would be under normality. They thus use heavy – tailed distributions for modelling. This means that the extreme observations are significantly more extreme than would be expected for normal distributions. Kurtosis is sensitive to tail weight and therefore, high kurtosis is nearly synonymous with having a heavy tailed distribution (David & David, 2015). Heavy-tailed distributions are important models in finance because returns and other changes in market prices usually have heavy tails.

(Markus & Christian, 2007) discover that compared to the normal distribution, the distribution of the S&P500 index returns is fat-tailed. This means that the probability of large losses and gains is higher than would be observed in a Gaussian distribution. It would thus not be advisable to use this for describing the ups and downs of activities in Financial Markets. The history of heavy tailed distributions in finance originates in a model known as the alpha stable model proposed by (Mandelbrot, 1963). It is the first alternative to the Gaussian Law. It has attractive theoretical properties such as the stability property and domains of attraction. While the fat–tailedness of the alpha stable distributions makes it already an attractive candidate for modelling financial returns, the concept of the domains of attraction provides a further argument for its use in finance, as under the relaxation of the assumption of a finite variance of the continuously arriving return innovations the resulting return distribution at lower frequencies is generally an alpha stable distribution (Markus & Christian, 2007). The t distribution might be regarded as the strongest competitor to the alpha stable distribution, shedding also more light on the empirical tail behaviour of returns.

Heavy tails in cryptocurrencies

(Osterrieder, 2017) concludes that cryptocurrencies exhibit heavy tails. This is after fitting both the normal distributions and a set of heavy tailed distributions to 7 of the top 15 cryptocurrencies in 2017. For the heavy tailed fit, the results are similar and he opted to choose the student t distribution when fitting exchange rate returns of cryptocurrencies.

(Linda, 2018) agrees with (Osterrieder, 2017) by stating that the returns of cryptocurrencies are extremely volatile, exhibiting heavier tail behaviour than traditional fiat currencies. Negative skewness does not imply negative returns, as a significant area left of the mean can still be positive. The paper focuses on four cryptocurrencies, Bitcoin, Ripple, Ethereum and Litecoin. The kurtosis values make it clear that all four cryptocurrencies come from distributions with many extreme values compared to the Normal Distribution. (Linda, 2018) summarizes that the behaviour of the log returns appeared to be best described by Generalized Hyperbolic Distributions.

(Batista, 2018) attempts to determine whether a power law like the empirical power law for securities in many different countries holds for cryptocurrencies by estimating the distribution tails of their returns. He concludes that not only do they have fat-tailed distributions but also that the degree of heavy-tail is relatively high in comparison to an average U.S. based large-cap company.

Log Returns of Financial Assets

Log returns have some more favourable properties for statistical analysis than the simple net returns. They are called log returns because they are the difference between the natural log of the asset price at time t and the natural log of its price at the previous step in time (Quigley, 2008). This is calculated as shown below,

$$r_t = \ln(1 + R_t)$$

$$= \ln\left(\frac{S_t}{S_{t-1}}\right)$$

$$= lnS_t - lnS_{t-1}$$

The statistical properties of log returns are better behaved than simple returns. They are the most popular form of returns to be studied when investigating financial assets. The log return of a long period of time is the sum of the log returns of the shorter periods within the long period. The log return over a year is the sum of the daily log returns in the year. The log return over an hour is the sum of the minute log returns within the hour (Pat, 2012). Miskolczi (2017) concludes that the logarithmic return has an advantage against the simple return, namely that the multi-period logarithmic return can be calculated as a sum of the one-period logarithmic returns, while the multi-period simple return is the product of the one-period simple returns, which can lead to computational problems for values close to zero. He also notes that that the simple return of a portfolio is the sum of the weighted simple returns of the constituents of the considered portfolio. In contrast, the logarithmic return can only be approximated by the sum of the weighted logarithmic returns of the constituents of the considered portfolio.

Empirical Literature

Several studies have been done on the Statistical properties of Cryptocurrencies to investigate the statistical distribution with a focus on risk management with regards to portfolio optimization.

The Maximum Likelihood Estimator

The MLE is one of the most used estimation methods for finding estimators. It selects values of the model that maximize the function of likelihood. It simply estimated the parameters of a model. It works on the basis that t for each different data point, there is a unique probability for that data point to have occurred. Different parameters are chosen and ideally, different probabilities are expected for each data point. This means a different likelihood for every parameter. MLE chooses the parameter that would result in the highest value of the likelihood probability. It is used over other methods of estimation like the least square regression and generalized method of moments because it is the most efficient estimator, its results are unbiased in larger samples and it provides consistent but flexible approach which suits a variety of applications even when assumptions are violated (Erica, 2020).

Fitted Distributions

Other studies on cryptocurrencies have tried fitting distributions and some of the well fitting distributions have been the Generalized Hyperbolic Distribution, the Laplace Distribution and the Normal Inverse Gaussian Distribution (Chan, Chu, Nadarajah, & Osterrieder, 2017). The student t distribution is also a common distribution that has been fitted (Linda, 2018). This paper seeks to fit the Generalized Pareto Distribution, the Student t distribution, the Normal Inverse Gaussian Distribution, the Generalized Hyperbolic Distribution and the Gumbel Distribution.

The Generalized Pareto Distribution

This distribution is defined as:

$$G_{\xi,\beta(u)}(x) = \begin{cases} 1 - \left(1 + \frac{\xi x}{\beta}\right)^{-\frac{1}{\xi}} & \text{if } \xi \neq 0 \\ 1 - \exp\left(\frac{-x}{\beta}\right) & \text{if } \xi = 0 \end{cases}$$

The support of x is $x \geq \frac{-\beta}{\xi}$ if $\xi \geq 0$, and $0 \leq x \leq \frac{-\beta}{\xi}$ when $\xi < 0$

This distribution is largely discussed in the Extreme Value Theory Approach. It is mainly used to model exceedance over a threshold. (J.D.Holmes & W.W.Moriarty, 1999) discusses the generalized Pareto distribution (GPD) and its application to the statistical analysis of extreme wind speeds. Its main advantage is that it makes use of all relevant data on the high wind gusts produced by the storms of interest, not just the annual maxima, and it is not necessary to have a value for every year to carry out the analysis. Wind data is heavy tailed as discussed in (Katz, 2000).

The student's t distribution

The distribution is defined as:

$$f(x) = \frac{K(v)}{\sigma}\left[1 + \frac{(x-\mu)^2}{\sigma^2 v}\right]^{-(1+v)/2}$$

for $-\infty < x < \infty$, $-\infty < \mu < \infty$, $\sigma > 0$ and $v > 0$, where $K(v) = \sqrt{v}B(v/2,1/2)$ and B $(\cdot,\cdot)$ denotes the beta function defined by

$$B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}dt$$

(Chan, Chu, Nadarajah, & Osterrieder, 2017) use the distribution to fit cryptocurrency data of Bitcoin, Dash, Dogecoin, Litecoin, MaidSafeCoin, Monero and Ripple. The study found out that the distribution failed to fit the data closely based on the Information Criteria.

(Linda, 2018) also fits the distribution to Bitcoin, Ripple, Ethereum and Litecoin and it is not mentioned as the best fit distribution for the examined cryptocurrencies according to the Kolmogorov – Smirnov Statistic.

The Normal Inverse Gaussian distribution

The density is given by:

$$f(x) = \frac{(\gamma/\delta)^{-1/2}\alpha}{\sqrt{2\pi}K_{-1/2}(\delta\gamma)} \exp\,[\beta(x-\mu)][\delta^2 + (x-\mu)^2]^{-1}K_{-1}\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right)$$

(Linda, 2018) fitted the distribution with the expectation that it would fit the data of cryptocurrencies well because it has a tail heaviness and an asymmetry component. She goes ahead to conclude that it fits the Bitcoin, Dash, Dogecoin, Litecoin, MaidSafeCoin, Monero and Ripple data well when the Value at Risk and Expected Shortfall are considered. Only slight discrepancies are seen at the tips of the tails.

In (Chan, Chu, Nadarajah, & Osterrieder, 2017), for 3 out of 7 cryptocurrencies which are Dash, Monero and Ripple, the normal Inverse Gaussian distribution gives the best fit. This is assessed in terms of the Q-Q Plot, P-P plots, the Kolmogorov – Smirnov Statistic, the Anderson – Darling test and the Crammer – Von Mises test.

The Generalized Hyperbolic Distribution

The distribution is defined as:

$$f(x) = \frac{(\gamma/\delta)^{\lambda}\alpha^{1/2-\lambda}}{\sqrt{2\pi}K_{\lambda}(\delta\gamma)} \exp\,[\beta(x - \mu)][\delta^2 + (x-\mu)^2]^{\lambda-1/2}K_{\lambda-1/2}\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right)$$

For,

$-\infty < x < \infty, -\infty < \mu < \infty, -\infty < \lambda < \infty, \delta > 0, \alpha > 0$ and $\beta > 0$, where $\gamma = \sqrt{\alpha^2 - \beta^2}$

The generalized hyperbolic distributions allow a more realistic description of asset returns than the classical normal distribution. It has as subclasses hyperbolic as well as normal inverse Gaussian (NIG) distributions which have recently been proposed as basic ingredients to model price processes.

(Chan, Chu, Nadarajah, & Osterrieder, 2017) explain that this distribution is accommodating to semi heavy tails. It contains several heavy tailed distributions in its own particular cases which makes it popular in Finance. In the conclusion of their work, this distribution was best fitting for Bitcoin, Litecoin, Monero and Ripple.

The Gumbel Distribution

The distribution is defined as:

$$F(x; \beta, \mu)/\sigma = 1 - \left(1 + \left(e^{e^{(x-\mu/\sigma)}} - 1\right)^{\beta}\right)^{-1}$$

$$f(x; \beta, \mu)/\sigma = \frac{\beta}{\sigma} e^{(x-\mu)/\sigma} e^{e^{(x-\mu)/\sigma}} \left(e^{e^{(x-\mu)/\sigma}} - 1\right)^{\beta-1} \left(1 + \left(e^{e^{(x-\mu)/\sigma}} - 1\right)^{\beta}\right)^{-2}$$

$$Q(u) = \mu + \sigma \ln \ln \left\{1 + \left(\frac{u}{1-u}\right)^{1/\beta}\right\}$$

Where $-\infty < x < \infty, -\infty < \mu < \infty, 0 < \beta\sigma < \infty$ and $0 \leq u \leq 1$.

(Cooray, 2009) explains that the Gumbel distribution is useful for modelling and fitting a wide range of data sets that aren't well-suited to other distributions. It uses the maximum likelihood method to estimate parameters.

(Ghimire, 2021) fits the Gumbel distribution to data of 772 cryptocurrency daily returns among other 21 distributions noting that this distribution which belongs to the extreme value distribution family has been staple in early academic research on the distribution of returns of cryptocurrencies topic. The paper notes that the extreme value distributions which includes the gumbel distribution were not particularly well performing. This, in their conclusion, contradicts the assertions of prior research.

Goodness of fit test

Goodness of fit tests indicate whether it is reasonable to assume that a random sample comes from a specific distribution (Ricci, 2005). They go through a process of null and alternative hypotheses.

H_o: The data comes from the distribution stated

H_A: The data does not come from the stated distribution

Histograms and Q-Q plots are used to provide insights of distributions that might be good fit. Visualization alone, however, might not be enough evidence to make conclusions and therefore, the Kolmogorov – Smirnov is used. This test is one of the reserved tests for continuous distributions.

The Kolmogorov – Smirnov test.

It is a test used to decide if a sample comes from a population with a specified distribution. It determines whether the underlying distribution differs from the hypothesized distribution. It is defined as:

$$\mathrm{KS} = \sup_{x} \left| \frac{1}{n} \sum_{i=1}^{n} I\{x_i \leq x\} - \hat{F}(x) \right|$$

where I {·} denotes the indicator function and $\hat{F}(.)$ the maximum likelihood estimate of F(x).

(Linda, 2018) calculates this statistic by comparing the empirical distributions of the log returns to one of the theoretical distributions explained and created by the parameters for the distribution at hand and the random samples. For the best distribution obtained in the paper, the p values of the Kolmogorov – Smirnov test are stated to show the values come from the same distribution.

CHAPTER 3 : METHODOLOGY

Introduction

This chapter explains how the study proposed is to be achieved. It defines the research design, the population and sampling with its techniques, data collection in type, source and methods and the analytic techniques. The procedures followed in carrying out the empirical analysis are discussed.

Research Design

This study will adopt a descriptive research design. This means that it will focus on finding the "what" of the research subject. It will be quantitative research. This is because it will collect quantifiable information for statistical analysis of the sample. The log returns of cryptocurrencies are the quantifiable information thus the quantitative and descriptive research design. The study will focus on fitting statistical distribution to the data from a set of fat tailed distributions and discriminate them to find the best fit.

Population and Sampling

The study intends to use daily adjusted closing prices data of 6 highly ranked cryptocurrencies by market capitalization provided adequate data is available as listed in the Yahoo Finance site where the data is to be obtained. The rank of cryptocurrency is normally evaluated by its market capitalization. Market capitalization ranks the relative size of a cryptocurrency by multiplying the spot price by the circulating supply which approximates the coins circulating in the market (Linda, 2018). Another factor in choosing is the availability of adequate data. The 6 cryptocurrencies are Bitcoin(BTC), Ethereum(ETH), Tether(USDT), Dogecoin(DOGE), XRP(XRP) and Litecoin(LTC-USD). The sample data used consists of daily time series from January 2016 to December 2020. All the sampled cryptocurrencies were already in existence by January 2016 making it the best start time of analysis. The end time of analysis is set to be December 2020 to allow for use of 5 complete years of analysis.

Data Collection

The data to be used is secondary data. The log returns of the cryptocurrencies will be calculated from the daily prices of the cryptocurrencies. The log returns will be calculated using Microsoft Excel. This is by the function

$$R_i = \ln\left(\frac{S_t}{S_{t-1}}\right)$$

Where:

R_i is the log returns of the cryptocurrencies

S_t is the price of the cryptocurrency today

S_{t-1} is the price of the cryptocurrency yesterday

The resulting figure is the log return. The data will be obtained from yahoo finance which is a credible source and thus the data may be regarded as reliable and valid.

Data Analysis

To meet the research objectives, this study seeks to find a statistical distribution that fits the log returns data of the cryptocurrencies. Fitting distributions consists in finding a mathematical function which represents in a good way a statistical variable (Ricci, 2005). There are four steps outlined in fitting distributions. These are hypothesizing families of distributions to use, estimating the parameters, evaluating the quality of fit and carrying out statistical tests for the goodness of fit (Ricci, 2005). To achieve the research objectives, this study will follow the steps to analyse the cryptocurrencies data.

The data analysis will be done using both Microsoft Excel and R Studio program. The first step will be a brief description of the data to depict statistics like the skewness, variance and mean. The parameters will be estimated by Maximum Likelihood Estimation. This method is used over other methods of estimation like the least square regression and generalized method of moments because of its efficiency as an estimator. It has unbiased results in larger samples and provides consistent but flexible approach which suits a variety of applications even when assumptions are violated (Erica, 2020). The distributions described in the distributions fitted section will then be fitted, then the best fitting distribution will be selected depending on the goodness of fit tests also described. The theoretical framework applied in the data analysis is described after the Data description.

Data Description

The descriptive statistics of the data gives an overview which allows a simple interpretation of the data and interaction with the data beforehand. The table shows the descriptive statistics of the sample cryptocurrencies data that will be used in the study.

The skewness is positive except for the Tether which is negatively skewed. This shows that the data would be better modelled using fat tailed distributions. The high kurtosis shown by most of the cryptocurrencies also shows that the tails of the data greatly differ from the tails of a normal distribution.

Table 1: Cryptocurrencies Sample Adjusted Closing Price Data 2016 to 2020 Descriptive Statistics

	Ethereum	*XRP*	*Dogecoin*	*Tether*	*Bitcoin*	*Litecoin*
Mean	241.0263	0.2890	0.0023	1.0011	6122.3059	57.0920
Standard Error	5.4177	0.0077	0.0000	0.0002	110.6050	1.2595
Median	194.9090	0.2461	0.0024	1.0003	6398.5400	48.2672
Mode	241.0263	0.2890	0.0002	1.0000	6122.3059	57.0920
Std Deviation	231.5722	0.3280	0.0019	0.0095	4727.6350	53.8368
Variance	53625.69	0.1076	0.0000	0.0001	22350532.9	2898.401
Kurtosis	3.0461	23.205	8.0382	28.2630	1.4221	4.7749
Skewness	1.5708	3.7518	1.8693	-2.7579	0.8835	1.8474
Range	1395.482	3.3727	0.0170	0.1643	28637.3897	355.3392
Minimum	0.9371	0.0051	0.0001	0.9136	364.3310	2.9968
Maximum	1396.420	3.3778	0.0171	1.0779	29001.7207	358.3360
Sum	440354.9	527.93	4.1967	1829.0721	11185452.9	104307.0
Count	1827	1827	1827	1827	1827	1827

Table 2: Cryptocurrencies Sample Log Returns Data 2016 to 2020 Descriptive Statistics

	Ethereum	*XRP*	*Dogecoin*	*Tether*	*Bitcoin*	*Litecoin*
Mean	0.0036	0.0020	0.0019	0.0000	0.0023	0.0020
Standard Error	0.0015	0.0017	0.0015	0.0001	0.0011	0.0013
Median	0.0006	-0.0023	0.0000	0.0000	0.0021	-0.0003
Mode	0.0000	-0.0106	0.0000	0.0000	0.0000	0.0000
Std Deviation	0.0629	0.0705	0.0628	0.0057	0.0490	0.0570
Variance	0.0040	0.0050	0.0039	0.0000	0.0024	0.0032
Kurtosis	10.8561	39.4969	13.6364	24.8860	56.9926	11.9648
Skewness	-0.1024	2.3310	0.9633	0.4563	-0.3383	0.7128
Range	1.0046	1.6436	1.0112	0.1092	1.2445	0.9605
Minimum	-0.5507	-0.6163	-0.4929	-0.0526	-0.6203	-0.4491
Maximum	0.4539	1.0274	0.5183	0.0566	0.6243	0.5114
Sum	6.6571	3.6087	3.4145	0.0006	4.2013	3.5705
Count	1826	1826	1826	1826	1826	1826

For the log returns, the skewness is also mostly positive except for Ethereum and Bitcoin. This also suggests that the data would be better modelled using fat tailed distributions. This is again supported by the high kurtosis which shows that the tails of the log returns data are different from the tails of a normal distribution.

The Maximum Likelihood Estimation

To fit the distributions described below, the method of Maximum Likelihood is used to find the estimates for the parameters which will be based on the cryptocurrencies data. If $x_1, x_2, \dots, x_n$ is a random sample of observed values of the prices of cryptocurrencies, and if $\Theta = (\theta_1, \theta_2, \dots, \theta_k)$ are parameters specifying the distribution of the cryptocurrencies, then the maximum likelihood estimates of Θ are those maximizing the likelihood

$$L(\Theta) = \prod_{i=1}^{n} f(x_i; \Theta)$$

or

$$\ln L(\Theta) = \sum_{i=1}^{n} \ln f(x_i; \Theta)$$

Distributions Fitted

The Generalized Pareto, Student's t, normal inverse Gaussian, generalized hyperbolic and gumbel distributions will be employed. The properties of the distributions are described below.

The Generalized Pareto Distribution

The Pareto distribution is generalized by adding a scale parameter. The distribution function is given by:

$$F(x) = 1 - \left(\frac{b}{x}\right)^a, x \geq b$$

The mean and variance are therefore:

$$\mathbb{E}(X) = b\frac{a}{a-1} \text{ if } a > 1$$

$$\text{var}(X) = b^2 \frac{a}{(a-1)^2(a-2)} \text{ if } a > 2$$

The student's t distribution

$$f(x) = \frac{K(v)}{\sigma}\left[1 + \frac{(x-\mu)^2}{\sigma^2 v}\right]^{-(1+v)/2}$$

for $-\infty < x < \infty$, $-\infty < \mu < \infty$, $\sigma > 0$ and $v > 0$, where $K(v) = \sqrt{v}B(v/2,1/2)$ and $B(\cdot, \cdot)$ denotes the beta function defined by

$$B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}dt$$

The normal inverse Gaussian distribution

$$f(x) = \frac{\frac{\gamma}{(\delta)^{-\frac{1}{2}}\alpha}}{\sqrt{2\pi}K_{-\frac{1}{2}}(\delta\gamma)} \exp[\beta(x-\mu)]\,[\delta^2 + (x-\mu)^2]^{-1} K_{-1}\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right)$$

For $-\infty < x < \infty, -\infty < \mu < \infty, \delta > 0, \alpha > 0$ and $\beta > 0$, where $\gamma = \sqrt{\alpha^2 - \beta^2}$

The $K_v(.)$ denotes the modified Bessel function of the 2nd order v:

$$K_v(x) = \begin{cases} \frac{\pi xsc(\pi v)}{2}[1_{-v}(x) - I_v(x)], & \text{if } v \notin \mathbb{Z} \\ \lim_{\mu \to v} K_\mu(x), & \text{if } v \in \mathbb{Z} \end{cases}$$

$I_v(.)$ is of the 1st order:

$$I_v(x) = \sum_{k=0}^{\infty} \frac{1}{\Gamma(k+v+1)k!}\left(\frac{x}{2}\right)^{2k+v}$$

The gamma function used is defined by:

$$\Gamma(a) = \int_0^\infty t^{a-1} \exp(-t)\, dt$$

The generalized hyperbolic distribution

$$f(x) = \frac{(\gamma/\delta)^\lambda \alpha^{1/2-\lambda}}{\sqrt{2\pi}K_\lambda(\delta\gamma)} \exp\,[\beta(x - \mu)][\delta^2 + (x-\mu)^2]^{\lambda-1/2} K_{\lambda-1/2}\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right)$$

For $-\infty < x < \infty, -\infty < \mu < \infty, -\infty < \lambda < \infty, \delta > 0, \alpha > 0$ and $\beta > 0$, where $\gamma = \sqrt{\alpha^2 - \beta^2}$

The Gumbel distribution

$$F(x; \beta, \mu)/\sigma = 1 - \left(1 + \left(e^{e^{(x-\mu/\sigma)}} - 1\right)^\beta\right)^{-1}$$

$$f(x; \beta, \mu)/\sigma = \frac{\beta}{\sigma} e^{(x-\mu)/\sigma} e^{e^{(x-\mu)/\sigma}} \left(e^{e^{(x-\mu)/\sigma}} - 1\right)^{\beta-1} \left(1 + \left(e^{e^{(x-\mu)/\sigma}} - 1\right)^\beta\right)^{-2}$$

$$Q(u) = \mu + \sigma \ln\, \ln \left\{1 + \left(\frac{u}{1-u}\right)^{1/\beta}\right\}$$

Where $-\infty < x < \infty, -\infty < \mu < \infty, 0 < \beta\sigma < \infty$ and $0 \leq u \leq 1$

The Gumbel distribution's log likelihood function is as follows

$$l(\boldsymbol{\theta}) = n\ln\beta - n\ln\sigma - n\frac{\mu}{\sigma} + \sum_{j=1}^{n}\frac{x_j}{\sigma} + \beta\sum_{j=1}^{n} e^{(x_j-\mu)/\sigma}$$
$$+(\beta-1)\sum_{j=1}^{n}\ln\left(1 - e^{-e^{(x_j-\mu)/\sigma}}\right) - 2\sum_{j=1}^{n}\ln\left(1 + \left(e^{e^{(x_j-\mu)/\sigma}} - 1\right)^{\beta}\right)$$

Model Discrimination

The Kolmogorov – Smirnov statistic will be used to evaluate the significance of the fit by use of the p-value. The null and alternative hypotheses are:

H_o: The data comes from the distribution stated

H_A: The data does not come from the stated distribution

At the 5% significance level, the p-value will inform the decision whether to reject or not reject the null hypothesis.

The Kolmogorov-Smirnov statistic

$$\mathrm{KS} = \sup_{x}\left|\frac{1}{n}\sum_{i=1}^{n} I\{x_i \le x\} - \hat{F}(x)\right|$$

where I $\{\cdot\}$ denotes the indicator function and $\hat{F}(.)$ the maximum likelihood estimate of F(x);

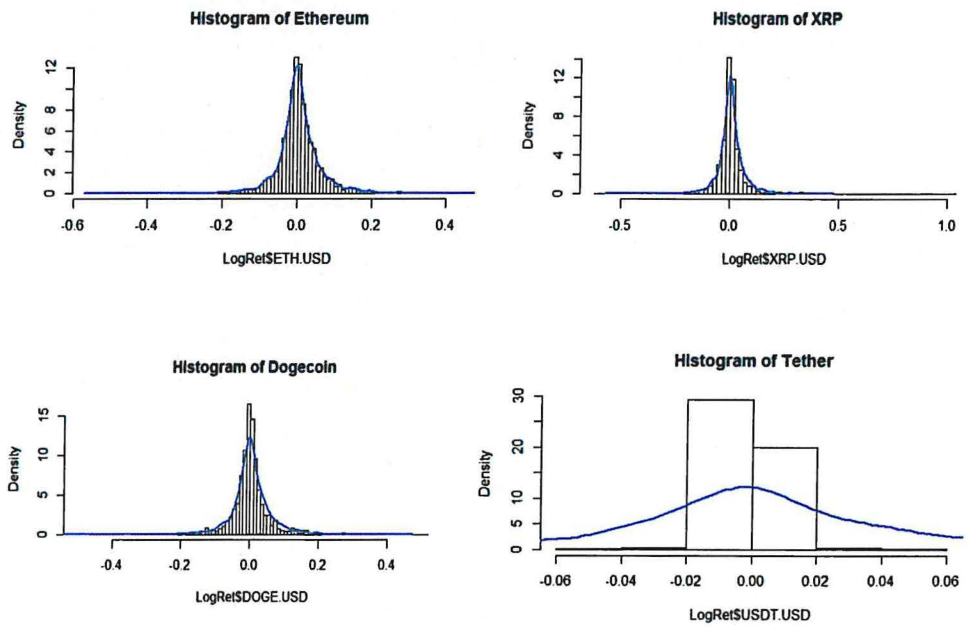
CHAPTER 4 : DATA ANALYSIS

Introduction

This study used Cryptocurrency data of 6 cryptocurrencies which are Bitcoin (BTC), Ethereum (ETH), Tether (USDT), Dogecoin (DOGE), XRP (XRP) and Litecoin (LTC-USD). This data was readily available from Yahoo Finance where historical data could be downloaded for all the cryptocurrencies for the desired period of study. This data was obtained in Microsoft Excel format and the adjusted closing prices extracted therein. From the adjusted closing prices, the log returns of the cryptocurrencies were obtained. The data is from 1st January 2016 to 31st December 2020. The log returns obtained are thus from 2nd January 2016 to 31st December 2020. The log returns of the cryptocurrencies was analysed separately in pursuit to obtain the best model for each cryptocurrency. The data was analysed using R software.

Data Visualization

Figure 1 shows the histograms of the daily log returns of the 6 cryptocurrencies.



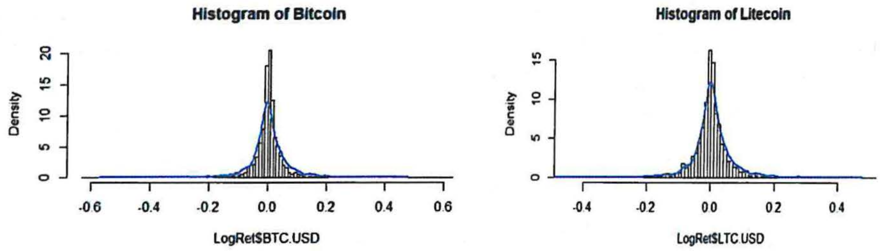


Figure 1: Histogram representation of the log returns of the cryptocurrencies

Normality Test

From the Shapiro Wilk method for normality test, the p-values are insignificant showing that the sample data does not come from a normal distribution. The p- values are all less than 2.2e-16.

The curve imposed on the Histogram shows that there is a wide range of values in the tails of the histograms. In all the histograms, the range of data which are outliers is on both sides. The tails are thus heavier with more bulk showing that occurrence of extreme observations is higher than it would be in a normal distribution. This is in parallel with the observed kurtosis figures which are excessively positive.

Parameter Estimation

The parameters of the distributions were obtained using the Maximum Likelihood Estimate method. Table 3 below shows the estimated parameters and the Maximum Log Likelihood function. The highest likelihood function shows the most suitable distribution.

Table 3: Estimated Parameters for fitted distributions

Distribution	Parameter	Ethereum	XRP	Dogecoin	Tether	Bitcoin	Litecoin
Generalized Pareto	Xi	0.1440	0.0465	0.3325	0.4040	0.2067	0.2300
	Beta	0.0366	0.0015	0.0274	0.0021	0.0205	0.0292
	LLF	2012.77	1925.95	1977.66	3535.95	2720.63	2072.24
Student t's	Sigma	0.0312	-0.0229	0.0233	0.1630	0.0183	0.0263
	DF	2.1193	1.5809	1.5798	3.0739	1.7631	1.9262
	Mu	0.0010	0.0021	0.0012	0.0109	0.0027	0.0003
	LLF	2810.28	3023.57	3005.49	7792.86	3589.13	3025.88
Normal Inverse Gaussian	Mu	-0.0012	-0.0031	-0.0020	-0.0154	0.0027	-0.0014
	Delta	0.0313	0.0218	0.0227	0.0574	0.0173	0.0258
	Alpha	7.9829	4.9151	5.5143	7.2306	8.1320	7.7641
	Beta	1.2181	1.1176	0.9309	1.4781	-0.1855	1.0056
	LLF	2824.37	3054.84	3031.70	7807.52	3601.70	3044.54
Generalized Hyperbolic	Lambda	-0.0847	-0.4418	-0.2119	1.3466	-0.0248	0.2478
	Alpha	0.2522	0.1107	0.1390	0.0003	0.1353	0.1345
	Mu	-0.0018	-0.0031	-0.0020	0.00003	0.0025	-0.0015
	Sigma	0.0615	0.0668	0.0623	0.4676	0.0444	0.0553
	Gamma	0.0053	0.0051	0.0038	0.2517	-0.0004	0.0034
	LLF	2826.56	3055.01	3035.69	10800.23	3607.16	3054.7
Gumbel Distribution	Sigma	0.1027	0.1155	0.0916	0.0102	0.1159	0.0847
	Mu	-0.0025	-0.0321	-0.0284	-0.0030	-0.0270	-0.0258
	LLF	-1753.1	-1577.5	-1934.25	-6013.3	-1651.0	-2083.9

For easier comparison, table 4 below is an extract of the LLF figures.

Table 4: Extract of the Log Likelihood Function of each distribution and cryptocurrency

Distribution	*Ethereum*	*XRP*	*Dogecoin*	*Tether*	*Bitcoin*	*Litecoin*
GP	2012.77	1925.95	1977.66	3535.95	2720.63	2072.24
Student t's	2810.28	3023.57	3005.49	7792.86	3589.13	3025.88
NIG	2824.37	3054.84	3031.70	7807.52	3601.70	3044.54
GHYP	**2826.56**	**3055.01**	**3035.69**	**10800.23**	**3607.16**	**3054.7**
Gumbel	-1753.1	-1577.5	-1934.25	-6013.3	-1651.0	-2083.9

The bold values are the highest likelihood functions. It is evident that the Generalized Hyperbolic distribution is the most suitable distribution for explaining the cryptocurrencies as it has the highest log likelihood in all cases.

Q-Q Plots

The Q-Q Plots of the best fitting distributions for the 6 cryptocurrencies are shown in figure 2 below.

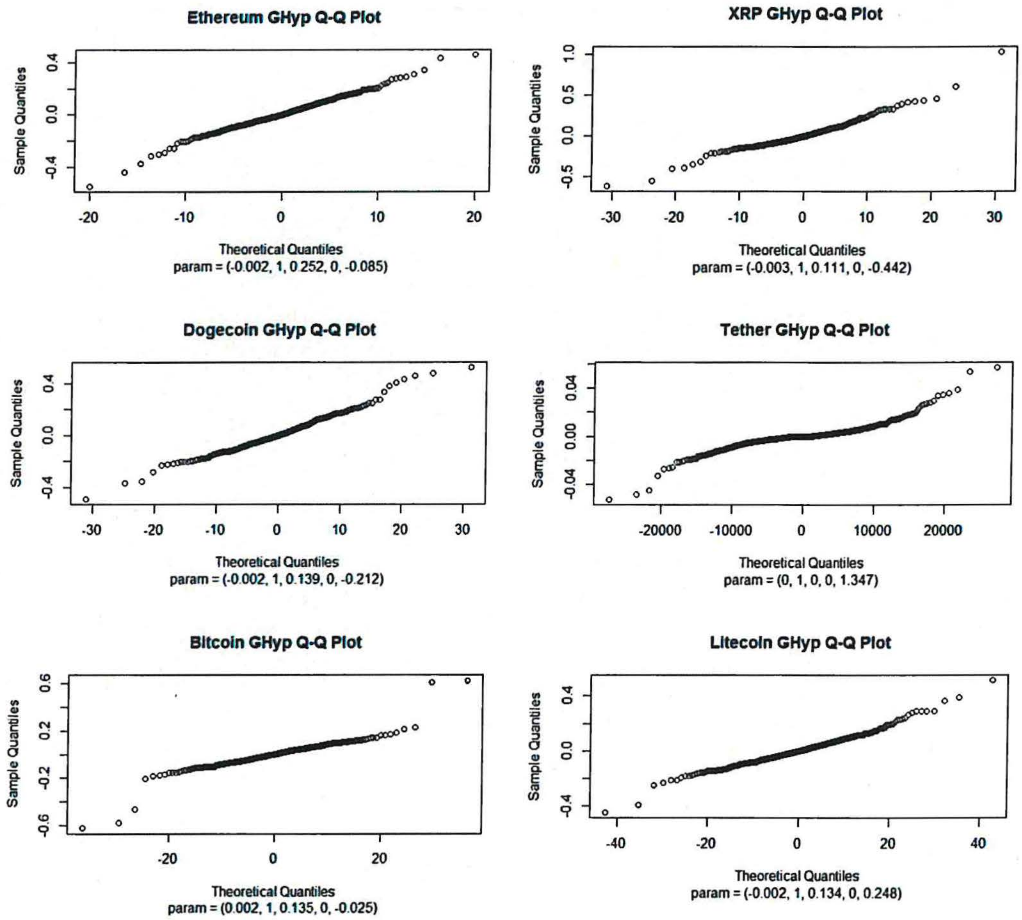


Figure 2: The Q-Q Plots of the best fitting distributions of the log returns of cryptocurrencies

P-P Plots

The P-P Plots of the best fitting distributions for the 6 cryptocurrencies are shown in figure 3 below. The distributions capture the middle, upper and lower parts well.

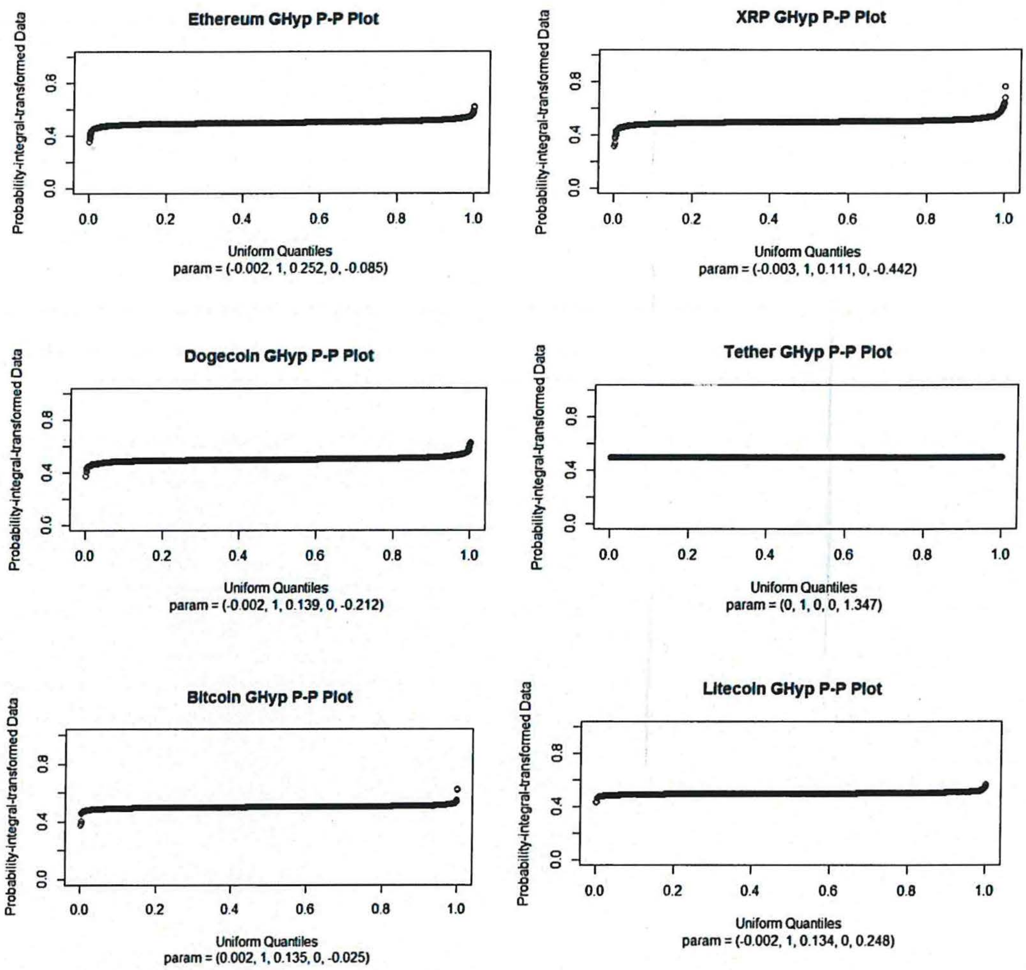


Figure 3: The P-P Plots of the best fitting distributions of the log returns of cryptocurrencies

Goodness of Fit Test

The p-values for the best fitting distributions are represented in the table below.

Table 5: p-values obtained from the test statistics

Cryptocurrency	Kolmogorov-Smirnov Test
Ethereum	0.6940
XRP	0.8935
Dogecoin	0.5391
Tether	< 2.2e-16
Bitcoin	0.6293
Litecoin	0.9999

The figures show that all the best fitting distributions are adequate at the five percent significance level except for Tether.

Summary

The generalized hyperbolic distribution is the best fit for all the log returns data of the cryptocurrencies used. This shows that it is the most appropriate distribution for explaining the behaviour of cryptocurrencies. The p-values obtained from the test statistics show adequacy at the five percent confidence level except for the log returns of Tether. This proposes fitting a different set of distributions in pursuit of one that best fits the log returns of Tether.

Based on past studies, there has not been one distribution that best explains the log returns of cryptocurrencies. This study suggests that the Generalized hyperbolic distribution is the best distribution in explaining the log returns of cryptocurrencies since it fits the sampled cryptocurrencies log returns data.

CHAPTER 5 : CONCLUSION

Introduction

The objective of this research was to find the distribution that best fits the log returns of cryptocurrencies. This helps in estimating the risk and returns since the statistical properties imply the risk properties. This in turn will allow investment managers to make informed decisions in portfolio optimization since demand and interest in cryptocurrencies is rising. Blockchain and cryptocurrencies have found a wide spectrum of application scenarios in various types of industries especially finance and asset management. The study especially fitted heavy tailed distributions since cryptocurrencies exhibit heavy tails as concluded by (Osterrieder, 2017).

The first step entailed hypothesizing families of distributions to use whereby the generalized pareto distribution, the student t's distribution, the normal inverse gaussian distribution and the generalized hyperbolic distribution were selected since they have been used before to explain heavy tailed data. The parameters were then estimated using the Maximum Likelihood Estimation method. This method is more efficient in estimation over others like the least squares regression and the generalized method of moments. The Kolmogorov-Smirnov Statistic was used as the goodness of fit test on the fitted distributions.

Conclusion

This paper studied the log returns of the daily adjusted closing prices of 6 highly ranked cryptocurrencies by market capitalization, which were Ethereum, XRP, Dogecoin, Tether, Bitcoin and Litecoin. The sample data was from January 2016 to December 2020. The finding of the study was that the Generalized hyperbolic distribution is the most suitable since it fitted all cryptocurrencies' data sampled. This decision was informed by the fact that the Generalized Hyperbolic distribution had the Maximum Log Likelihood Function. The p-values obtained showed significance of the fit except for Tether.

Findings of past studies as seen in the Literature review section (Chan, Chu, Nadarajah, & Osterrieder, 2017), show that there is no one distribution that explains all the cryptocurrencies. This study agrees with the past studies since two distributions best fit the small sample of cryptocurrencies considered. This study, however, agrees with (Chu, Nadarajah, & Chan, Statistical Analysis of the Exchange Rate of Bitcoin, 2015), that the Generalized hyperbolic distribution gives the best fit, as assessed by the log likelihood value, since it fits most of the sampled data.

The objectives of the study were achieved since fat tailed distributions were fit to sample cryptocurrencies data and the Generalized hyperbolic distribution was selected as the best fit distribution of the log returns of the cryptocurrencies. This distribution brings out the statistical properties of the log returns and may thus be used to infer the risk properties. This is helpful in optimization of portfolios for investment managers who would like to invest in cryptocurrencies in determining the weights to assign to the cryptocurrencies. It will also make an informed decision whether to invest or not especially depending on the risk tolerance of the investor.

Recommendations

The distribution obtained in the study are suitable for the top cryptocurrencies by market capitalization. For investors that do not seek to invest in this group of cryptocurrencies, it is recommended that they try to fit distributions to the log returns of the desired cryptocurrencies to ascertain their statistical properties and behaviours. This would allow them to know the expected changes in the portfolio. It is however highly probable that the distributions will fall on the normal inverse gaussian or the generalized hyperbolic distributions since they are closely related in terms of their properties. The distributions would be helpful to other investors with similar investment goals.

Limitations of the Study

The distributions fitted were out of the normal and popular distributions normally fitted in studies and thus were laborious to fit. As opposed to distributions that are easily fitted using the "fitdist" packages in R, these distributions required numerous packages. This called for more time used in studying different packages in the R programme.

These different distributions were also difficult to evaluate using different test statistics. This is because there was no inbuilt way of evaluation as seen in the "fitdist" package and thus the distributions are evaluated separately. The study could thus not calculate the statistics for all the distributions fitted but could only calculate for those that best explained the data to obtain the p-value and prove significance. It would have been desirable to use more than one statistic but the methodology to use others like the Anderson Darling and the Crammer von Mises needs to be developed.

The study went ahead to fit the different distributions per cryptocurrencies' data as opposed to a multivariate method where all the data would be fitted to a specific distribution at once. That would be helpful since it would give the distribution for all the cryptocurrencies data and would also be easy to use a large set of data for the work would thus not be repetitive.

Suggestion for Further Research

For people interested in finding the suitable distribution for cryptocurrencies data, they may investigate a joint distribution of the log returns of cryptocurrencies which would be aided by using multivariate processes.

Further research would also be carried out to point out a simple way of using the Kolmogorov-Smirnov, Anderson-Darling and Crammer von Mises statistics on the different fat tailed distributions fitted to cryptocurrencies data.

References

Arfeen, M. A., Pawlikowski, K., McNickle, D., & Willig, A. (2013). *The role of the Weibull distribution in Internet traffic modeling.* Shangai: IEEE.

Batista, Q. (2018, July 2). *Cryptocurrency Returns Are Heavy-Tailed.* Retrieved from Quant Econ Notes: https://notes.quantecon.org/submission/5b3a5a50b9eab00015b89f8b

Camblor, P. M., Carleos, C., & Corral, N. (2014). Cramér-Von Mises Statistic for Repeated Measures. *Revista Colombiana de Estadística*, 45-67.

Chan, S., Chu, J., Nadarajah, S., & Osterrieder, J. (2017). A Statistical Analysis of Cryptocurrencies. *Risk and Financial Management*, 23.

Chatfield, C., & Collins, A. J. (1980). Introduction to Multivariate Analysis. In C. Chatfield, & A. J. Collins, *Introduction to Multivariate Analysis.* Chapman & Hall.

Chu, J., Chan, S., Nadarajah, S., & Osterrieder, J. (2017). GARCH Modelling of Cryptocurrencies. *Risk and Financial Management*, 15.

Chu, J., Nadarajah, S., & Chan, S. (2015). Statistical Analysis of the Exchange Rate of Bitcoin. *Plos One*, 24-25.

Cooray, K. (2009). Generalized Gumbel Distribution. *Journal of Applied Satistics*, 171-179.

David, R., & David, S. M. (2015). *Statistics and Data Analysis for Financial Engineering.* New York: Springer.

Erica. (2020, September 21). *Aptech.com.* Retrieved from Data Analytics Blog: https://www.aptech.com/blog/beginners-guide-to-maximum-likelihood-estimation-in-gauss/

Forbes. (2019, March 7). *Forbes.* Retrieved from How Can Financial Institutions Make Use Of Blockchain?: https://www.forbes.com/sites/darrynpollock/2019/03/07/forget-cryptocurrencies-how-can-financial-institutions-make-use-of-blockchain/?sh=67c868515fff

Forbes. (2020, November). *Forbes Advisor.* Retrieved from Forbes: https://www.forbes.com/advisor/investing/what-is-cryptocurrency/

Fryzlewicz, P. (2005). Modelling and Forecasting Financial Log-Returns as Locally Stationary Wavelet Processes. *Journal of Applied Statistics*, 503-528.

Ghimire, S. S. (2021). *A fitting return to fitting returns: Cryptocurrency Distributions Revisited.* Newcastle: Northumbia University.

J.D.Holmes, & W.W.Moriarty. (1999). Application of the generalized Pareto distribution to extreme value analysis in wind engineering. *Journal of Wind Engineering and Industrial Aerodynamics*, 1-10.

Jäntschi, L., & Bolboacă, S. D. (2018). Computation of Probability Associated with Anderson–Darling Statistic. *Mathematics MPDI*, 88.

Katz, R. W. (2000). Do Weather or Climate Variables and Their Impacts. *Environmental and Societal Impacts Group*, 1-4.

Linda, M. (2018). *Tail Risk in Cryptocurrencies.* Delft: Technische Universiteit Delft.

Mandelbrot, B. (1963). New Methods in Statistical Economics. *Journal of Political Economy*, 421-440.

Markus, H., & Christian, P. (2007). *Financial Economics, Fat - Tailed Distributions.* Munich: German Research Foundation.

Miskolczi, P. (2017). *Note on Simple and Logarithmic Return.* Debrecen: Centre - Print.

Osterrieder, J. (2017). The Statistics of Bitcoin and Cryptocurrencies. *International Conference on Economics, Finance and Statistics*, 285-289.

Pat. (2012, January 23). *Portfolio Probe.* Retrieved from Investment Technology for the 21st Century: https://www.portfolioprobe.com/2012/01/23/the-distribution-of-financial-returns-made-simple/

Quigley, L. (2008). *Statistical Analysis of the Log Returns of Financial Assets.* Eagna Chun Gnimh.

Ricci, V. (2005). *Fitting Distributions with R.* Free Software Foundation.

Swift, C. (2019, September 07). *Kurtosis and Bitcoin: A Quantitative Analysis.* Retrieved from Hackernoon: https://hackernoon.com/kurtosis-and-bitcoin-uc55n3w5x

Telkes, C. (2018). *The Comparison of Risk Measures on Claim Distributions.* Ankara: Middle East Technical University.

Timm, N. H. (2002). Applied Multivariate Analysis. In N. H. Timm, *Applied Multivariate Analysis.* New York: Springer.

Tsay, R. S. (2005). *Analysis of Financial Time Series.* New Jersey: John Wiley & Sons.

Yuan, Y., & Wang, F. Y. (2018). Blockchain and Cryptocurrencies: Model, Techniques and Applications. *IEEE Transactions on Systems* , 8.